Subcontractors for Tractors: Theory and Evidence on Flexible Specialization, Supplier Selection, and Contracting*

Tahir Andrabi
Pomona College
tandrabi@pomona.edu

Maitreesh Ghatak
London School of Economics
m.ghatak@lse.ac.uk

Asim Ijaz Khwaja
Harvard University
akhwaja@ksg.harvard.edu

December, 2002

Abstract

We develop a simple model of flexible specialization under demand uncertainty. A buyer faces multiple suppliers with heterogeneous quality who have the option of selling to other buyers. The more specific a seller's assets are to the buyer, the lower is his flexibility to cater to the outside market and this cost is greater for higher quality suppliers. Therefore even if a buyer typically prefers high quality suppliers, some low quality suppliers might be kept as marginal suppliers because of their greater willingness to invest more in assets specific to the buyer, especially in the presence of contracting problems and high uncertainty. We then examine a primary dataset on contracts between the largest tractor assembler in Pakistan and its suppliers and examine how the extent of asset specificity and other supplier characteristics affect contractual outcomes such as prices and distribution of orders and find evidence that the more dedicated suppliers are indeed of lower quality.

1 Introduction

What are the things that a firm should make in-house and what are the things should it buy from outside? According to the transactions costs and property rights literature an important

*We thank Abhijit Banerjee, George Baker, Robert Gibbons, Oliver Hart, Alexander Karaivanov, Michael Kremer, Rocco Macchiavello, W. Bentley MacLeod, Kaivan Munshi, Canice Prendergast, Tomas Sjöström, Jeffrey Williamson, Chris Udry and several seminar audiences for feedback on a previous version of the paper. We are grateful for all the support and information provided by the Lahore University of Management Sciences and Millat Tractors Ltd. All errors are our own.

determinant governing the boundaries of the firm is the importance of relationship-specific investments.[1]These generate "quasi-rents", and if contracting is costly, the possibility that the parties might behave opportunistically after the investments are sunk to try to get a larger share of these rents. The boundaries of the firm are chosen optimally the minimize the resulting efficiency losses.

However, a question that has received little attention is how relationship-specific investments affect contracts when the boundaries of the firm are given. This question is important given that many industries are characterized neither by vertically integrated firms nor by a set of independent buyers and suppliers but as *networks* where suppliers provide specialized inputs to several buyers selling related but different products and buyers have more than one supplier for the same input. The resulting investment pattern on the part of suppliers has been characterized by flexible specialization which is considered to be an optimal response to demand uncertainty and costly capacity-building or inventory-holding.[2] In such environments the interesting variation is not in the choice of integration by the buyer but in the terms of the contracts faced by a set of sellers who vary in terms of how specific their assets are with respect to the main buyer.

Moreover, the existing literature treats specificity as being driven purely by technology. In a network or cluster setting, given that investment is characterized by flexible specialization, the degree of asset specificity with respect to any particular buyer is also partly a matter of choice. For the same product, different suppliers to the same buyer have assets that vary in terms of how specific or dedicated they are.[3] In this paper we ask the question: what types of suppliers are more likely to invest more in specialized capacity in relation to one particular buyer and what types of suppliers are likely to invest more in flexible capacity? If we allow for the possibility that asset specificity also partly reflects supplier heterogeneity in terms of quality, how does it affect our interpretation of the effect of specificity on dependent variables such as the contracts between a particular buyer and his sellers?

The goal of this paper is take a first step towards addressing these questions by developing a simple theoretical model which analyzes the costs and benefits of undertaking relationship-specific investments, and how these are likely to vary among suppliers with heterogeneous qualities. We

[1]Coase (1937), Klein, Crawford and Alchian (1973), and Williamson (1985) are the classic references in the transactions costs literature. Hart (1995) and Holmström and Roberts (1998) provide excellent reviews of the property rights literature.

[2]See Piore and Sabel (1984) for a discussion of flexible specialization and Kranton and Minehart (2000) for a formal analysis of when such networks are optimal relative to vertical integration.

[3]See Asanuma (1989) for a case study of the Japanese auto-manufacturing industry.

then examine a detailed primary data set of contracts between one large buyer, Millat Tractors Limited (MTL), the largest and oldest tractor-producing firm in Pakistan, and its suppliers who vary in terms of how dedicated their assets are in relation to the former and interpret our findings in terms of the theoretical model. Following the institutional setting, we will call the buyer the "assembler", and the suppliers, "vendors".

Our theoretical model has three key ingredients. First, relationship-specific investments (as opposed to general investments) increase the surplus within the relationship, but lowers the flexibility of a supplier to cater to the outside market which is costly when demand is uncertain. Second, vendors are of different qualities. For simplicity, we assume they are of two types, high and low. A high quality vendor generates a higher level of surplus for the same level of investment. Third, we assume that the marginal cost of undertaking relation-specific investments is higher for a higher quality supplier since such a supplier presumably has a wider range of skills and higher returns from retaining flexibility. This means, for the same level of orders higher quality suppliers invest less than lower quality suppliers. Therefore even if a buyer might typically prefer high quality suppliers, some low quality suppliers might be kept as marginal suppliers because of their greater willingness to invest more in assets specific to the buyer, especially when demand is very uncertain. If in addition, following the Grossman-Hart-Moore (GHM) property rights framework, we allow the investments to be non-contractible and subject to hold-up problems, then the relative attractiveness of low type vendors increases. The outside option is relevant to a vendor not only for the possibility that the demand facing the buyer could be low but also in bargaining over ex post surplus. This makes high types invest even less than low types compared to the first-best, and in addition the assembler has lower bargaining power in dealing with them.

In the empirical section we examine how specificity affects contractual outcomes. MTL manufactures almost none of its required parts in-house and obtains them from local suppliers. For each part, there are multiple suppliers. There is great variation in contracts offered - both the price offered to and quantities ordered from different vendors supplying the same part in the same year differs significantly (on average, a 25% difference in price while quantities ordered differ on average by a ratio of 3). The data set is attractive for several reasons. First, the focus on a single large buyer ensures that the comparison between contracts is meaningful. Second, we were able to get access to detailed contractual outcomes of interest including prices paid to each supplier for a given product (tractor part) and quantities scheduled for each part from the supplier over a decade. Finally, given the assembler often has multiple vendors supplying the same part, we are able to make a cleaner comparison by contrasting contracts between two vendors with different degrees of specificity but which supply the same part. Our comparisons are therefore not confounded with other effects that may be specific to a product yet not related to relationship specificity.

We find that price differences are consistently correlated with vendor characteristics, such as their age, size, distance from MTL, and in particular, prices are lower for vendors with greater relationship-specificity. We also find the level of orders received by some vendors is high and relatively stable (a "first-preference" vendor) while for others it is low and fluctuating (a "second-preference" or marginal vendor).[4] Next we obtain a striking, and surprising finding: vendors whose assets are more specific to MTL receive lower prices and yet are treated as marginal suppliers. We do find that vendors with greater specificity have lower unit production costs, which might explain these vendors receiving lower prices, but other things being the same, the cheaper vendor should be made the first-preference vendor. It turns out that other things are not the same. Cost is not the only consideration of the assembler. It cares a great deal about timely and defect-free delivery. This suggests that quality differences between vendors can explain why MTL does not always buy from the cheapest vendor. Indeed, we find that vendors with greater asset-specificity perform worse both in terms of timely and defect-free delivery. We use our theoretical model to argue that vendor heterogeneity is the key to understand these findings.

Our work is related to the theoretical literature on property-rights. The key distinction is that in our set up only the vendors undertake investments, and so optimal ownership is not the key question. The key question is how can we explain heterogeneity among vendors in terms of extent of relationship-specific investments and, conditional on this, the heterogeneity in the price and the orders received by different vendors. Also, in the Grossman-Hart-Moore framework specific investments have a (weakly) positive marginal effect both on the surplus within the relationship, and the investor's outside option which is not the case in our framework Rajan and Zingales, 1998 analyze this possibility but their focus is how it affects the pattern of optimal ownership. Finally, in this literature specific investment is purely technology driven, and firm heterogeneity and selection issues are under-emphasized.

Our paper is also related to the recent literature on buyer-seller networks. Kranton and Minehart (2000) analyze the choice between vertical integration and networks of suppliers in model with

[4] Asanuma (1989) uses the same terms in his case study of the Japanese auto-industry where he reports a similar hierarchy of subcontractors in terms of distribution of orders. The hierarchy in the Japanese case though, extends to tiers of subcontracting.

flexible specialization, and the strategic investment incentives of individual firms that lead to their formation. Our paper shares with this paper the focus on the advantage of having flexible assets in the presence of demand uncertainty and costliness in maintaining capacity. We focus on the complementary question about how a given assembler would choose its portfolio of vendors between those who are flexible specialists versus those who have dedicated (or tied) assets, and how the price and orders for the same parts will be distributed among them.

Our work is also closely related to the empirical literature on asset specificity and how it affects the nature of contracting.[5] In this literature, the effect of asset specificity is typically shown either on contract duration or on certain contract provisions. For example, Joskow (1987) finds that an increase in the importance of relationship-specific investments lead to parties making longer *ex ante* commitments using data on contracts between coal suppliers and electric utilities in the U.S. Lyons (1994) studies firms involved in subcontracting in the UK engineering industry and reports evidence that the use of formal contracts is positively associated with how specific the investments are and other variables measuring technological complexity and potential opportunism by customers. Gonzalez, Arrunda and Fernandez (1999) examine the determinants of the extent of subcontracting in the construction industry in Spain and report evidence that as specificity grows, firms tend to subcontract less. Woodruff (2002) looks at the Mexican footwear industry and studies the effect of an increase in the importance of specific investments of the retailers (as opposed to footwear manufacturers) on the likelihood of vertical integration. Baker and Hubbard (2001) address the question of how the contractibility of actions affecting the value of an asset affects the pattern of asset ownership by studying the impact of the introduction of on-board computers in the US trucking industry. While we too study the effect of asset specificity on contracts, given the environment we have, the dependent variables we focus on are not the presence of formal contracts (as in Lyons) or their duration (as in Joskow), the decision to subcontract (as in Gonzalez, Arrunda and Fernandez), whether to integrate vertically (Woodruff) or who should own a given asset (as in Baker and Hubbard) but the prices and quantities of orders, and their variability over time and across subcontractors. This way, our work complements this literature. More generally, our work is related to the recent empirical literature on contracting where controlling for unobserved heterogeneity is an important theme (Chiappori and Salanie, 2001).

There is also a large but less formal literature on the Japanese subcontracting system which is somewhat similar to the environment we study. In particular, different suppliers to the main

[5] Chiappori and Salanie (2000) and Shelanski and Klein (1995) provide reviews of this literature.

assembler have differing degrees of dedicated assets; there is intense price competition between multiple vendors and a long term stable vendor pool exists for the auto producers (Asanuma, 1989). This work too does not focus on contractual variations in prices and quantities.

Finally, we view our work as a contribution to the emerging literature on contracting and organizational choice in the industrial sector in developing countries. The presence of significant uncertainty and transactions costs in these economies provide a fertile ground for testing many predictions of the theory of contracts and organizations. While a rich and growing empirical literature on contracting and organizational choice exists in the context of agriculture in developing countries, there is very little work in the context of industry (exceptions include Banerjee and Duflo, 2000, McMillan and Woodruff, 1999, and Banerjee and Munshi, 2000).[6] Banerjee and Duflo study the customized software industry in India and show that the chosen contract as well as contractual outcomes vary with firm characteristics that proxy reputation, especially the age of a firm. McMillan and Woodruff use data from Vietnam and show that inter-firm trade credit is more likely when the delivering firm trusts its client, using several measures of trust including the length of the relationship and information from third parties.

We too have data on the age of the vendors and the length of their relationship with the assembler, but our results do not appear to support a reputation or trust based story as these two papers. Banerjee and Munshi study the knitted garment industry in Tirupur, India. While subcontracting is quite extensive in this case, the authors focus on the question of how efficiently capital is allocated in this industry, as opposed to organizational and contracting issues.

The plan of the paper is as follows. In section 2 we discuss the institutional environment. In section 3 we present the theoretical model. In section 4 we describe the data and the present the empirical results. In section 5 we discuss these results in the light of the theoretical model. Section 6 concludes.

2 Institutional Environment

In this section we briefly discuss the institutional setting to motivate the theoretical model, based mainly on the relationship between MTL and its local suppliers about which we have detailed information, but drawing examples from other case studies wherever possible.

MTL is licensed to produce two models of Massey-Ferguson tractors, the MF-240 and MF-

[6]See Mookherjee (1999) for a recent survey.

375. A national import deletion policy introduced in the late sixties encourages the substitution of imported parts by locally manufactured ones. Almost all of MTL's manufacturing is outsourced to local vendors: it manufactures in-house only 7 out of the 500 tractor parts that are locally produced.[7] However, the main engine components are still imported. MTL has multiple vendors for each part, usually two to three. This feature is the same in the Japanese auto-industry. Asanuma (1989) reports that each buyer seeks to have more than one - typically two or three - suppliers for each part and hold them in parallel. Each vendor typically supplies to several buyers, which includes other tractor assemblers, the market for replacement parts, and automobile manufacturers.

MTL has developed a stable vendor pool with 200 total vendors. There has been very little turnover in these vendors - it has lost only 4 vendors in the last eighteen years. MTL's vendors are based in the two largest cities in the country, Lahore and Karachi and MTL itself is based in Lahore. Karachi is the larger and more industrialized city and hence many of MTL's initial vendors were in Karachi. Our sample consists of 28 randomly selected vendors from this pool. The majority of the vendors supply more than one part to MTL.

An important feature in the environment faced by MTL is the high degree of uncertainty in yearly sales the tractor industry in Pakistan faces. Figure 1 shows that while MTL's annual tractor production during the period under study, 1989-99 was at 10,000 (17000 being total domestic production) there was a high degree of volatility in production. In fact this volatility seems less to be MTL specific but instead faced by the market as a whole (MTL's market share remains fairly stable at around 60% during this period). Moreover, Figure 1 shows that this volatility seems to be specific to the tractor market: While car sales have steadily increased over the years, tractor sales have fluctuated greatly from year to year. There are several reasons for this uncertainty, ranging from erratic government policy (concerning the financing of tractor purchases by government banks, imports, taxes) to demand fluctuations driven by unstable agricultural output growth.

2.1 Vendor Selection and Contracting

A vendor is first selected by MTL after a rigorous examination of its technological capability and track record. The vendor is then typically asked to match a sample part in all its specifications

[7] A medium sized tractor assembled by MTL requires around 900 components, of which 400 are imported. A *component* normally comprises of several *parts* (e.g., a set of bolts required for one wheel is treated as a component by MTL). The vast majority of local vendors produce metalwork parts. In 1994, 54% of the total value of parts procured by MTL was from local vendors. See Amir, Isert and Khan (1995).

and if it is approved, costs are estimated and prices are negotiated. Given the rigorous selection process, uncertainty about a vendor's quality does not appear to be an important consideration. Indeed, most vendors in our sample are quite old, with 80% established in 1980s.

The contract is issued at the time the price of the part is agreed upon. While the price agreed upon for the period of the contract (generally a year) is binding for both parties, there is no commitment on quantity.[8] The fact that contracts tend to relatively incomplete is not unique to this particular case. Asanuma (1989) reports that in the case of the Japanese auto-industry contracts between buyers and suppliers are annual but relatively vague in the sense that it merely states general obligations of the two parties that should be obeyed irrespective of specific transactions, and typically, do not even specify prices.

The vendor only has a rough assurance that it will receive a "reasonable" annual order. The actual supply schedule is determined later based on demand conditions. Such schedules are issued on a quarterly basis. As a result, we use quarterly aggregation for the quantity measures in our empirical findings. Asanuma (1989) reports that in the Japanese auto-industry too, the quantity of a given part which a supplier is asked by the buyer to deliver in the course of the next month varies from month to month, depending on the demand for the car for which the part is needed for. While price is not explicitly made contingent on quantity, it appears that both MTL and the vendors honor their fixed-price contracts despite the large variation they may potentially face in actual quantities bought and sold.

Pricing and order scheduling is carried out for each vendor and often for even each part (unless the parts are natural complements or form a "product-pack"), separately. In other words, price negotiations are sequential rather than simultaneous.

2.2 Nature of Specific Investments

We interviewed MTL engineers and several vendors in our sample who specialize in various types of metal work (forging, casting, and machining) to get a sense of the nature of asset specificity in this environment.[9] According to them specificity takes two forms which, in formal terms, can be

[8]Price renegotiations within the contract period happen rarely unless there is some large unexpected change in input prices (e.g., a hike in raw steel prices).

[9]Forging involves converting metal into desired shapes by heating and then applying pressure. Casting involves pouring molten metal into a mold and allowing it to harden into the shape of familiar engine parts. Machining involves making slots and holes, smoothing the rough edges and providing the required precision in parts that are produced in the casting process.

interpreted as physical capital specificity and human capital specificity.

Physical capital specificity in turn takes two forms, namely, the choice between general-purpose machines and special purpose machines, and the choice of the manufacturing process, e.g., how a machine is "tooled".

Equipments such as dies that are used in forging to make metal parts of specific shapes needed for sections of the body of a particular tractor model, have no alternative uses and so the degree of specificity is not really a choice in these contexts. For some equipments, however, there is a degree of choice of specificity. A multi-drill boring unit can only be used to produce a restricted range of parts, but according to the vendors we interviewed it can increase accuracy by up to 35% as compared to using standard single-drill boring units which are more general-purpose.

The second form of physical capital specificity concerns choosing the way in which equipments are tooled for specific uses. All general-purpose equipment have to be "tooled" in certain ways before they are able to produce a specific part. In machining, tooling is essentially calibrating a machine so that it produces the finished part according to particular specifications. The machines are not computerized and therefore tooling is done manually through trial and error and it takes time for workers to get adjust to new specifications. As a result it is costly to shift the machine to another use. If demand from a particular source is steady and high then it is efficient to incur the fixed costs of tooling for one particular use, because that leads to lower variable costs. Otherwise, it makes sense to choose flexible manufacturing processes that allows them to smaller volumes and larger product mixes given the volatility in demand (both volume and product mix). This involves smaller fixed costs in terms of tooling, but larger variable costs. German and Roth (1997) discuss similar choices facing automotive engine valve suppliers in Argentina who face a more uncertain environment than some of their global competitors. They prefer a process that is more flexible in that tooling and other fixed costs are lower, and so it is easier to shift from one product to another, but variable costs are higher (e.g., it requires more machining work to achieve the desired level of precision). The latter choose processes that require a large investment in tooling (i.e., large fixed costs), but is very efficient for producing large volumes

According to the vendors interviewed, human capital specificity is also important. Vendors have to match the imported sample given to them by MTL and this involves considerable time, effort and learning-by-doing. Moreover, since specialized machine manufacturers do not always exist locally, sometimes the vendor has to develop a specialized machine himself.[10] Skilled labor is

[10]For example, one of the vendors we interviewed, had developed a special press just to create a specific part.

also a major constraint and it takes time to train a machinist to produce specific parts. In addition, because vendors are dependent on MTL for technical and managerial advice, they have to make their operations and organizational setup compatible to that of MTL. This resonates with what Asanuma (1989) calls *relationship-specific skills* in his discussion of subcontractors in the Japanese auto-industry - these are skills that the supplier needs to develop to respond efficiently to the specific needs of a particular buyer.

While the above discussion underlines the significance of relationship-specific investments, for the theoretical analysis an important issue is whether these investments are contractible. There are several reasons why some of these investments are likely to be non-contractible. First, in a developing country such as Pakistan the contracting environment is very poor and few business transactions take place under complete formal contracting. Indeed, a detailed examination of contracts reveals a high degree of incompleteness - the only aspect of the relationship explicitly contracted is price. Second, even if one assumes a good contracting environment, the above description of relation-specific investments suggests that it is hard to verify and hence contract upon some of them. Specifically, while the choice of which type of machine (general or special purpose) to buy can, in principle, be contracted upon, it is much harder to contract upon the choice of manufacturing process involving how a machine is tooled. Moreover, specific investments in human capital are clearly likely to be hard to contract on.

2.3 Quality checks and Performance Monitoring

There are two aspects of a firm's supply of part that MTL is mainly concerned about, quality and timely delivery.

The main quality issue that the Quality Control (QC) division worries about is a part failing to meet measurement specifications. The incorrect measurements result either from vendor error or using equipments beyond their acceptable lifetime. To contain the former, detailed quality checks are carried out when a vendor first starts supplying a part and for the latter, routine quality checks are carried out at regular intervals to ensure that the vendor maintains quality. Additionally, at the end of the yearly contract, the QC division evaluates the past performance of vendors. This is used in future price negotiations, and in deciding the future level of orders.

Our field interviews suggest that ensuring timely delivery of the quantity ordered is a major concern for MTL. Given the high demand uncertainty that MTL faces, it is unwilling to commit on quantity *ex ante*. But the absence of explicit commitment on quantity implies that neither party is in violation of the formal contract even if it does not buy or supply anything. To give incentives to vendors to ensure timely delivery, those who consistently do not meet the required orders are given lower quantity orders in the future, and in extreme cases, discontinued as a supplier.

3 Theory

An assembler faces stochastic demand for his product. It is unable to make some parts in house and needs to outsource. There are many vendors of differing quality available in the market for each part. How would the assembler optimally allocate its orders? What will be the resulting pattern of asset specificity? Will the assembler prefer to have only high quality vendors? If the cost of doing so is that they are less willing to make relationship-specific investments due to uncertain demand, would it like to have some lower quality vendors around who might be more willing to undertake such investments? Who will be chosen to be the first-preference vendor and who will be chosen to be the second preference vendor? Below we provide a simple theoretical model where vendors differ in *ex ante* quality and choose to undertake relationship specific investments and derive the optimal portfolio of the assembler in terms of vendor quality and specific investments, and also distribution of orders.

Suppose that there is a single assembler who buys parts produced by vendors and converts them into final output. For simplicity, we will treat all parts as the same and assume that the assembler needs one unit of an input to produce one unit of output. The assembler faces uncertain demand: with probability $\alpha \in [0, 1]$ demand is high ($= 2$ units) and with probability $1 - \alpha$, demand is low ($= 1$ unit) .

Each vendor has the capacity to produce one unit of the part, but until that capacity constraint is reached, unit costs are constant and do not depend on the level of production. We do not endogenize *why* there are multiple vendors and take a particularly simple form of decreasing returns to scale. An alternative reason could be guarantee against supply bottlenecks.

Vendors are of two types, high or low, i.e., $\theta_i \in \{\underline{\theta}, \bar{\theta}\}$ where $1 \geq \bar{\theta} > \underline{\theta} \geq 0$. Let $\triangle\theta \equiv \bar{\theta} - \underline{\theta}$. We will use the normalization $\bar{\theta} + \underline{\theta} = 1$ which is without loss of generality. All through we assume that the type of vendor i, θ_i, is known and observable to all parties, i.e., there is no learning or adverse selection problem. The type of a vendor affects the quality of its output, as well as the value of its outside option. There are many vendors of both types in the population. Everyone is assumed to be risk-neutral.

At the beginning of the period the parties meet and vendors undertake once-and-for-all investments. Trade takes place at the end of the period after these investments are undertaken. Let $x_i \in [0,1]$ denote the level of relationship-specific investment (henceforth, specific investment) undertaken by vendor i. For example, we can think of x_i as the percentage of its total capacity a vendor keeps tooled up for producing parts specific to MTL on demand and, the rest being kept flexible. The cost of undertaking the specific investment is

$$c(x_i) = \frac{1}{2}x_i^2.$$

The assembler's expected payoff from a unit of the input from vendor i is:

$$V(x_i, \theta_i) = a + b\theta_i + c_0 x_i$$

where $a > 0$, $b > 0$, and $c_0 \in (0,1)$. An increase in the vendor's type increases the value of the input to the assembler by improving quality (e.g., by reducing the likelihood of there being some defect). Also, specific investments improve the expected value of the input to the assembler. In addition, it reduces the cost of producing the input. In particular, the unit cost of producing a part by vendor i is

$$\gamma(x_i) = 1 - c_1 x_i$$

where $c_1 \in (0,1)$. Notice that since $c_1 \in (0,1)$ and $x_i \in [0,1]$, unit costs are always non-negative.

Let us define

$$\gamma \equiv c_0 + c_1$$

as the total marginal return from increasing x. We further assume that $\gamma \in (0,1]$ which, we will see shortly, ensures that $x \in [0,1]$.

Vendor i's outside option is assumed to take the following form:

$$\bar{u}(x_i, \theta_i) = \theta_i(1 - u x_i)$$

where $u \in [0,1)$, which implies that the value of the outside option is always positive, irrespective of the extent of specific investment. This specification of the outside option assumes that the higher is the type of a vendor, the greater is the value of his outside options; the greater is asset-specificity, the lower is the value of outside options; and the marginal loss from specialization is higher for a high type vendor. The first two properties are straightforward. The last property too is intuitive:

the better types are those who have better prospects outside the relationship because they are more *versatile:* they can cater to a diverse range of buyers and can switch quickly from activity to the other. Therefore the marginal cost of constraining themselves to one particular buyer is higher for them. This property is crucial to our analysis - it implies that for the same level of expected orders from the assembler, low type vendors always invest more.[11]

Let

$$S_i = V(x_i, \theta_i) - \gamma(x_i) = a + b\theta - 1 \ + \gamma x_i$$

denote the **gross** *ex post* surplus (i.e., not taking into account the cost of undertaking the specific investment and the opportunity cost of using the asset in terms of passing up the outside option) within a vendor-assembler relationship when the assembler buys one unit of the part from vendor i whose type is θ and whose investment choice is x_i. In order to avoid a tedious multiplication of cases we make the following assumption regarding the parameters that guarantees that *ex post* it is always efficient for the assembler and the vendor to trade, i.e., $S_i - \bar{u}^i(x_i, \theta) \geq 0$ for all $x_i \in [0, 1]$ and $\theta \in \{\underline{\theta}, \bar{\theta}\}$:

Assumption 1

$$a - 1 + b\theta > \theta \text{ for } \theta \in \{\underline{\theta}, \bar{\theta}\}.$$

The left-hand side is of the inequality is gross *ex post* surplus inside the relationship and the right-hand side is the outside option of the vendor when the level of investment is zero. For positive levels of investment the left hand side is larger, and the right hand side is smaller, and so it would be even more profitable to trade.

3.1 The First-Best

We first analyze the case where the investment is contractible and is chosen the maximize the expected joint surplus of the assembler and the vendor. Let $\beta \leq 1$ be the demand faced by a vendor from the assembler, to be chosen endogenously. This can be either a certain demand of β units, or the probability that he is called to supply one whole unit. We will use the latter interpretation. If x is contractible and the assembler plans to buy β units from a vendor of type θ,

[11]This is clearly a consequence of $\frac{d\gamma(x_i)}{dx_i}$ being independent of θ, but would also follow if $\frac{\partial\gamma(x_i)}{\partial x_i}$ did positively depend on θ, but to a lesser extent than $\frac{d\bar{u}^i}{dx_i}$, i.e., so long as $\frac{\partial}{\partial\theta}\left(\frac{d\gamma(x_i)}{dx_i}\right) < \frac{\partial}{\partial\theta}\left(\frac{d\bar{u}^i}{dx_i}\right)$. We consider such an extension at the end of this section.

x will be chosen to maximize *ex ante* expected joint surplus between the vendor and the assembler (ignoring the subscript i):

$$s(x) = \beta\left(V(x,\theta) - \gamma(x)\right) + (1-\beta)\bar{u}(x,\theta) - c(x).$$

which yields the following optimal choice of x:

$$x^*(\beta,\theta) = \max\left\{\beta\gamma - (1-\beta)\theta u, 0\right\}. \tag{1}$$

Since $\gamma < 1$, $x^*(\beta,\theta) < 1$. Let $\hat{\beta}(\theta)$ be the critical value of β such that $x^* = 0$, i.e., $\hat{\beta}(\theta) = \frac{\theta u}{\gamma+\theta u}$. Notice that $\hat{\beta}(\theta) < 1$ and that it is increasing in θ. Our first result follows immediately upon inspection:

Result 1: *Under the first-best:*

(i) *The level of investment is increasing in the level of orders.*

(ii) *The higher is the type of the vendor, the lower is the level of investment for the same level of order.*

(iii) *For the range of orders where both types of vendors invest this gap decreases as the level of orders increases and disappears when the level of orders is* 1.

The first part of the result follows from the fact that the specific investment is useful only when the asset is used to produce for the assembler, and so the level of specific investment is increasing in the level of orders. The second part follows from the fact that the marginal social return from relationship specific investment is lower the higher is a vendor's type. The fact that low type vendors invest more for the same expected order than high type vendors make them potentially attractive. The third part follows from the fact that as the level of orders (β) increases, the outside option gets lower and lower weight, and so the gap between the investment of a high type and a low type vendor shrinks so long as they are both choosing positive levels of investment. When $\beta = 1$ because the outside option gets no weight in the choice of investment, this gap disappears.

Given the equilibrium value of x_i, the equilibrium value of the **net** *ex ante* expected surplus from the relationship is:

$$s^*(\beta,\theta) = \left[\beta\{V(x^*(\beta,\theta),\theta) - \gamma(x^*(\beta,\theta))\} + (1-\beta)\bar{u}(x^*(\beta,\theta),\theta) - \frac{1}{2}\left(x^*(\beta,\theta)\right)^2\right] - \theta.$$

This is the maximum expected surplus that is generated if a vendor of type θ and the assembler decide to enter into a relationship net of the maximum surplus the vendor could have generated by being full time outside the relationship, i.e., $\max \bar{u}(x,\theta) - c(x) = \theta$. Under the first-best, the type of a vendor the assembler would deal with will be chosen to maximize total expected surplus. Recall that by Assumption 1 the expected surplus within the relationship is higher than the outside option without considering the investment. Under the first-best, the level of investment will be chosen to maximize expected surplus taking into account its positive effect on surplus inside the relationship, its negative effect on surplus outside the relationship and the cost of undertaking the investment. As a result, even when we consider the investment the expected surplus within the relationship, is going to be higher than the outside option for both types of the vendors, i.e., $s^*(\beta,\theta) > 0$. But that does not mean both types of vendors will necessarily be chosen. The gain in replacing a lower quality vendor with a higher quality vendor inside the relationship must exceed the loss in replacing a higher quality vendor with a lower quality vendor in the outside sector. To see this in the simplest possible way, suppose there is no investment in the model. Let $a > 1$. Since $b > 0$ by Assumption 1 it is efficient for the assembler to trade with either the high type or the low type. However, for $b < 1$ it is more efficient that the assembler matches with a low type vendor and high type vendors work elsewhere. The total surplus from this allocation would be $(a - 1 + b\underline{\theta}) + \overline{\theta}$ which is greater than that under the alternative allocation, $\left(a - 1 + b\overline{\theta}\right) + \underline{\theta}$. For $b > 1$ by a similar argument, the assembler would match with a high type vendor and for $b = 1$ it is going to be indifferent.

Let

$$s^*(\beta) = \max\{s(\beta,\underline{\theta}), s(\beta,\bar{\theta})\}$$

denote the upper envelope of the net *ex ante* expected surplus from the relationship when the type of vendor of the vendor can be chosen for any given level of order β. Our next result characterizes how should an assembler distribute its orders among the vendors that are chosen:

Result 2: *Given Assumption 1 under the first-best the assembler would give one vendor a certain order of* 1, *and the other vendor an order of* 1 *with probability* α *and* 0 *with probability* $1-\alpha$.

Proof: See the appendix.

We will refer to the vendor with a high and stable order as the “first-preference” vendor and the vendor with a lower and fluctuating order as the “second-preference” vendor. This result shows is that it is in fact optimal to have a “first-preference” and a “second-preference” vendor as opposed to spreading the orders between the vendors in a more equal manner. The intuition behind it follows from the fact that the expected joint surplus within a given vendor-assembler relationship is an increasing and convex function of the level of orders, β. This is depicted in Figure 2. The fact it is increasing is straightforward. An increase in the orders increases the weight attached to the more profitable activity (i.e., trade between the vendor and the assembler) relative to the less profitable activity (i.e., trade between the vendor and the outside market). For the range of β for which we have an interior solution, there is also an indirect effect, as the investment level is also likely to change if the level of orders change. But this effect is exactly zero by the envelope theorem. The marginal gain from increasing orders is exactly equal to the (per unit) *ex post* surplus from trade between the vendor and the assembler (i.e., $S_i - \bar{u}(x_i, \theta)$). For the range of β for which we have a corner solution, this is a constant and so expected joint surplus within the relationship is linear and increasing in orders. For the range of β for which we have an interior solution, this is increasing in x which in turn is increasing in β so that the expected joint surplus within the relationship is strictly convex and increasing in orders for each type of vendor.[12] Since the type specific joint surplus functions are increasing and convex, so must be their upper envelope $s^*(\beta)$ which is the relevant joint surplus function given that the assembler can choose different types of vendors for different ranges of orders. Therefore the assembler should try to place as much order with one vendor as possible, and pass the residual order to another vendor. Since each vendor has a capacity constraint of one unit of output, in equilibrium the assembler buys from more than one vendor - otherwise he would buy everything from one vendor.

The following result characterizes the choice of the type of vendor:

Result 3: *Given Assumption 1 under the first-best:*

(i) For $b \leq 1$ *both the first and second-preference vendors are going to be low types.*

(ii) For $b > \underline{b}$ *where* $\underline{b} > 1$ *both the first and second-preference vendors are going to be high types.*

(iii) For $b \in (1, \underline{b}]$ *the first-preference vendor is going to be a high type vendor and the second preference vendor is going to be a high type vendor if* α *is very low or very high but a low type vendor otherwise.*

Proof: See the appendix.

The key parameter in this characterization is b, which is the marginal return from higher quality inside the relationship. Recall that the marginal return from higher quality outside the relationship

[12] Analytically, this property is similar to the profit function of a competitive firm being convex in the output price.

is fixed at 1. When $b \leq 1$, while higher quality is preferred both inside the relationship and outside, it is more valuable outside. As a result, it is not efficient for the assembler to try to attract high type vendors. Indeed, as we noted earlier, in this case if we ignore investment, low types would be preferred. But since they invest at least as much as the high types for any given level of orders (Result 1) considering investment just reinforces the preference for low type vendors. With $b > 1$ there is, a trade off. Now the marginal social return from higher quality, ignoring investment, is higher inside the relationship than outside. But high types invest less for the same level of orders than low types. When b is high enough (i.e., $b \geq \underline{b}$ where $\underline{b} \in (1, 1.5)$) the first effect dominates and the high types are always preferred. But for intermediate values of b (namely, $b \in (1, \underline{b})$) it is possible that low types become attractive since they invest more. Recall from Result 1 that the gap between the investment levels of the high type and the low type decreases as the level of orders increases. For very low levels of orders neither type of vendors invest, and so for $b > 1$ the high types would be preferred. For high levels of orders, the gap between the investments of high and low types is very small, and so once again the high types would be preferred. For intermediate levels of orders the gap between the investments of high and low types is large, and low types could be preferred.

Recall from Result 2 that one vendor is going to be given a certain order (the "first-preference" vendor) and the other vendor is going to be given the residual order α (the "second-preference" vendor). So for $b \leq 1$ both the first-preference and the second-preference are going to be low types, and similarly, for $b \geq \underline{b}$, both the first-preference and the second-preference are going to be high types. For intermediate values of b, the first-preference vendor is going to be a high type. If α is high the second-preference vendor is going to be a high type and the same is true if α is low. But if α takes an intermediate value, the second-preference vendor will be a low type vendor. Since α is the probability of the high demand state, and we are considering a binary distribution, the variance is $\alpha(1 - \alpha)$ which is high for intermediate values of α and low for high or low values of α. So this result tells us that the presence of greater uncertainty makes having a mixed portfolio of vendors more likely.

Now we proceed to characterize the investment levels undertaken by first and second-preference vendors:

Result 4: *The first-preference vendor will always undertake a higher level of investment than the second-preference vendor, and will therefore have lower unit costs.*

Proof: See the appendix.

If both first-preference and second-preference vendors happen to be the same type, naturally the former will invest more than the latter. However, if the first-preference vendor is high type, and the second-preference vendor is low type then the comparison is not straightforward - for the same order the high type invests less than the low type, but he happens to get a higher order. It turns out that in this particular instance, the comparison is actually straightforward. Being first-preference means you always receive an order of 1, and so the outside option gets no weight in the choice of investment under the first-best. Indeed, when the order is 1, the investment of a high type and a low type are the same (Result 1). But we know that a low type vendor with order $\alpha < 1$ invests less than a low type vendor with order 1. So even when the first-preference vendor is a high type and the second-preference vendor is a low type, in the first-best the former will invest more than the latter.

Remark 1: Suppose we allow the specific investment and the type of the vendor to be substitutes or complements *within* the relationship as opposed to just being substitutes in the outside option.[13] Does this qualitatively affect our results? Let us modify the above model such that $V(x, \theta) = a + b\theta + c_0 x - \delta\theta x$ where the parameter δ allows x and θ be complements ($\delta < 0$) or substitutes ($\delta > 0$) as opposed to being independent ($\delta = 0$) which is what we assumed in the benchmark model. This yields the following optimal choice of x:

$$x^*(\beta, \theta) = \max\{\beta(\gamma - \delta\theta) - (1 - \beta)\theta u, 0\}.$$

Clearly, if the investment and the type of the vendor are complements (substitutes) then the investment advantage of the low types decrease (increase) compared to the above model but the basic classification of alternative cases remain valid. However, when the investment and the type of the vendor are substitutes ($\delta > 0$) there is an interesting implication for the case where both high types and low types are chosen in equilibrium (i.e., b is greater than 1 but not too large and α is neither too large nor too small). To rule out extreme cases, let us assume $\gamma > \delta$. Otherwise a high type vendor (e.g., $\bar{\theta} = 1$) will not undertake any specific investments even if he faces a certain demand. Now the first-preference vendor (a high type) will choose an investment level

[13] Both are plausible *a priori.* To be selected by the assembler the quality must cross a certain threshold, and so those with lower quality must invest more to compensate which would imply x and θ are substitutes. Alternatively, having a good vendor work for you is likely to be more productive the more he invests in the relationship which would imply x and θ are complements.

of $\gamma - \delta\bar{\theta}$. However, the second-preference vendor (a low type) will choose an investment level of $\alpha(\gamma - \delta\underline{\theta}) - (1 - \beta)\underline{\theta}$. Now we have the interesting possibility that the second-preference vendor might invest more even though he has a lower demand. For example, let $\bar{\theta} = 1$ and $\underline{\theta} = 0$ and the required condition for this becomes $\gamma(1-\alpha) < \delta$ which is possible even though $\gamma > \delta$ by assumption.

Remark 2: Throughout we have assumed that all parties are risk neutral and showed that the assembler does not want to split a given level order between two vendors unless capacity constraints bind. If vendors are risk averse then that would add a cost to this kind of an allocation. The assembler would then want to give both vendors some orders in all states of the world. Unless the vendors are extremely risk averse, however, we would still expect one vendor to emerge as a first-preference vendor and another as a second-preference or marginal vendor.

3.2 The Second Best

Now we turn to the case where x is subject to transactions costs. Following the Grossman-Hart-Moore property-rights framework let us assume that x is observable but not verifiable. The price for vendor i, p_i, is negotiated after the investments are sunk, and the parties are assumed to adopt the Nash bargaining solution. The assembler bargains with each vendor separately and independently. If bargaining breaks down with a particular vendor after the investment is undertaken, the vendor is able to walk out of the relationship and earn its outside option $\bar{u}(x, \theta)$. The assembler can, in principle, find another vendor and buy the part from him. There are several costs of doing this - there are costs of screening and training a new vendor (which we have not modeled), the new vendor would not have had the time to invest, and there will be some loss of surplus due to delay in delivering to the final consumer.[14] For simplicity we assume that these costs are significant and so the assembler has to forego the sale of one unit of the final good in the current period if bargaining breaks down with one vendor.[15] In general, we can allow the assembler to replace a vendor by buying from another vendor if bargaining breaks down, and earn a surplus of $\mu(V(0,\theta) - \gamma(0) - \bar{u}(0,\theta)) = \mu(a - 1 + (b-1)\theta) \equiv \bar{\Pi}$ where $\mu \in (0,1)$ and $1 - \mu$ is a measure of the cost of delay. This will simply add a constant term to the payoffs, and will not substantively

[14]According to Williamson's notion of relationship-specificity, even if one party does not undertake any up front expenditures at all, his *ex ante* relationship-specific investment could just be the choice of a partner or a standard or anything else that limits his later options.

[15]This *ex post* bargaining power of the vendor is perfectly consistent with *ex ante* competition among vendors to be able to trade with the assembler, and is a consequence of relationship specificity.

affect the subsequent analysis.

The gross *ex post* surplus within the relationship if trade takes place is $V(x_i, \theta_i) - \gamma(x_i)$. The vendor's profit from dealing with the assembler per unit of the part (conditional on trade taking place) using the standard Nash-bargaining formula is

$$\pi_i \equiv p_i - \gamma(x_i) = \frac{V(x_i, \theta_i) - \gamma(x_i) + \bar{u}(x_i, \theta_i)}{2}. \tag{2}$$

The profit the assembler makes in his relationship with vendor i per unit of the part (conditional on trade taking place) is

$$\Pi_i \equiv V(x_i, \theta_i) - p_i = \frac{V(x_i, \theta_i) - \gamma(x_i) - \bar{u}(x_i, \theta_i)}{2}. \tag{3}$$

which is non-negative by Assumption 1.

The vendor would choose x_i to maximize

$$\beta\pi_i + (1-\beta)\bar{u}(x_i, \theta_i) - c(x_i).$$

The vendor's first-order condition is

$$x^{SB}(\beta, \theta) = \max\left\{\frac{\beta}{2}\gamma - (1 - \frac{\beta}{2})\theta u, 0\right\}. \tag{4}$$

where the superscript SB indicates that this is the optimal second-best allocation. As $\gamma < 1$, $x^{SB}(\beta, \theta) < 1$. From the previous section, we know that if $\frac{\beta}{2} \leq \hat{\beta}(\theta)$, or, $\beta \leq 2\hat{\beta}(\theta) = \frac{2\theta u}{\gamma + \theta u}$ then $x^{SB} = 0$. Let $\beta_0 \equiv 2\hat{\beta}(\underline{\theta})$ and $\beta_1 \equiv 2\hat{\beta}(\overline{\theta})$. As $\hat{\beta}(\theta)$ is increasing in θ, $\beta_0 < \beta_1$. We make the following additional assumption:

Assumption 2

$$\gamma > u.$$

This assumption ensures that high type vendors undertake *some* positive level of investment in the second-best situation when demand is high enough. Formally, Assumption 2 implies that $\beta_1 < 1$. For $\beta > \beta_1$ the investment levels of both types of vendors are positive, for $\beta_1 \geq \beta > \beta_0$, the investment level of the high type is 0, but that of the low type is positive, and for $\beta_0 \geq \beta \geq 0$ the investment levels of both types are zero.

Result 5: *Under the second-best:*

(i) The level of investment is lower for any given level of orders and any given vendor type compared to the first-best.

(ii) The level of investment is increasing in the level of orders.

(iii) The higher is the type of the vendor, the lower is the level of investment for the same level of order.

(iv) The gap between the investment levels of the two types of vendors for any given level of orders is larger than the first-best.

(v) For the range of orders where both types of vendors invest this gap decreases as the level of orders increases and but stays positive even when the level of orders is 1.

Proof: See the appendix.

The first part of the result illustrates the classic under-investment result due to hold-up. A vendor underinvests not only because he expects a fraction of the surplus from his investment to be appropriated by the assembler in the *ex post* bargaining game, as in standard hold-up models, but in addition because a higher level of investment directly reduces his outside option and hence, his bargaining power. The second and the third parts are similar to that in the previous section. The fourth part follows from the fact that as investment increases, the value of a high type vendor's outside option decreases at a higher rate compared to a low type vendor and because of this, their marginal loss from investing is higher. This was true also under the first-best, but now a decrease in the outside option leads to both lower flexibility and lower bargaining power. As a result, for the same level of orders, high type vendors invest less than low types, as under the first-best but the gap between their investment levels is higher. For the fifth part, as before as demand increases the weight assigned to the outside option falls and so the gap between the investments of the high type and the low type decreases. However, even when demand is certain (i.e., $\beta = 1$) the outside option gets some weight unlike in the first-best, because of the value of the outside option in bargaining and as a result, the gap between the investment levels of the two types of vendors remain positive.

Using (3), let the profit *per unit of the part* that the assembler makes from trading with a vendor whose type is θ and who gets an order of β be denoted as:

$$\Pi(\beta, \theta) \equiv \frac{(a - 1 + (b - 1)\,\theta) + (\gamma + \theta u)\, x^{SB}(\beta, \theta)}{2}. \tag{5}$$

Let

$$\Pi^*(\beta) = \max\{\Pi(\beta, \underline{\theta}), \Pi(\beta, \bar{\theta})\}$$

denote the upper envelope of the profit per unit schedule facing the assembler who can choose the type of vendor that is most profitable to deal with for any given level of order β. The expected profit of the assembler from a vendor who is given an order of β is

$$\beta\Pi^*(\beta).$$

When deciding on how to allocate orders, the assembler will consider this function, and not expected joint surplus as under the first-best. Now we are ready to characterize how should an assembler distribute its orders among the vendors under the second-best:

> **Result 6:** *Given Assumptions 1 and 2 under the second-best the optimal way to distribute orders would be to give one vendor a certain order of* 1, *and the other vendor an order of* 1 *with probability* α *and* 0 *with probability* $1-\alpha$.
>
> **Proof**: See the appendix.

The intuition is as follows. Unlike the first best, the assembler does not care about *ex ante* expected joint surplus anymore. Rather, he cares about his share of the surplus in the post-investment bargaining game, i.e., $\beta\Pi^*(\beta)$. The greater is the level of orders, the higher is the investment effect and the higher is the assembler's *per unit* profits, $\Pi^*(\beta)$. In our set up, this relationship is linear - either per unit profits do not depend on the level of orders (if the level of investment has a corner solution) or it is linearly increasing in the level of orders because the level of investment is an increasing and linear function of the level of orders. This implies that the profit rate of an assembler from a given type of vendor is an increasing and (weakly) convex function of the level of orders. Since the assembler can choose the type of a vendor, the true profit rate is the upper envelope of the profit rates for each type, which is therefore a convex function. Since the *level* of profits is just the rate of profits times the level of orders, it too must be a convex and increasing function of the level of orders. Given this the way the assembler should distribute its orders among the vendors follows using the same arguments as under the first-best.

The following result characterizes the choice of the type of vendor:

> **Result 7:** *Given Assumptions 1 and 2 under the second-best:*
>
> *(i) For* $b \leq 1$ *both the first and second-preference vendors are going to be low types.*
>
> *(ii) For* $b > \bar{b}$ *where* $\bar{b} > \underline{b}$ *both the first and second-preference vendors are going to be high types.*

(iii) For $b \in (b_0, \bar{b}]$ where $b_0 < \bar{b}$ the first-preference vendor is going to be a high type vendor and the second preference vendor is going to be a high type vendor if α is very low or very high but a low type vendor otherwise.

(iv) *For $b \in (1, b_0]$ the first-preference vendor is going to be a low type vendor and the second preference vendor is going to be a high type vendor if α is very low, but a low type vendor otherwise.*

Proof: See the appendix.

The only difference with the corresponding result for the first-best case (Result 3) is that the parameter region for which an assembler might choose to deal with both high type and low type vendors is larger now. This is because two additional factors diminish the attraction of high type vendors relative to low type vendors under the second-best. First, the assembler has lower bargaining power in dealing with them and so even if the high types invested as much as the low types for the same order, the relative attraction of the high type is less. Second, the investment gap between high types and low types is greater than under the first-best (Result 5(iv)). So for $b \leq 1$, like under the first-best, there is no reason to prefer high types and both the first and second preference vendors are going to be low types. For $b > 1$, now the threshold level of b such that only high types are preferred, $\bar{b}$, is higher than the one under the first-best, $\underline{b}$. For the case where $b > 1$ there are several possibilities which are illustrated using Figure 3 where we plot per unit profits $\Pi^*(\beta)$. When both type of vendors do not invest (i.e., $\beta \in [0, \beta_0]$) then high types are preferred, as before. Over the range where the low type vendor invests (i.e., $\beta \in [\beta_0, \beta_1]$) but the high type vendor does not, the relative attractiveness of low type vendors increases as their investment is increasing in β. It reaches a maximum at the level of demand $\beta = \beta_1$ from which point onwards the high type starts investing. For the range where both types of vendors choose positive investment levels (i.e., $\beta \in [\beta_1, 1]$), Π increases at a faster rate with respect to β for the high type vendor. This follows directly from the fact that an increase in the size of the order β leads to a greater increase in x for higher types since their investment is more sensitive (in a negative way) to the weight assigned to the outside option. This means, the relative attractiveness of low type vendors fall in this range. However that the difference in the profit rate when $\beta = 1$ cannot be as large as when none of the vendors invest, since the low type vendor always invests more than the high type vendor. Three cases are possible depending on parameter values and these are depicted in Figures 3(a), 3(b) and 3(c). Figure 3(a) considers the case $b > \bar{b}$, and here high types are always preferred and so both the first preference and the second-preference vendor would be high types. Figure 3(b) considers the case $b_0 < b \leq \bar{b}$, and here low types are preferred for intermediate values of β but high types are preferred for high and low values of β. Here the first preference vendor would be high type, for intermediate values of α the second preference vendor would be low type, and for very high or very low values of α, it would be a high type. Finally, Figure 3(c) considers the case $1 < b \leq b_0$, and here high types are preferred only for low values of β (such that no one invests), otherwise low types are preferred. In this case the first preference vendor would be low type, and the second preference vendor would be a low type for high or intermediate values of α and a high type for low values of α.

Finally, we proceed to characterize the investment levels undertaken by first and second-preference vendors:

> **Result 8:** *The first-preference vendor may undertake a lower level of investment than the second-preference vendor, and have higher unit costs.*
>
> **Proof**: See the appendix.

Like in the previous section if both first-preference and second-preference vendors happen to be the same type, naturally the former will invest more than the latter. However there even if the first-preference vendor was high type and the second-preference vendor was low type, the former would invest more than the latter. The reason was, in the first-best when the order is 1, the investment of a high type and a low type are the same. But under the second best, as Result 5 indicates, even when demand is certain (i.e., $\beta = 1$) a high type vendor invests less than what a low type vendor would have invested for the same level of orders.

We end this section by making a few remarks about some features of the model.

Remark 3: Why doesn't the assembler own the asset instead of the vendors owning it? In this model, we are assuming that the assembler does not undertake any significant specific investments with a specific vendor the *ex ante* choice of which could be affected by *ex post* bargaining. This is justified by the institutional setting. MTL makes tractors based on blueprints provided by MTL's foreign partner, Massey-Ferguson and gives the vendor an imported sample of a relevant part. Any technological support provided is not specific to a particular vendor, but to all vendors that supply that part. Given this, ownership by the assembler will only help to reduce the vendor's share of *ex post* surplus, which would dampen his investment incentives.[16] However, there is another force at

[16]This assumes at least part of the investment is embodied in the human capital of the vendor, as in the Grossman-

work which would favor ownership by the assembler even if he does not undertake any relationship specific investments. If vendors own the asset, the level of investment they undertake is lower than standard hold-up models because lower investment boosts up their outside option. This potentially adverse effect of ownership has been pointed out by Rajan and Zingales (1998).

Remark 4: The "just-in-time" management system is fully dependent on minimizing inventory holding, which requires timely delivery from the vendors. Given that demand is stochastic and the time of delivery is not contracted upon *ex ante,* the assembler would want to give incentives to the vendors to ensure timely delivery. The assembler is obviously concerned about "undersupply" (i.e., the quantity delivered being less than quantity ordered) but given the goal of minimizing inventory holding, "oversupply" is likely to be a problem as well (i.e., the quantity delivered being more than quantity ordered). By a simple dynamic extension of the model we can show the following : in order to induce effort to ensure timely delivery, both undersupplying and oversupplying need to be punished, but the greater the outside option of a vendor, the greater will be his incentive to sell to the outside market on the side at the expense of delays to the assembler. As a result such vendors are going to be punished more heavily for undersupplying, which implies that vendors with low outside options are going to be punished relatively more heavily for oversupplying.

4 Evidence

The theory offers several predictions based on the importance of vendor quality inside the relationship (b), the level of demand uncertainty facing the buyer (α), and the nature of contractibility of relationship specific investments. In this section we analyze primary data collected on contracts between the largest tractor assembler in Pakistan, Millat Tractors limited (MTL), and some of its vendors. Our aim is to find out how contractual outcomes and performance measures are related to asset specificity, controlling for various factors, including other vendor characteristics and then interpret our results in terms of the theory in the subsequent section.

4.1 Data

Our empirical study is based on two data sources. First, we use data from a survey of the automobile industry vendors conducted by the Lahore University of Management Sciences (LUMS) in 1997 to construct vendor attributes. The 28 vendors in the sample were randomly selected from the set

Hart-Moore framework.

of all of MTL's vendors. Second, our primary dataset consists of vendor-assembler contracts (for only those vendors in the secondary source) and additional part level information from the records and archives at MTL. This dataset gives detailed contractual features and is merged with the LUMS dataset. The contract data is available for the period 1989-1999. The parts are fairly representative, ranging from low priced simple products such as tractor clips (Rs. 0.2) to high priced complex products such as the transmission case (Rs. 5306).

We restrict attention to parts for which MTL bought from more than one of the 28 vendor's in our sample. This restriction allows us to address the main question of interest, namely how prices and quantities vary across vendors supplying the *same* part. Otherwise we would be comparing single vendors that supply different products, and even though we allow for a separate intercept for each part, we would not be able to separate out the effect of a firm's characteristics from the effect of technology. The drawback of this restriction is that it leaves us with only 19 out of our original 28 vendors and therefore produces potentially higher standard errors. However, it enables us to provide a much cleaner interpretation of our results. With this restriction we are left with 39 different parts; 37 of the 39 parts in this restricted data set the number of sample-vendors per part is 2, and for the remaining two it is 3. In Table A1 we list the parts in our sample and show how many of the vendors for whom we have attribute data produce that particular part. Note that this does not mean that the majority of parts do not have multiple vendors but that, while there are multiple vendors for the excluded parts, our secondary data source does not have data on these vendors. In fact for a majority of the parts that MTL outsources it generally has at least two vendors.[17]

Table 1 gives the summary statistics for some of the main variables used for the restricted dataset. The definition the primary variables used in the analysis are:

P = Agreed contract price for a given vendor and contract year;

C = MTL engineer's estimate of cost of production of a particular part for a given vendor in a given year;[18]

[17]The choice of how may vendors to have for a given part and whether this effects contractual outcomes is an interesting question which, given the nature of the data we cannot address satisfactorily. However, it appears that the number of vendors supplying a part is largely time-invariant and so such effects are already accounted for by using part-specific fixed effects.

[18]For a subset of the data, we have vendor-specific cost estimates for each part made by MTL engineers before negotiation of the contract. These cost figures should be interpreted as economic costs since they include a standard markup.

Q_S = Quantity scheduled from the vendor during a quarter;

Q_R = Quantity received from the vendor during a quarter;

TQ_S = Total quantity scheduled for a given part from all MTL vendors (including those not in our sample) during a quarter;[19]

TQ_R = Total quantity received for a given part from all MTL vendors (including those not in our sample) during a quarter;

R = Fraction of the vendor's quarterly received quantity that is rejected by the MTL quality inspection section;

Age = Vendor age;

$Specificity$ = The percentage of the vendor's physical assets/equipment that would become idle (i.e. would have to be scrapped) if MTL stopped buying from it;

$Size$ = Size of the vendor's labor force (in 1995);

$Distance$ = Distance of the vendor (km) from MTL;

$City$ = A dummy variable that equals 1 if the vendor is located in Karachi and 0 if in Lahore;

The attribute that is going to be of particular interest to us is the measure of specificity. To motivate the empirical analysis we need to relate the measure of specificity to the theoretical model. Since prices do not depend on the level of orders, this measure tells us what percentage of the *value* of a vendor's assets will become zero if MTL stops buying from him and corresponds closely to the measure of asset specificity suggested in Williamson (1985) i.e. the aggregate level of quasi-rents created by the investment. In terms of our previous notation, $S = V(x, \theta) - \gamma(x)$ is the measure of the value of the asset owned by a vendor if MTL buys from him, and $\bar{u}(x, \theta)$ is the measure of the outside option of the vendor if MTL stops buying from him where x is the level of relationship specific investment and θ, the vendor's quality. Therefore, our measure of specificity corresponds to:

$$\text{Specificity} = (S - \bar{u})/S = 1 - \frac{\bar{u}}{S}.$$

[19] For the parts we have in our sample, the vendors included in our sample on average account for 80% of total scheduled quantity (the range being 16-100%), the rest coming from vendors not in our sample.

Thus specificity is a function, $f(x,\theta)$, of x and θ. Since the outside option of a vendor is decreasing in the extent of relationship-specific investment, while the surplus within a relationship is increasing in it, it follows directly that a higher level of x will lead to a higher degree of specificity holding θ constant. An increase in θ holding x constant reduces specificity if the percentage change in the outside option is higher than the percentage change in the surplus within the relationship. That is, the condition for $\frac{\partial\left(\frac{\bar{u}}{S}\right)}{\partial\theta}$ to be positive is $\frac{1}{\bar{u}}\frac{\partial\bar{u}}{\partial\theta} > \frac{1}{S}\frac{\partial S}{\partial\theta}$. Since we interpret θ as a measure of the versatility of the vendor, this is an intuitive condition.[20] However, x in turn may depend on θ as determined by the different scenarios in our model and so the correlation between specificity and vendor quality is ambiguous.

Table 2 shows correlations between various vendor characteristics of interest: Vendors located in Karachi have significantly higher levels of specificity and this is supported by anecdotal evidence from MTL; some of MTL's earliest vendors were in Karachi and they took the initiative in developing parts locally and invested heavily in assets specific to MTL. In addition, from the various other attributes that are available from the LUMS survey, the only other attribute that is significantly correlated with specificity is the extent of excess capacity. While all vendors faced excess capacity to some degree in the year of the survey, it is significantly negatively correlated to a vendor's specificity measure.

4.2 Results

An important part of the setting and in the data is that there is considerable price and quantity variation among vendors supplying the same part. Explaining what vendor attributes are related to this variation thus forms the primary empirical question we address. In particular, we are interested in examining how a vendor's degree of specificity to MTL determines this variation.

First we would like to establish that such there is in fact significant variation in the prices: restricting ourselves to cases where two different vendors are supplying a given part *in the same year* we get that on average one vendor gets a 25% higher price than the other. Doing the same for quantity supplied shows that on average a vendor is scheduled three times as much as another vendor supplying the same part in the same year. There are also substantial yearly fluctuations in aggregate quantity for each part. We compute the coefficient of variation (CV) for the total quarterly quantity supplied by the vendor for each part in our multiple-vendor restricted sample. The mean value of the part-CV for all the parts is 0.99 for (quarterly) scheduled quantity, suggesting

[20]It is satisfied for our model. Formally, $\frac{\partial\left(\frac{\bar{u}}{S}\right)}{\partial\theta} = \frac{(1-ux)(a+b\theta-1+\gamma x)-\theta(1-ux)b}{S^2} > 0$.

that there is considerable variation in orders and that MTL passes a significant part of its total sales shock to its vendors. This variation could be partly capturing growth over time but quantity plots suggest that this is not the case, i.e., there are no consistent time trends. We now turn to the empirical specification used to explain such variation across vendors.

As a result of the restriction described earlier, we are guaranteed that we have more than one vendor for each part in our sample. This enables us to use part and time fixed effects to effectively contrast contractual outcomes for two different vendors supplying the same part in the same year. This is the tightest restriction that we can apply, though at the cost of reducing our sample size. Part fixed effects control for potentially important and confounding aspects specific to the part (such as its technological nature, how critical it is, the number of vendors supplying it provided that this does not change over the sample period etc.) and time fixed effects controls for common period-specific shocks such as inflation, demand, government policy changes. We will mostly be estimating equations of the form:

$$C_{ijt} = \alpha_i + \tau_t + \sum_j \beta_j X_j + error. \tag{6}$$

C_{ijt} = (level of) contractual feature/outcome for part i, vendor j, in period (year or quarter) t.

X_j = time-invariant vendor characteristic.

α_i = part-specific intercept

τ_t = period-specific (year/quarter) intercept.

In addition, some specifications will also consider interaction terms between vendor attributes and lagged or aggregate contractual outcomes on the right hand side.

One caveat is that the specification above treats vendor characteristics as time invariant. At a practical level this is dictated by the fact that we have information on some vendor characteristics such as specificity for only one year (1997 for specificity). However, we are not unduly concerned at this since we believe that these attributes do not change substantially over the short time period under study: 11 years, with 96% of the data coming from the period 1993-1999. Moreover, even if they do change, as long as they change uniformly (vendors do not keep switching rank) there is unlikely to be any systematic bias in our results. In addition, the assumption of inertial vendor attributes is supported for cases where we do have another year of vendor characteristics (for example, correlations between vendor size observations in 1985, 1990 and 1995 are all above 0.95 and highly significant).

The main results are summarized below. We generally look at the logarithm of these variables rather than levels, because different products have very different levels of prices and quantities, and the aim of the analysis is to explain the percentage changes in level differences across vendors (i.e., log differences) and not the absolute level of differences. While we are primarily interested in vendor specificity, we also mention robust results for other vendor attributes. Since we want to distinguish vendor quality (as potentially captured by the specificity measure) from learning-by-doing effects that affect all firms or economies of scale with respect to size, we control for the age of a firm and its size in all regressions. Since the mean age of vendors is 15 years in our sample (the range being 3-34 years) asymmetric information or reputation based factors are unlikely to be strong influences.[21] For all regression tables, the first column reports the result with a smaller set of basic controls (specificity, age, size) and the second column adds more controls (the distance from MTL, a city dummy). The estimated coefficients below are from the last column unless noted otherwise.

1. **Vendor Specificity ("tiedness"):**

   (a) *Level Outcomes*:

   - **Cost:** Table 3 shows that per unit production costs are lower for tied vendors. Since the estimated coefficient for specificity is -0.0016 and the dependent variable is log cost, this implies an increase in specificity from 0 to 100 decreases level costs by approximately 15%. Alternatively, a standard deviation increase in tiedness lowers costs by 6.6%.[22]
   - **Price:** Tied vendors get a lower price. Table 4 shows that an increase in the measure of specificity from 0 to 100 decreases price by 19%; a standard deviation increase in tiedness results in a 7.9% lower price level.[23] We assume that the quality of the lot supplied does not affect current prices. This is reasonable, as the price of each part to be supplied by a vendor is agreed upon for the entire year before delivery begins.

[21] We also have data on years of relationship. It is highly correlated with the age of a vendor and our results are similar if we use this variable instead of age.

[22] In general, the magnitude of the effect is calculated from a one standard deviation increase in specificity starting at 0 specificity. The percentage change is then evaluated at mean of the dependent variable.

[23] Note that we have an unbalanced panel i.e. for all years the set of parts are not the same. One concern is that our results are based only on a comparison of the same part for two vendors across *different* years. However, we can check for this by allowing for interacted part-year dummy variables. Doing so gives similar results allaying our concern, although, as expected, the standard errors are higher.

- **Quantity:** Table 5 (Column 1) shows that tied vendors get lower quantity orders. A standard deviation increase in tiedness is associated with a 5.5% lower quarterly vendor scheduled quantity. However, this result weakens once we add city and distance controls.

(b) **Coefficient of Variation:** Table 6 (Column 1) shows that tied vendors are given more unstable quantity schedules. A standard deviation increase in tiedness is associated with a 5.4% higher coefficient of variation of vendor quarterly scheduled quantity. The result weakens with city and distance controls. Unlike other specifications, there is no time-varying component in these regressions since each observation is at the vendor-part level. This explains the smaller sample size.

(c) **Rejections:** Table 7 shows that tied vendors' supply is of lower quality. There is a positive association between vendor specificity and the proportion of the vendor's order that is rejected because it fails to meet MTL's quality standards. A standard deviation increase in tiedness is associated with a 43.6% increase in the proportion rejected. However it should be noted that this proportion is generally quite low, with a mean value of 1% (the increase is from 1% to 1.43%) suggesting that in general the MTL quality control is fairly effective.

(d) **Quantity Elasticity:** The scheduled quantity of tied vendors is more elastic to the overall scheduled quantity (Table 8), but for such vendors received quantity is relatively less elastic to both the vendor's own scheduled quantity (Table 9) and the total scheduled quantity (Table 10). Table 8 shows that a standard deviation increase in tiedness is associated with a 3.7 *percentage points* increase in elasticity (i.e. from an elasticity of 0.739 to 0.776).[24] Tables 9-10 show that a standard deviation increase in tiedness is associated with a 8.7 percentage points decrease (from 0.474 to 0.387) and a 7.8 percentage points decrease (from 0.289 to 0.211) in elasticity with respect to own scheduled and total scheduled quantity respectively. If the relative share of a vendor in total quantity scheduled was always kept constant, then the average elasticity of vendor scheduled quantity to total scheduled quantity would be one, which is, in general, not true otherwise. Therefore, it seems that tied vendors face a more fluctuating order schedule (Table 8). However, it also appears that tied vendors have poorer delivery performance

[24] Estimated by adding to the coefficient of log total scheduled quantity (0.7392) the product of the coefficient of the interaction term (0.0009) and a one standard deviation change in specificity.

(Table 9). While MTL gives them a schedule that is more responsive to its own overall needs, these vendors are less responsive to MTL's requests. A concern in only considering responsiveness of a vendor with respect to its own scheduled quantity is that it may look worse since it is given a harder task. Table 10 addresses this concern by looking at vendor responsiveness to total scheduled quantity (from all MTL vendors, including those not in the sample).

(e) **Under and Over-supply:** Tables 11 and 12, provide further evidence for the poorer delivery performance of tied vendors suggested by the elasticity results above. These tables show that tied vendors are more likely to both under and oversupply quantity (i.e. the quantity received from them is less or more than what MTL ordered from them). Table 11 shows that a standard deviation increase in tiedness is associated with a 24% increase in the quantity that is undersupplied by the vendor. Table 12 (Column 1) shows that a standard deviation increase in specificity is associated with a 10% increase in quantity oversupplied but this result loses significance once we add city and distance controls.[25]

(f) **Dynamic Incentives:** Given that timely delivery is an issue for MTL, we also look for evidence of dynamic incentives given by MTL (Tables 13-14). In particular we check whether poor delivery performance in the previous quarter is penalized (in terms of quantity scheduled this quarter) in the current one and specifically whether these penalties differ with the vendor's specificity level. We find that while tied vendors are penalized less for undersupplying last period, they are penalized relatively more for oversupplying Table 13 shows that a standard deviation increase in the amount undersupplied last quarter by a completely untied vendor (one with 0 specificity) results in the vendor getting a 37% decrease in quantity scheduled this quarter while this decrease is only 14% for a fully tied vendor (one with 100% specificity). Table 14 shows that an untied vendor is rewarded for oversupplying last quarter: A standard deviation in the amount oversupplied last quarter results in an 11.5% increase in scheduled quantity this quarter. However a completely tied vendor is strictly punished for oversupplying last period. A standard deviation in the amount oversupplied last quarter results in a 6.3% decrease in scheduled quantity this quarter.

[25]Table A2 in the appendix considers the full sample and treats under and oversupplying as symmetric and gets similar results.

2. *Vendor Age:*

   Table 3 shows that older vendors have lower cost estimates suggesting returns to learning over time. Table 4 shows that older vendors get a lower price: A standard deviation increase in age (6.5 years) results in a 1.64% lower price. Tables 5 and 8 show a weak positive relationship between vendor age with log of quantity scheduled, and Table 6 shows a weak negative relationship with the coefficient of variation of scheduled quantity. Table 7 suggests however, that older vendors tend to be of a poorer quality, at least in terms of the fraction of their supplied quantity that is rejected by MTL. An extra year of age increase rejections by 15.2% (from a mean value of 1% to 1.52%). Table 11 offers another dimension along which older vendors are poorer i.e. older vendors undersupply more: an extra year increases the amount undersupplied by 1.6%. However, older vendors tend not to oversupply (Table 12).

3. *Vendor Size:*

   Tables 3 and 4 shows a weak negative association of vendor size with cost and price. Tables 5 and 8 show no robust relationship between vendor size and log scheduled quantity. Table 6 provides weak evidence that larger vendor face a more variable schedule, and Table 7 shows that they are significantly associated with lower rejections. Tables 11-12 weakly suggest that larger vendors are more likely to both undersupply and oversupply.

4. *Other controls:*

   Distance and city are conflated since the distance when the city is Karachi is significantly higher (1300 km versus a mean distance of 122 km for Lahore firms). Moreover, there is little variation in distance for the different vendors in Karachi. So the distance coefficient should be interpreted as a "local (near Lahore) distance effect" and the city coefficient is likely to capture a Karachi fixed effect. The evidence suggests that Karachi vendors face weakly higher prices and significantly higher costs (Tables 3-4). They get a significantly lower scheduled quantity (Table 5). Table 6 also suggests that Karachi vendors are given a more unstable scheduled quantity order. Tables 11-12 suggest that Karachi vendors undersupply less but oversupply more. For the local (within Lahore) distance effect, vendors that are more distant from MTL have higher costs (Table 3), a lower quantity scheduled from them (Table 5), and a more unstable quantity schedule (Table 6). There is weak evidence that they have lower quality in terms of proportion of received quantity rejected (Table 7), and undersupplying (Table 11).

4.3 Concerns About the Measure of Specificity

Since specificity is essential to our analysis it is worth addressing some concerns regarding the empirical measure of specificity we use, which is the percentage of the vendor's physical assets/equipment that would become idle if MTL stopped buying from it. Since the size of a vendor's plant is not correlated with this measure (and in any case we control for vendor size in our regressions), this measure is not picking up the fact that different vendors have different sizes.

The main concern is does our measure of specificity reflect the importance of specific investments or is it capturing some omitted variable such as a vendor's dependence on the assembler or his confidence in its own abilities. We address this concern below.

First, we show that the measure is correlated with objective assessments of technological processes. To do so we used MTL engineers rankings of technological processes used by our vendors by their degree of specificity. In order of increasing specificity these were casting, machining and forging. Using their classification we then grouped vendors according to the processes they used and summarized the mean specificity for each category: Vendors that did casting had a mean specificity of 0, those that did machining had a mean of 62 and those that did forging a mean of 66. Note that since our estimation uses part-specific fixed effects, the variation in specificity we are capturing is both the variation in processes two vendors use to make the same part (and such variation is too fine to be captured by our crude main process rankings) but also (since the specificity measure is vendor and not part specific) the (mean) specificity of all processes a vendor requires given the set of all the parts that it supplies to MTL. These facts reassure us that the specificity measure is indeed has some technological content.

Second, while the above shows our specificity measure has objective technological information, we also check for an obvious mis-interpretation with "sales reliance" on MTL. The vendor survey also asks the vendor for its percentage sales to MTL. Reassuringly, we find that a vendor's percentage sales to MTL and specificity are not correlated, suggesting that this mis-interpretation is not an issue. One might expect that a vendor that has less dedicated assets to MTL also be making few sales to MTL and vice versa. Neither relation is necessarily true. First, a vendor could be making all its sales to MTL but it could be using generic machinery to produce the parts and as such have low specificity. As a result, if MTL stops buying from it, it could switch to another buyer. Indeed, in our field interviews we found a vendor involved in machining that makes all its sales to MTL, but since it uses lathe machines that can easily be switched to produce parts for another buyer if the need arises, it reported a specificity of zero. Now consider the converse. It would seem that a vendor that has a high specificity should be making most of its sales to MTL. However, this is not necessarily the case. Sales are quite variable in this environment. As a result, all vendors face excess capacity, and indeed, those with greater specificity face significantly higher excess capacity.[26] So a vendor could be very tied to MTL but could still be making most of its sales for the part it produces for MTL in the replacement market when it faces excess capacity. Therefore, at a given point in time a vendor's relative sales to MTL need not be very correlated with its specificity to MTL.

A somewhat different concern is that even though our measure is similar to the standard measures of asset specificity used in the empirical literature on transaction cost economics, namely, a manager is asked on a scale of 1-5 or 1-7 the degree to which an asset has value in outside uses (Shelanski and Klein, 1995), a limitation of these measures is that they are likely to be subjective. For example, for the same objective measure of specificity, a more risk averse person could report a higher cardinal measure. While we cannot directly address this, we argue that this is not a serious concern since our results are robust to treating specificity as a binary (high or low) measure. This is not surprising given that distribution of specificity is fairly bimodal - more than half the vendors in our restricted sample either claim to be fully tied ($= 100\%$) or untied ($= 0, 1\%$).[27]

Another empirical concern is that specificity is, theoretically, vendor-part specific and not just vendor specific. However, the specificity measure used is the overall specificity of the vendor and there is no obvious way to decompose it into part-specific specificity measures. Nevertheless, a vendor's specificity choice is likely to be correlated across the set of parts it supplies as the asset specificity measure is likely to include an organizational/managerial component i.e. a vendor is tied to MTL not just because of the choice of its machinery for a particular part, but also because its operations/organizational setup is more suited to MTL's needs.

Finally a few words about the appropriateness of using specificity as an independent variable. Since investment decisions that determine a vendor's specificity (choice of technology, tooling etc.) are substantial and likely to have been undertaken at an initial stage in the vendor's life or at least prior to its producing the part in question (i.e. when a vendor first entered the business, or developed a new part) and therefore, *before* prices are negotiated or quantities are ordered or received, we do not have to worry about any effect of these outcomes on specificity. In other words,

[26]The average capacity utilization among the MTL vendors in the LUMS survey in 1995 was 75%.

[27]The remaining are mostly in the 30-40% range giving a smaller mode in the middle.

we argue that specificity can be safely regarded as "exogenous" to our outcomes of interest.

5 Discussion

In this section we discuss the interpretation of the empirical results in light of the theoretical model. As the model suggested, once we allow for vendor heterogeneity, it is likely that the specificity measure will pick up vendor characteristics and it is endogenous in that sense. While we cannot separate out the effect of investment from vendor heterogeneity[28], we can use the (potentially biased) empirical estimates to infer whether and how unobserved vendor heterogeneity does matter in interpreting the effects of specificity on contractual outcomes. We also discuss if there are alternative explanations of our results.

5.1 Interpreting the Evidence in Terms of the Theoretical Model

From the theoretical analysis recall that both in the first-best and in the second best, when b is small the buyer chooses only low quality vendors and when b is large, only high quality vendors. In these cases the variation in specificity only captures variation in x. However, for intermediate values of b, this is no longer the case and in general, variation in specificity will also capture vendor heterogeneity. We argue below that the results suggest that first-preference vendors are high quality vendors and the second-preference vendors are low quality vendors, even though the latter have higher levels of specificity. This corresponds to a specific case under the first-best (case (iii) of Result 3) and under the second-best (case (iii) of Result 7) where in addition demand is very uncertain (i.e., α is neither very high nor very low). We first show that our results are consistent with this case and then argue why they are not compatible with the other cases.

Tied (higher specificity) vendors have lower costs (Table 3), and are offered lower prices (Table 4). This presents us with the puzzle that if tied vendors are cheaper, why does MTL buy from untied vendors? A possible answer to this question could be that tied vendors have fixed capacities, and so untied vendors might have to be called in to supply residual amounts. However, the evidence suggests exactly the opposite. Not only do tied vendors get a smaller order (Table 5), but that they receive a less stable order (Tables 6,8). Also, data from the LUMS survey also suggests that

[28]Note that putting in fixed effects for vendors will not help since both quality and asset-specificity are likely to be time-invariant. Having an instrument for vendor type too will not solve the problem, since according to the model, it would also affect the choice of specificity.

tied vendors face significantly higher levels of excess capacity. In the terminology of our theoretical setup, therefore, less tied vendors are "first preference" vendors. All these results are consistent with case (iii) of Result 3 and with case (iii) of Result 7. MTL buys from both high and low quality vendors, with the former being the first preference vendor but the latter reporting a higher level of specificity. Now for case (iii) of Result 3 , the first-preference vendor actually chooses a higher level of x than the second-preference vendor. But this is consistent with the second preference vendor reporting a higher level of specificity, if her quality is very low (as $\underline{\theta}$ goes to zero, the specificity measure for the second preference vendor goes to 1 while that of the first-preference vendor will be some positive fraction). Moreover, if x and θ are substitutes in the surplus within the relationship, then as Remark 1 points out, the second-preference vendor could invest more than the first-preference vendor and then it is obvious why it would have a higher level of specificity (even if $\underline{\theta}$ is not too low). For case (iii) of Result 7, the first-preference vendor could choose a lower level of x than the second-preference vendor even if x and θ are not substitutes in the surplus within the relationship and in this case too the second-preference vendor would have a higher level of specificity.

Moreover, these cases predict that tied vendors should have an inferior performance record since they are of lower quality although this is partly mitigated by vendors undertaking higher specific investments.[29] Tables 7 and 9-12 provide evidence that more tied vendors perform worse in terms of various performance measures: Not only do parts supplied by more tied vendors face greater rejections (Table 7) but these vendors also respond less to MTL's quantity orders (Tables 9, 10) and are also more likely to under and oversupply (Tables 11, 12).

It is interesting to note here that Asanuma (1989) reports some facts from his case study of suppliers of automobile parts in Japan that are similar to our findings. In particular he reports that the buyers try to keep business as continuous as possible with better performing suppliers and other suppliers are treated as marginal suppliers. According to him the more intermittent and uneven the demand for the final product, the more necessary it is for this firm to keep some marginal suppliers as a capacity buffer.

Finally, turning to the dynamic results, the extension of the theoretical model (Remark 4) suggests that vendors with lower outside options (tied vendors in case (iii) of Results 3 and 7) should be punished relatively more for oversupplying than undersupplying and vice-versa for higher

[29]Recall that in the model we allowed V, the value generated in the relationship, to depend positively on x in addition to θ i.e. vendors can improve their perfomance through their investments.

outside option vendors. Tables 13 and 14 show that this indeed is the case.

Having argued that our empirical results are consistent with a particular case of our model, we briefly argue they are incompatible with the other cases. If MTL chooses to deal with only high type vendors or only low type vendors then that cannot explain the empirical result that suggests vendors with higher specificity are treated as second preference vendors. Moreover, it is also not consistent with vendors reporting higher specificity performing worse. If these were the relevant cases then vendors have the same quality, and since a higher x improves performance we would expect more tied vendors to perform better, not worse. Under the second-best there is an additional possibility (case (iv) of Result 7): the first preference vendor is a low type. It therefore should unambiguously have a higher specificity level than the second preference vendor who can be either low or high quality depending on the level of uncertainty. This is inconsistent with our empirical findings on two counts. First, the first preference vendor in fact has a lower specificity level, and second, it is observed to be of better quality than the second preference vendor.

5.2 Alternative Explanations

In this section we discuss whether there are alternative, perhaps simpler, theoretical explanations of the key empirical findings.

First, let us consider if such an explanation can be provided without relying on the role of specific investments. In order to be consistent with our finding that tied vendors are willing to be treated as marginal vendors an explanation would require that tied vendors are of lower *ex ante* quality. One may argue that our specificity measure only captures how dependent a vendor is on MTL i.e. low quality vendors report higher specificity because their outside options are worse and not because they are unable to supply to other buyers because of their specific investments. This argument is not as plausible given that we have already argued above that our specificity measure is in fact capturing technological specificity and not sales dependance. Moreover, if this argument is valid, then in the first-best MTL would deal with only high quality vendors or only low quality vendors, which is contrary to what the evidence suggests. Therefore, one would also need to explain why MTL chooses to buy from both low and high quality vendors. One explanation for this is a shortage of high quality vendors and so MTL is forced to deal with some low quality ones, especially when faced with a positive demand shock.

However, it is unlikely that high quality vendors are in short supply across the board for all parts given MTL's pre-eminent position as the largest tractor assembling firm in the country and that MTL carefully selects its pool of (200) active vendors out of the 2,000 auto-parts making firms in Karachi and Lahore. Moreover, even for its existing high quality vendors, our data shows that neither are their sales exclusively to MTL, nor are they producing near plant capacity, further suggesting that MTL *chooses* not to have them supply more.[30] Thus while it cannot be completely ruled out, the scarcity of good vendors is unlikely to be an important consideration. Alternatively, we could make restrictive assumptions on the demand function facing MTL. In particular, it is possible that MTL deliberately demands both low and high quality parts from its vendors.[31] For instance, this may be due to a segregated market i.e. MTL might be selling two different qualities of tractors under the same label and with the demand from customers who value quality less being more stochastic. However, this seems quite unlikely since MTL can only sell tractors at uniform prices predetermined through negotiations with the government and it is implausible that it can consistently "fool" less quality conscious consumers to buy inferior tractors at the same price. Also, MTL could be employing non-price mechanisms of discriminating among different types of customers but we find no evidence for any such mechanisms.

Second, even if we were to rely on the role of specific investments, can one build an explanation without making the somewhat unconventional assumption that these investments hurt outside options and assume instead that the investment and the type of a vendor are substitutes *within* the relationship? Following the simple extension of the basic model in Remark 1, let us assume that $\delta > 0$ and $u = 0$ (the latter implying that investments do not affect outside options). It is straightforward to see that high types invest less than low types for the same orders as in our framework. However, the gap between the investment of the high type and the low type is now *decreasing* in the level of orders. This means, if low types are hired at all, they are more likely to be used as first-preference vendors which, as we have argued already, is inconsistent with our empirical findings.

Third and finally, the performance measures that we see are negatively correlated with the extent of specificity may not capture *ex ante* differences in quality and investment as we have argued, but *ex post* differences in behavior. That is, being more tied, receiving lower prices, and a low and fluctuating order schedule makes a vendor perform worse than he would otherwise. However, as we have argued above, it is plausible that greater specific investments should lead to

[30] This assumes that the different sources of demand they face are not perfectly correlated.

[31] Assuming that MTL's quality control process cannot perfectly detect all defective parts and MTL is aware that a low quality vendor is more likely to supply defective parts that make their way into the final product.

greater efficiency. Indeed we do observe lower costs for tied vendors. Also, if vendors are of the same quality ex ante, then we would expect those who receive a higher and more stable level of orders to invest more, not less which is exactly the opposite of what the evidence suggests.

Alternatively one may argue that untied vendors have greater exposure to changes in the outside market which facilitates learning and innovation. These considerations are likely to have limited relevance in our context, since a supplier has to follow blueprints obtained from Massey-Ferguson which generally do not change.[32] Even if one does accept that tying might affect the quality of parts supplied (i.e. the proportion of rejections) it should not affect the responsiveness of the vendor (i.e., the problems of undersupplying and oversupplying). If anything, their outside options being worse, it would be cheaper for MTL to give them incentives to behave.

However, even if we accept the argument that greater specificity and the contractual package that comes with it makes a vendor perform worse, it raises a more serious question - if two vendors have the same *ex ante* quality why would one of them tie himself to MTL if doing so clearly makes it worse off in all respects? Different attitudes towards risk cannot explain this, because if anything, tying exposes a vendor to greater risk by putting all its eggs in the same basket. If MTL faced a stable pattern of demand, and in turn gave a steady flow of orders to a tied vendor, then this would not be an issue, but neither happens to be the case. An alternative explanation would be that MTL provides assistance to the tied vendors in other forms, for example, loans. The limited evidence we have on this issue does not support this explanation.[33]

6 Conclusion

The relationship between a tractor assembling firm in Pakistan and its subcontractors offers important insights about contracting and asset specificity within a manufacturer-supplier network. The presence of demand uncertainty makes undertaking relationship-specific investments costly on the part of suppliers. This cost is likely to be more, the more able and versatile the supplier. Therefore there is a chance for low quality suppliers to survive because of their greater willingness to undertake specific investments. This explains the puzzle we observe in the data that the assembler treats suppliers with greater asset-specificity as marginal suppliers even though they are cheaper. Our

[32] Unlike cars, models of tractors do not change very often.

[33] The LUMS survey asked each vendor whether MTL provided any financial support to the vendor. We find that vendors received very little financial support and there was no correlation between financial assistance received and the extent of specificity.

work suggests that asset specificity should not always be viewed as purely technology-driven which has been the dominant view in the organizations literature. It could also capture heterogeneity among firms and this implies one has to be careful in interpreting the effect of asset-specificity on contractual and performance outcomes. More broadly, the contracting relationships in the kind of network setting considered in this paper raises many interesting theoretical and empirical questions which we hope will be taken up in future research.

References

[1] Amir, I., K. Isert, and S. Khan (1995): "Note on the Local Suppliers to the Tractor Industry in Pakistan", Mimeo. Lahore Institute of Management Studies.

[2] Asanuma, B. (1989): "Manufacturer-Supplier Relationships in Japan and the Concept of Relationship-Specific Skill", *Journal of the Japanese and International Economies*, 3, 1-30.

[3] Baker, G. P. and T. N. Hubbard (2001): "Contractibility and Asset Ownership: On-Board Computers and Governance in U.S. Trucking", Mimeo. Harvard Business School and University of Chicago Graduate School of Business.

[4] Banerjee, A.V. and E. Duflo (2000): "Reputation Effects and the Limits of Contracting: A Study of the Indian Software Industry", *Quarterly Journal of Economics*.

[5] Banerjee, A.V. and K. Munshi (2000): "Networks, Migration, and Investment: Insiders and Outsiders in Tirupur's Production Cluster", MIT Department of Economics Working Paper No. 00-08.

[6] Chiappori, P.A., and B. Salanie (2001): "Testing Contract Theory: A Survey of Some Recent Work", Invited Lecture, World Congress of the Econometric Society, Seattle.

[7] Coase, R. (1937): "The Nature of the Firm", *Economica*, 4, 386-405.

[8] German, L. and R. Roth (1997): "Parts Manufacturing in Developing Economies", IMVP Working Paper #w-0174a, M.I.T.

[9] Ghani, J. (1997): "Inter-Firm Linkages in Pakistan's Auto Industry", Mimeo. Lahore Institute of Management Studies.

[10] Gonzales, M., B. Arrunada, and A. Fernandez (1999) : "Causes of Subcontracting : Evidence from Panel Data on Construction Firms", Mimeo, University of Oviedo, Spain and Pompeu Fabra University.

[11] Holmström, B. and J. Roberts (1998): "The Boundaries of the Firm Revisited", *Journal of Economic Perspectives*, Fall, 12 (4), p.73-94.

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

[13] Joskow, P. (1987): "Contract Duration and Transactions Specific Investment: Empirical Evidence from Coal Markets," *American Economic Review.*

[14] Klein, B., R. Crawford, and A. Alchian (1978): "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process", *Journal of Law and Economics*, 21(2), pp. 297-326.

[15] Kranton, R.E. and D. F. Minehart (2000): "Networks versus Vertical Integration", *Rand Journal of Economics*, Vol. 31, No. 3.

[16] Lyons, B.R. (1994): "Contracts and Specific Investment : An Empirical Test of Transaction Cost Theory." *Journal of Economics and Management Strategy.*

[17] McMillan, J. and C. Woodruff (1999) : "Interfirm relationships and informal credit in Vietnam", *Quarterly Journal of Economics*, pp. 1285-1320.

[18] McMillan, J. (1995): "Reorganizing Vertical Supply Relationships", in *Trends in Business Organizations*, Ed. Horst Siebert, Tubingen: J.C.B. Mohr.

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

[20] Mookherjee, D. (1999) : "Contractual Constraints on Firm Performance in Developing Countries", Mimeo. Boston University.

[21] Piore, M. J. and C. Sabel (1984): *The Second Industrial Divide.* New York: Basic Books.

[22] Rajan, Raghuram and Luigi Zingales [1998], "Power in a Theory of the Firm", *Quarterly Journal of Economics,* 113, 387–432.

[23] Shelanski, H. A. and P. G. Klein (1995) : "Empirical Research in Transaction Cost Economics: A Review and Assessment". *Journal of Law, Economics, and Organization*, Vol.11, No. 2, pp. 335-61.

[24] Whinston, M. W. (2002). "On the Transaction Cost Determinants of Vertical Integration", mimeo, Northwestern University.

[25] Williamson, O. (1985): *The Economic Institutions of Capitalism,* New York: The Free Press.

[26] Woodruff, C. (2002): "Non-contractible investments and vertical integration in the Mexican footwear industry", *International Journal of Industrial Organization.*

7 Appendix

Proof of Result 2: We prove this in two steps. First we show that for $\beta \in [0, \hat{\beta}(\theta)], s^*(\beta, \theta)$ is an increasing and linear function of β. For $\beta \in [\hat{\beta}(\theta), 1]$, $s^*(\beta, \theta)$ is an increasing and strictly convex function of β. Next we show that given this property, the assembler would give one vendor a certain order of 1, and the other vendor an order of 1 with probability α and 0 with probability $1 - \alpha$.

Step1: We know that $s^*(\beta, \theta) = \beta(a + b\theta - 1) + (1 - \beta)\theta + \beta\gamma x^*(\beta, \theta) - (1 - \beta)u\theta x^*(\beta, \theta) - \frac{1}{2}x^*(\beta, \theta)^2 - \theta$. Notice that $\beta\gamma x^*(\beta, \theta) - (1-\beta)u\theta x^*(\beta, \theta) - \frac{1}{2}x^*(\beta, \theta)^2 = \{\beta\gamma - (1 - \beta)u\theta\} x^*(\beta, \theta) - \frac{1}{2}x^*(\beta, \theta)^2 = \frac{1}{2}x^*(\beta, \theta)^2$ using (1). Therefore, the expected joint surplus function can be written as

$$\begin{aligned} s^*(\beta, \theta) &= \beta\{a + (b - 1)\,\theta - 1\} \text{ for } \beta \leq \hat{\beta}(\theta) \\ &= \beta\{a + (b - 1)\,\theta - 1\} + \frac{1}{2}\{\beta\gamma - (1 - \beta)u\theta\}^2 \text{ for } \beta \geq \hat{\beta}(\theta). \end{aligned}$$

For $\beta \leq \hat{\beta}(\theta)$, $\frac{\partial s^*}{\partial \beta} = (a - 1 + (b - 1)\,\theta)$ which is positive by Assumption 1 and independent of β. For $\beta \geq \hat{\beta}(\theta)$, by the envelope theorem, $\frac{\partial s_i}{\partial \beta} = S_i - \bar{u}^i(x_i^*(\beta, \theta), \theta) + \frac{\partial s_i}{\partial x_i}\frac{\partial x_i}{\partial \beta} = S_i - \bar{u}^i(x_i^*(\beta, \theta), \theta) = (a - 1 + (b - 1)\,\theta) + \{\beta\gamma - (1 - \beta)u\theta\}(\gamma + u\theta) > 0$ by Assumption 1. Differentiating with respect to β again we find that $s_i^*(\beta, \theta)$ is strictly convex in $\beta : \frac{\partial^2 s_i}{\partial \beta^2} = (\gamma + \theta u)^2 > 0$.

Step 2: By Step 1 the function $s^*(\beta, \theta)$ is always (weakly) convex and it is strictly convex for $\beta > \hat{\beta}(\theta)$. As $s^*(\beta) = \max\{s(\beta, \underline{\theta}), s(\beta, \bar{\theta})\}$, it too is a convex function of β. The expected level of demand facing the assembler is $1 + \alpha$. Let $\rho(1 + \alpha)$ be allocated to one vendor and $(1 - \rho)(1 + \alpha)$ be allocated to the other vendor subject to the capacity constraints, $\rho(1 + \alpha) \leq 1$ and $(1 - \rho)(1 + \alpha) \leq 1$, i.e., $\rho \in [0, \frac{1}{1+\alpha}]$. As $s^*(\beta)$ is convex in β, it directly follows from the property of convex functions that $\frac{1}{2}s^*(1) + \frac{1}{2}s^*(\alpha) \geq \frac{1}{2}s^*(\rho(1+\alpha)) + \frac{1}{2}s^*((1-\rho)(1+\alpha))$ (this is easy to see graphically), or, $s^*(1) + s^*(\alpha) \geq s^*(\rho(1+\alpha)) + s^*((1-\rho)(1+\alpha))$ for all $\rho \leq \frac{1}{1+\alpha}$.**QED**.

Proof of Result 3:

The net expected joint surplus consists of two components, one independent of investment, and one due to investment. The component $\beta\{a + (b-1)\,\theta - 1\}$ is independent of investment and the component $\frac{1}{2}\{\beta\gamma - (1-\beta)u\theta\}^2$ is due to investment. The latter component is relevant only when demand exceeds some threshold level, i.e., $\beta \geq \hat{\beta}(\theta)$. As far as the former component is concerned, the difference between a high type and a low type vendor is $\beta\,(b-1)\,\triangle\theta$, which positive, zero, or negative according as $b > 1, b = 1$ and $b < 1$. Since low type vendors always invest more than high type vendors, if $b \leq 1$ then high type vendors will never be strictly preferred. From Result 2 we know that the first-preference vendor is going to be given a certain order of 1 and the second-preference vendor the residual order α. In this case both these vendors would be low type vendors. Below we consider the case there $b > 1$.

We need to consider three ranges of values of β.

First, low ranges of orders, i.e., $\beta \leq \hat{\beta}(\underline{\theta})$. In this case $x^* = 0$ for both types and since $b > 1$ high types would be strictly preferred.

Second, intermediate ranges of orders, i.e., $\hat{\beta}(\underline{\theta}) < \beta \leq \hat{\beta}(\bar{\theta})$ where the high type vendor chooses a zero level of investment, but the low type vendor chooses a positive level of investment. The difference in surplus is now: $\beta\,(b-1)\,\triangle\theta - \frac{1}{2}\{x^*(\beta, \underline{\theta})\}^2$. Since $x^*(\beta, \underline{\theta})$ is increasing in β, the maximum value it can take in this range of β is at $\hat{\beta}(\bar{\theta}) = \frac{\bar{\theta}u}{\gamma+\bar{\theta}u}$, for which it takes the value $\frac{1}{2}\left(\frac{\bar{\theta}u}{\gamma+\bar{\theta}u}\gamma - \frac{\gamma}{\gamma+\bar{\theta}u}u\underline{\theta}\right)^2 = \frac{1}{2}\left(\frac{u\gamma\triangle\theta}{\gamma+\bar{\theta}u}\right)^2$. Comparing with $\beta\,(b-1)\,\triangle\theta$ at $\hat{\beta}(\bar{\theta}) = \frac{\bar{\theta}u}{\gamma+\bar{\theta}u}$ we see that in this range of β, low types would be preferred if $(b-1) < \frac{1}{2}\frac{u\gamma^2}{\gamma+\bar{\theta}u}\frac{\triangle\theta}{\bar{\theta}}$. Since $\frac{\triangle\theta}{\bar{\theta}} \leq 1$, $\gamma < 1$, and $\frac{u\gamma}{\gamma+\bar{\theta}u} < 1$ (because it is equivalent to the inequality $\gamma(1-u) + \bar{\theta}u > 0$) the right-hand side cannot exceed $\frac{1}{2}$. So there is a range of values of $\underline{b} \in (1, 1.5)$ such that low types will be chosen for this range of orders. If $b \geq \underline{b}$, high types will always be chosen.

Third, high ranges of orders, i.e., $\beta > \hat{\beta}(\bar{\theta})$. Here $x^* > 0$ for all θ. As low type vendors invest more than high type vendors for the same β, they generate a higher surplus if we consider only the part of joint surplus that is due to investment. This difference is equal to $\frac{1}{2}\{\beta\gamma - (1-\beta)u\underline{\theta}\}^2 - \frac{1}{2}\{\beta\gamma - (1-\beta)u\bar{\theta}\}^2 = (1-\beta)u\{\beta\gamma - \frac{1}{2}(1-\beta)u\}\triangle\theta$ (using the normalization $\bar{\theta} + \underline{\theta} = 1$). Let us compare this with the difference in surplus if we consider the part of joint surplus that is independent of investment, i.e., $\beta\,(b-1)\,\triangle\theta$. It is clear that the higher is β the smaller is the term due to investment and the larger is the term independent of investment. For very high values of β, the high type vendors will always be preferred (the former term becomes zero and the latter term equals $(b-1)\triangle\theta > 0$). For values of β near the bottom of this interval, $\hat{\beta}(\bar{\theta})$, low types could be chosen using the argument in the previous paragraph, by continuity.

For the case $b > 1$, the high types are preferred for very high and very low levels orders. So the first-preference vendors are going to be high type. The second-preference vendor gets the residual order α. If α is very high or low, the second-preference vendor is going to be a high type vendor. If α is neither too high nor too low, and in addition $b \leq \underline{b}$, then low type vendors are going to be chosen as second-preference vendors. **QED.**

Proof of Result 4: By Result 2 one vendor gets a certain order of 1, and the other vendor gets an order of 1 with probability α. So if they are of the same type, by Result 1 the one with the certain order will invest more. So let us consider the remaining possibility where the first-preference vendor is a high type vendor and the second-preference vendor is a low type vendor. Because the first-preference vendor gets a certain order of 1, by (1) the level of investment is γ. The second-preference vendor, who gets an order of α, chooses an investment level of $\alpha\gamma - (1-\alpha)\underline{\theta}$ which is clearly less than γ. Since the unit cost of the part is decreasing in x, the remaining part of the result follows.**QED.**

Proof of Result 5: Parts (i)-(iii) follow upon inspecting (1) and (4). To prove (iv) let us calculate the difference in investment for the same order under the first-best and the second-best, $x^*(\beta,\theta) - x^{SB}(\beta,\theta) = \frac{\beta}{2}(\gamma + \theta u) > 0$. Since $x^*(\beta,\theta) - x^{SB}(\beta,\theta)$ is increasing in θ, the underinvestment problem is more serious for the high type vendors. The last part directly follows from:

$$x^{SB}(\beta,\underline{\theta}) - x^{SB}(\beta,\bar{\theta}) = \left(1-\frac{\beta}{2}\right)u\triangle\theta \qquad \text{for } 1 \geq \beta \geq \beta_1. \tag{7}$$

$$= \left\{\frac{\beta}{2}\gamma - (1-\frac{\beta}{2})\underline{\theta}u\right\} \qquad \text{for } \beta_1 \geq \beta \geq \beta_0. \tag{8}$$

$$= 0 \qquad \text{for } \beta_0 \geq \beta \geq 0. \tag{9}$$

QED.

Proof of Result 6: First we show that for $x^{SB}(\beta,\theta) = 0$, $\Pi(\beta,\theta)$ does not change as β changes, and for $x^{SB}(\beta,\theta) > 0$, $\Pi(\beta,\theta)$ is a linear and increasing function of β. The first part of the statement follows immediately from the fact that the only way β affects $\Pi(\beta,\theta)$ is through

$x^{SB}(\beta,\theta)$. Now consider the case $x^{SB}(\beta,\theta) > 0$. Using (4) and (5) we find that for $\beta \geq \beta_0$, $\frac{\partial \Pi(\beta,\underline{\theta})}{\partial \beta} = (\gamma + \underline{\theta}u)\frac{\partial x^{SB}(\beta,\underline{\theta})}{\partial \beta} = \frac{1}{2}(\gamma+\underline{\theta}u)^2$. Similarly, for $\beta \geq \beta_1$, $\frac{\partial \Pi(\beta,\bar{\theta})}{\partial \beta} = (\gamma + \bar{\theta}u)\frac{\partial x^{SB}(\beta,\bar{\theta})}{\partial \beta} = \frac{1}{2}(\gamma+\bar{\theta}u)^2$. Therefore $\frac{\partial \Pi(\beta,\theta)}{\partial \beta}$ is positive and independent of β for $x^{SB}(\beta,\theta) > 0$. This implies that the function $\Pi(\beta,\theta)$ is a (weakly) convex function of β. Next, observe that as $\Pi^*(\beta) = \max\{\Pi(\beta,\underline{\theta}), \Pi(\beta,\bar{\theta})\}$, it too is a convex function of β. Given that Π^* is increasing in β, this immediately implies that $\beta\Pi^*(\beta)$ is convex in β. The rest of the proof is identical to that of Result 3 if we replace the function $s^*(\beta,\bar{\theta})$ with $\beta\Pi^*(\beta)$.**QED.**

Proof of Result 7: Consider $\triangle\Pi(\beta) \equiv \Pi(\beta,\bar{\theta}) - \Pi(\beta,\underline{\theta})$. It readily follows from (7)-(8), and (5) (and the normalization that $\underline{\theta}+\overline{\theta} = 1$) that

$$\triangle\Pi(\beta) = \frac{1}{2}(b-1)\triangle\theta - \frac{1}{2}u\left\{(1-\beta)\gamma + (1-\frac{\beta}{2})u\right\}\triangle\theta \qquad \text{for } 1 \geq \beta \geq \beta_1. \tag{10}$$

$$= \frac{1}{2}(b-1)\triangle\theta - \frac{1}{2}(\gamma+\underline{\theta}u)\left\{\frac{\beta}{2}\gamma - (1-\frac{\beta}{2})\underline{\theta}u\right\} \qquad \text{for } \beta_1 \geq \beta \geq \beta_0. \tag{11}$$

$$= \frac{1}{2}(b-1)\triangle\theta \qquad \text{for } \beta_0 \geq \beta \geq 0. \tag{12}$$

If $b \leq 1$, low types would be preferred since $\triangle\Pi(\beta) < 0$ for all β. So let us focus on the case where $b > 1$. From (10)-(12), we know that $\triangle\Pi(\beta)$ is constant for $[0,\beta_0]$, decreasing for $[\beta_0,\beta_1]$ and increasing for $[\beta_1, 1]$. Given that $\triangle\Pi(\beta)$ is continuous it therefore directly follows that $\triangle\Pi(\beta)$ attains its minimum value at $\beta = \beta_1$. For the same reason, for $\beta \in [\beta_0, 1]$, $\triangle\Pi(\beta)$ reaches a maximum value at $\beta = 1$. This implies that a necessary condition for $\triangle\Pi(\beta) < 0$ for *some* $\beta \in [0,1]$ is $\triangle\Pi(\beta_1) < 0$. Now $\triangle\Pi(\beta_1) = \frac{1}{2}\left[(b-1) - \gamma u\frac{\gamma+\underline{\theta}u}{\gamma+\bar{\theta}u}\right]\triangle\theta$. It is straightforward to check that $\gamma u\frac{\gamma+\underline{\theta}u}{\gamma+\bar{\theta}u} > \frac{1}{2}\frac{u\gamma^2}{\gamma+\bar{\theta}u}\frac{\triangle\theta}{\underline{\theta}}$. Therefore, the threshold level of b for low types not to be preferred is higher than the one under the first-best.

Finally, a necessary condition for $\triangle\Pi(\beta) > 0$ for some $\beta \in [\beta_0, 1]$ is $\triangle\Pi(1) = \frac{\triangle\theta}{2}\{(b-1) - \frac{1}{2}u^2\}$. It is straightforward to check that $\frac{1}{2}u^2 < \gamma u\frac{\gamma+\underline{\theta}u}{\gamma+\bar{\theta}u}$. This means there is a critical value of b which is less than $\bar{b}$ but higher than 1 such that if b is less than this critical value, a low type vendor would be preferred for $\beta = 1$, and given the property of $\triangle\Pi(\beta)$, for all $\beta > \beta_0$. However, so long as both types of vendors do not invest (i.e., $\beta \leq \beta_0$) a high type vendor would be preferred since by assumption $b > 1$.**QED.**

Proof of Result 8: If a high type vendor is the first-preference vendor, his investment level is $\frac{1}{2}\gamma - \frac{1}{2}\bar{\theta}u$, and if the second-preference vendor is a low type vendor then his investment level is

$\frac{\alpha}{2}\gamma - (1-\frac{\alpha}{2})\underline{\theta}u$. The difference between the two investment levels is:

$$\frac{1-\alpha}{2}\gamma - \left\{\frac{1}{2}\bar{\theta} - (1-\frac{\alpha}{2})\underline{\theta}\right\} u = \frac{1-\alpha}{2}\gamma - \left(\frac{1}{2}\triangle\theta - \frac{1-\alpha}{2}\underline{\theta}\right) u.$$

By Assumption 2, $\gamma > u$, but so long as $\bar{\theta}$ is high enough relative to $\underline{\theta}$ this expression could be negative. **QED.**

Table 1. Summary Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Vendor Characteristics:[1]					
Vendor Specificity (%)	18[2]	43	43	0	100
Vendor Age (in 1995)	18	15	7	3	34
Vendor Size (Number of employees)	19	72	128	4	550
Distance to MTL (km)	19	334	548	7	1400
City (1=Karachi)	19	0.21	0.42	0	1
Part Order data:					
Annual[3] Average Part price (P)	273	243	741	0.2	4188
Log Annual Part Price	273[4]	3.30	2.20	-1.61	8.34
Quarterly Vendor Scheduled Quantity (Q_S)	758	2350	3750	0	31500
Quarterly Total Scheduled Quantity (TQ_S)	907	4601	6723	0	53100
Quarterly Vendor Received Quantity (Q_R)	790	3186	5383	0	81242
Quarterly Total Received Quantity (TQ_R)	907	6161	9580	0	117026
Log Quarterly Vendor Scheduled Q	558	7.55	0.96	4.61	10.36
Log Quarterly Total Scheduled Q	740	8.19	0.92	5.30	10.88
Log Quarterly Vendor Received Q	628	7.69	1.13	1.61	11.31
Log Quarterly Total Received Q	819	8.28	1.04	5.01	11.67
Quarterly Vendor Scheduled Q – Quarterly Vendor Received Q	802	-399	4253	-24,888	25,812
Undersupply[5] (absolute value)	390	1569	3347	0	25812
Oversupply[6]	449	2231	3914	0	24888
Vendor Scheduled Q Coefficient of Variation	56	0.99	0.38	0.49	2
Log Cost (log C)	64[7]	4.53	1.43	2.15	8.37
Proportion of Rejections[8] (R)	411	.01	.027	0	.29

Sources: Lahore University of Management Sciences (LUMS) survey 1997, Millat Tractors Limited (MTL) database.

Table 2: Vendor Attribute Correlations (Restricted to Multiple Vendors)

	Specificity	Age	Size	Distance	City
Specificity	1.00				
Age	-0.18	1.00			
Size	-0.33	0.58**	1.00		
Distance	0.34	0.51**	0.44*	1.00	
City	0.55**	0.09	-0.01	0.78***	1.00

*Significant at 10%, ** at 5%, *** at 1%

[1] The sample is restricted to only those parts for which we have multiple vendors in our sample. Henceforth, this restriction is referred to as "multiple vendors".
[2] In the multiple vendor list we have 19 vendors, but for one vendor data on specificity and age are missing.
[3] All annual data refers to the calendar year (as opposed to fiscal) year.
[4] The price and quantity data have different number of observations since they come from different MTL sources. The former comes from the Order Information Database and the latter from the Schedule-Receipts Database.
[5] Sample restricted to non-negative values of scheduled quantity minus received quantity for a vendor.
[6] Sample restricted to non-positive values of scheduled quantity minus received quantity for a vendor.
[7] The cost data comes from a smaller data-set which had internal "cost" estimates made my MTL engineers. These estimates are not accounting costs but also include a standard markup.
[8] The rejections data is only available for the years 2000 and 2001. Proportion of rejections is defined as the number of rejected parts divided by the total number of parts supplied and is defined per quarter.

Table 3: Determinants of log Cost Estimate

	(1)	(2)
Specificity	-0.0010***	-0.0016***
	(0.0002)	(0.0001)
Age	-0.0038**	-0.0035***
	(0.0012)	(0.0010)
Size	-0.0004	-0.0004**
	(0.0003)	(0.0002)
Distance		0.0022**
		(0.0008)
City		0.0841***
		(0.0177)
Fixed Effects	Part***	Part***
	Year***	Year***
Observations	61	61
R-squared	0.9980	0.9980

Table 4. Determinants of Average Log Annual Price

Variables	(1)	(2)
Specificity	-0.0007***	-0.0018***
	(0.0001)	(0.0006)
Age	-0.0071*	-0.0086**
	(0.0037)	(0.0034)
Size	-0.0003	-0.0005
	(0.0008)	(0.0009)
Distance[9]		0.0002
		(0.0010)
City		0.0988*
		(0.0501)
Fixed Effects	Part***	Part***
	Year***	Year***
Observations	273	273
R-squared	0.99	0.99

Robust Standard errors in parentheses
Errors clustered at the vendor level
***Significantly different from zero at 1%
**Significantly different from zero at 5%
* Significantly different from zero at 10%

[9] In all tables, the variable Distance is set equal to 0 if the vendor is in Karachi since there is no variation in distance for Karachi vendors. Thus together with the City dummy (=1 if vendor in Karachi), Distance should henceforth be interpreted as the Distance of non-Karachi vendors (i.e. "local distance").

Table 5. Determinants of log Scheduled Quantity

Variables	(1)	(2)
Specificity	-0.0013**	-0.0034
	(0.0006)	(0.0019)
Log Total Scheduled Q	0.7619***	0.7661***
	(0.0284)	(0.0281)
LogAge	0.2125**	0.1409
	(0.0951)	(0.1150)
LogSize	0.0658	0.0002
	(0.0386)	(0.0257)
LogDistance		-0.3142***
		(0.0379)
City		-0.5891***
		(0.1595)
Fixed Effects	Part***	Part***
	Quarter ***	Quarter ***
Observations	558	558
R-squared	0.90	0.91

Table 6. Determinants of Scheduled Quantity Coefficient of Variation

Variables	(1)	(2)
Specificity	0.0017*	0.0013
	(0.0008)	(0.0026)
Age	-0.0118	-0.0149
	(0.0160)	(0.0118)
Size	0.0015	0.0046*
	(0.0029)	(0.0025)
Distance		0.0122***
		(0.0039)
City		0.1835
		(0.2090)
Fixed Effects	Part	Part
Observations	55	55
R-squared	0.91	0.94

Robust Standard errors in parentheses
Errors clustered at the vendor level
***Significantly different from zero at 1%
**Significantly different from zero at 5%
* Significantly different from zero at 10%

Table 7. Determinants of Proportion of Rejections[10]

Variables	(1)	(2)
Specificity	0.00002	0.00011
	(0.00001)	(0.00002)***
Size	-0.00008	-0.00003
	(0.00002)***	(0.00001)**
Age	0.00139	0.00152
	(0.00020)***	(0.00023)***
Distance		0.00049
		(0.00032)
City		-0.00292
		(0.00239)
Fixed Effects	Part***	Part***
	Quarter	Quarter
Observations	397	397
R-squared	0.39	0.39

Table 8. Determinants of log Scheduled Quantity – Total Schedule Elasticity

Variables	(1)	(2)
Specificity	-0.0094**	-0.0111**
	(0.0031)	(0.0038)
Log Total Scheduled Q	0.7337***	0.7392***
	(0.0340)	(0.0329)
Log Total Scheduled Q*Specificity	0.0010**	0.0009**
	(0.0003)	(0.0003)
LogAge	0.1688	0.0990
	(0.1029)	(0.1191)
LogSize	0.0470	-0.0173
	(0.0352)	(0.0217)
LogDistance		-0.3086***
		(0.0353)
City		-0.5723***
		(0.1555)
Fixed Effects	Part***	Part***
	Quarter ***	Quarter***
Observations	558	558
R-squared	0.90	0.91

Robust Standard errors in parentheses
Errors clustered at the vendor level
***Significantly different from zero at 1%
**Significantly different from zero at 5%
* Significantly different from zero at 10%

[10] Recall that the rejections data is for the years 2000 and 2001.

Table 9. Determinants of log Received Quantity – Own Schedule Elasticity

Variables	(1)	(2)
Specificity	0.0124**	0.0025
	(0.0052)	(0.0034)
Log Scheduled Q	0.4434***	0.4744***
	(0.0918)	(0.1074)
Log Scheduled Q*Specificity	-0.0017**	-0.0020***
	(0.0006)	(0.0005)
LogAge	-0.1447	-0.4060***
	(0.2033)	(0.1217)
LogSize	0.1523*	0.0509
	(0.0706)	(0.0507)
LogDistance		0.2452
		(0.1527)
City		1.7908***
		(0.3505)
Fixed Effects	Part***	Part***
	Quarter***	Quarter***
Observations	479	479
R-squared	0.85	0.86

Table 10. Determinants of log Received Quantity – Total Schedule Elasticity

Variables	(1)	(2)
Specificity	0.0155***	0.0122
	(0.0033)	(0.0072)
Log Total Scheduled Q	0.2931***	0.2892***
	(0.0917)	(0.0869)
Log Total Scheduled Q*Specificity	-0.0022***	-0.0019***
	(0.0004)	(0.0005)
LogAge	0.4068	0.4430
	(0.2633)	(0.4927)
LogSize	0.2703**	0.3773
	(0.1112)	(0.2441)
LogDistance		1.1016
		(0.7236)
City		2.8665**
		(0.9397)
Fixed Effects	Part***	Part***
	Quarter***	Quarter***
Observations	518	518
R-squared	0.78	0.79

Robust Standard errors in parentheses
Errors clustered at the vendor level
***Significantly different from zero at 1%
**Significantly different from zero at 5%
* Significantly different from zero at 10%

Table 11. Determinants of Undersupply[11]
(Sample Restricted to Non-Negative values of Undersupplying)

Variables	(1)	(2)
Specificity	3.9253***	8.9634***
	(0.7187)	(1.8757)
Age	33.6682***	37.4110***
	(8.2361)	(9.2523)
Size	5.5829	9.3591**
	(3.3875)	(3.2054)
Distance		13.2401*
		(6.5339)
City		-17512.3531*
		(8489.4411)
Fixed Effects	Part***	Part***
	Quarter***	Quarter***
Observations	378	378
R-squared	0.63	0.63

Table 12. Determinants of Oversupply[12]
(Sample Restricted to Non-Negative values of Oversupplying)

Variables	(1)	(2)
Specificity	5.8274***	-7.1748
	(1.6637)	(6.1840)
Age	-47.9560	-63.3999**
	(31.2845)	(22.0871)
Size	6.4638	0.9490
	(5.8287)	(6.4147)
Distance		-18.5537
		(21.7324)
City		25090.6336
		(28211.0151)
Fixed Effects	Part***	Part***
	Quarter***	Quarter***
Observations	443	443
R-squared	0.60	0.60

Robust standard errors in parentheses
Errors clustered at the vendor level
***Significantly different from zero at 1%
**Significantly different from zero at 5%
* Significantly different from zero at 10%

[11] The LHS variable is quarterly vendor scheduled quantity minus quarterly vendor received quantity for cases where the former is greater or equal to the latter. Thus a positive coefficient on Specificity means that Tied vendors tend to under-supply.

[12] The LHS variable is the absolute value of quarterly vendor scheduled quantity minus quarterly vendor received quantity for cases where the former is less than or equal to the latter. Thus a positive coefficient on Specificity means that Tied vendors tend to over-supply.

Table 13. Dynamic Quantity Incentives – Undersupply
LHS = Quarterly Vendor Schedule Quantity
Sample Restricted to Non-Negative values of Lagged Undersupply

Variables	(1)	(2)
Specificity	-1.0743	-5.9205***
	(0.6176)	(1.1330)
Lagged |Scheduled – Received Q|	-0.2592***	-0.2582***
	(0.0403)	(0.0409)
Lagged |Scheduled – Received Q| * Specificity	0.0016***	0.0016***
	(0.0002)	(0.0002)
Quarterly Total Scheduled Q	0.5015***	0.5021***
	(0.1033)	(0.1037)
Age	6.1991	3.0475
	(7.2237)	(7.1556)
Size	6.7414**	4.1087
	(2.4998)	(2.5562)
Distance		-10.8352**
		(4.9315)
City		14394.2450**
		(6372.4664)
Fixed Effects	Part***	Part***
	Quarter***	Quarter***
Observations	306	306
R-squared	0.90	0.91

Table 14. Dynamic Quantity Incentives – Oversupply
LHS = Quarterly Vendor Schedule Quantity
Sample Restricted to Non-Negative values of Lagged Oversupply

Variables	(1)	(2)
Specificity	-1.9201	5.1422**
	(1.3111)	(1.8433)
Lagged |Scheduled – Received Q|	0.0643***	0.0652***
	(0.0107)	(0.0102)
Lagged |Scheduled – Received Q| * Specificity	-0.0010***	-0.0010***
	(0.0002)	(0.0002)
Quarterly Total Scheduled Q	0.4501***	0.4507***
	(0.0097)	(0.0097)
Age	56.4518***	62.8813***
	(15.1670)	(6.0068)
Size	5.6034	5.9359
	(4.0258)	(3.4423)
Distance		-17.1747
		(9.4875)
City		21463.3106
		(12123.6698)
Fixed Effects	Part***	Part***
	Quarter***	Quarter***
Observations	357	357
R-squared	0.91	0.91

Robust standard errors in parentheses
Errors clustered at the vendor level
***Significantly different from zero at 1%
**Significantly different from zero at 5%
* Significantly different from zero at 10%

Appendix

Table A1: Part Description and Sources
(Restricted Sample – Multiple Vendors)

	Part Description	Parts per Tractor	Vendor IDs (Specificity %)		
			Vendor A	Vendor B	Vendor C
1	ROD & YOKE ASSEMBLY	2	10 (0)	2 (80)	
2	PIN CLEVIS	8	10 (0)	2 (80)	
3	PIN COTTER	9	10 (0)	2 (80)	
4	WASHER	2	10 (0)	2 (80)	
5	ROD & YOKE ASSEMBLY2[13]	2	10 (0)	2 (80)	
6	SPRING	2	10 (0)	2 (80)	
7	CLIP	2	10(0)	2 (80)	
8	NUT BRAKE ROD	4	10 (0)	2 (80)	
9	WASHER BRAKE ROD	4	10 (0)	2 (80)	
10	TRUNION THREADED	4	10 (0)	2 (80)	
11	YOKE BRAKE ROD	4	10 (0)	2 (80)	
12	BRAKE ROD ASSEMBLY	2	10 (0)	2 (80)	
13	SUPPORT	2	10 (0)	2 (80)	
14	TRANSMISSION CASE	1	15 (100)	8 (0)	
15	HOUSING FRONT PTO B.	1	4 (100)	6 (0)	
16	FORK CLUTCH RELEASE	1	15 (100)	17 (0)	
17	HUB FRONT WHEEL	2	15 (100)	17 (0)	3 (100)
18	WATER PUMP	1	12 (0)	4 (100)	
19	ENGINE SUMP	1	15 (100)	8 (0)	
20	CYLINDER BLOCK	1	15 (100)	8 (0)	
21	IDLER GEAR	1	15 (100)	5 (30)	
22	GEAR CAM SHAFT	1	15 (100)	5 (30)	
23	GEAR FUEL PUMP	1	15 (100)	5 (30)	
24	NUMBER PLATE ASSY	1	10 (0)	2 (80)	
25	UPPER INST. PANEL[14]	1	7 (50)	16 100)	
26	LOWER INST. PANEL	1	7 (50)	16 (100)	
27	UPPER INST. PANEL 2	1	7 (50)	16 (100)	
28	LOWER INST. PANEL 2	1	7 (50)	16 (100)	
29	BATTERY BRACKET	1	18 (40)	16 (100)	
30	FRONT GRILL ASSEMBLY	1	18 (40)	9 (40)	
31	BATTERY BOX	1	18 (40)	16 (100)	
32	PIN CLEVIS2	4	3 (100)	10 (0)	
33	BRACKET TOP LINK	1	13 (100)	16 (100)	
34	FAN BLADE	1	2 (80)	1 (0)	
35	FAN BLADE2	1	2 (80)	1 (0)	
36	HINGE PIN ASSEMBLY	1	3 (100)	10 (0)	2(80)
37	PRE CLEANER	1	1 (0)	11 (1)	
38	PRE CLEANER + TUBE	1	1 (0)	11 (1)	
39	TOOL BOX	1	14 (100)	19 (30)	

[13] The number 2 is used to distinguish the two different tractor models that MTL produces.
[14] Inst. stands for instrument.

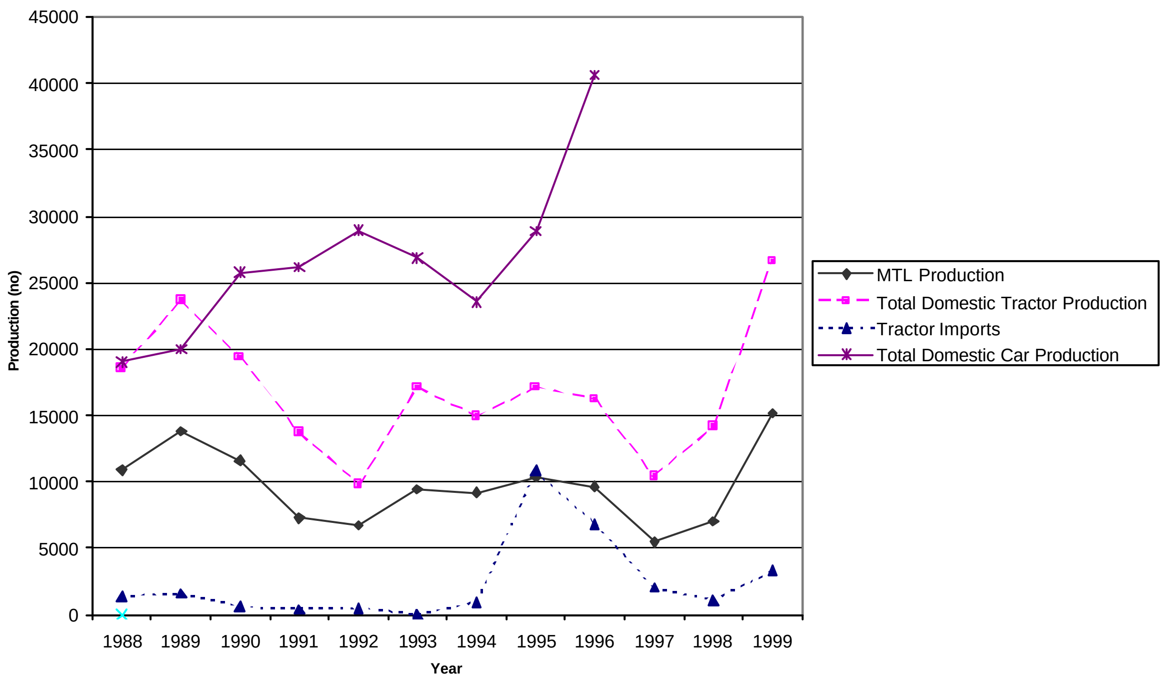
Table A2. Determinants of Over and UnderSupply[15]

Variables	(1)	(2)
Specificity	5.0446***	-1.9539
	(0.8359)	(1.6361)
Age	21.0951	12.7638
	(12.8636)	(7.4025)
Size	6.2366*	3.5811*
	(3.0064)	(1.6741)
Distance		-6.9481
		(5.8300)
City		565.2848***
		(110.1184)
Fixed Effects	Part***	Part***
	Quarter***	Quarter***
Observations	741	741
R-squared	0.58	0.58

Robust standard errors in parentheses
Errors clustered at the vendor level
***Significantly different from zero at 1%
**Significantly different from zero at 5%
* Significantly different from zero at 10%

Figure 1 : Pakistan Tractor and Automobile Industry

Source: Economic Survey of Pakistan,2000
Govt. of Pakistan, Ministry of Finance, Islamabad.
MTL. Annual Reports



[15] The LHS variable is the absolute value of quarterly vendor scheduled quantity minus quarterly vendor received quantity.

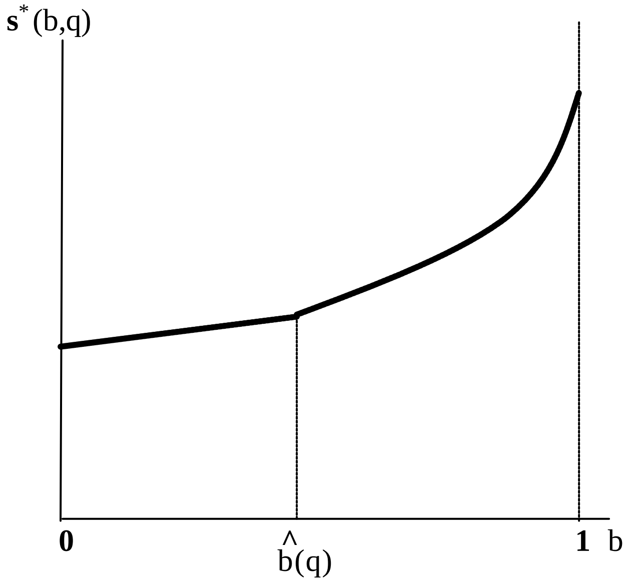


Figure 2

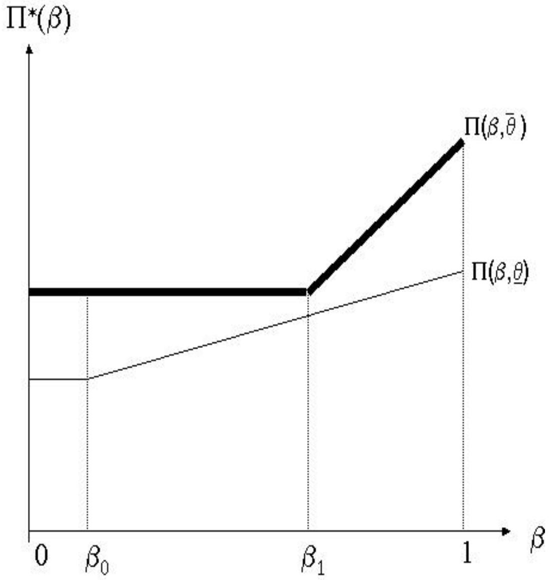


Figure 3 (a)

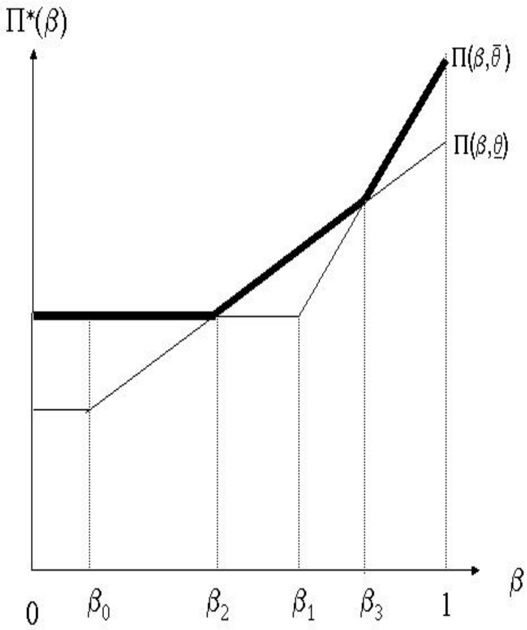


Figure 3 (b)

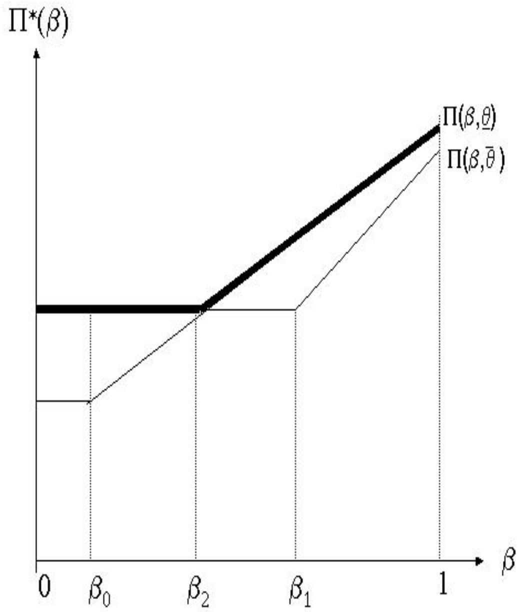


Figure 3 (c)