Giorgio ZANARONE

Fiat without Authority under Vertical Integration

This version, October 2008

Abstract

This paper develops a relational contracting model to show that fiat — a principal’s ability to dictate her agent’s performance — emerges in equilibrium under vertical integration, even when integration does not allocate distinctive formal authority to the principal. In a vertical structure, an efficient relational contract requires downstream managers to take actions that maximize aggregate profits, in exchange for future rents. If the manager of a vertically integrated unit reneges, she benefits from greater free time, but does not appropriate the associated increase in unit profits. Therefore, when the actions that maximize aggregate profits and the individual unit’s profits differ substantially — that is, when interest conflicts and spillovers between units are large — a manager’s promise to perform will be more credible under vertical integration than under separation.

Keywords: Ownership, Relational Contracts, Spillovers, Vertical Integration.

JEL codes: D23; L14; L22

* Universitat Pompeu Fabra; E-mail: giorgio.zanarone@upf.edu. I am grateful to Benito Arruñada and Robert Gibbons for continuous advice and encouragement, and to Gregory Corcos, Silke Forbes-Januszewski, Luis Garicano, Ricard Gil, Anna Grandori, Paolo Mariti, Eduardo Melero, Antonio Nicita, Emmanuel Raynaud and Dean Williamson for their comments and suggestions.
1. Introduction

In the last two decades a strong body of empirical evidence has emerged, suggesting that units at different stages in the chain of production tend to be vertically integrated when the potential interest conflicts between them are strong. For instance, franchisors own retail outlets that generate spillovers on the common brand (Brickley and Dark (1987), Lafontaine and Shaw (2005), Yeap (2006), Arruñada et al. (2008)), motor carriers own trucks whose poor maintenance would harm the carrier’s service and reputation (Nickerson and Silverman (2003)), and airline companies own regional carriers that serve routes between bad weather airports, in which frequent flight rescheduling preserves the network’s reputation but causes short-term losses to the regional (Forbes and Lederman (2008)).

These industries are all plagued by an agency problem: upstream principals (franchisors, motor carriers, airline companies) rely on downstream agents (franchisees, truck drivers, managers of regional airlines) to perform onerous tasks (serving customers, driving efficiently, adapting flight schedules under time constraints), whose benefits are shared between the upstream and downstream units. Importantly, agency is present independent of whether the units are vertically integrated or separated — for instance, a car manufacturer needs the retail outlets’ managers to implement service standards and spend sales effort both when these are franchisees, and when they are salaried employees. The question then is: why do agency problems conduce to vertical integration? What exactly does integration do to solve them?

---

1 See Lafontaine and Slade (1997, 2007) for detailed reviews of the empirical literature on vertical integration.
Standard incentive theories (Lutz (1995), Grossman and Hart (1986), Hart (1995)) are not well positioned to answer this question, as they predict that managers in vertically integrated firms, who do not appropriate the residual value of the assets they manage (Krueger (1991), Maness (1996)), have scarcer incentives to spend time and effort in production than if they owned the assets.²

As a possible explanation, transaction cost theory has suggested that agency conflicts disappear in vertically integrated firms, where employers use their power of fiat to direct employees (Coase (1937), Williamson (1971, 2000)), thus “economizing on contracting costs” (Klein et al. (1978), p. 299).³ While suggestive, the assumption that vertical integration conveys power of fiat seems often counterfactual. First, some agents’ actions are hard to verify in court, so it is unclear how formal authority over them could be exerted in the first place. Second, formal authority can also (and perhaps more cheaply) be allocated via contracts between independent firms (Alchian and Demsetz (1972), Hart (2008)). Indeed, distribution contracts allocate to upstream firms formal authority on both their employees and franchisees (Hadfield (1990), Arruñada et al. (2001, 2005), Zanarone (2008, 2009)). Also, contracts between major and regional airlines allocate to the former the right to change flight schedules, independent of whether the regionals are vertically integrated or not (Forbes and Lederman (2008)).

² Although property rights models such as Grossman and Hart (1986) and Hart (1995) have been mainly applied to study the managers’ incentives to increase firm-specific human capital, they can also be used to study the managers’ incentives to spend effort in production. See Holmstrom and Milgrom (1994), Holmstrom (1999), and Gibbons (2005) for thorough discussions of this point.
³ See Hart and Holmstrom (2002) and Baker et al. (2008a) for formal models, in which integration transfers control over a firm’s decisions, eliminating agency problems altogether. Also, see Masten (1988) and Williamson (1991) for legal arguments according to which vertical integration increases a firm’s formal authority to dictate decisions to its managers.
Alternatively, multi-task agency theories have suggested that, by muting the agent’s incentives, vertical integration can avoid imbalances in her allocation of effort across tasks. The problem with these models is that they require specific assumptions, such as agent’s risk-aversion (Holmstrom and Milgrom (1994), Bai and Tao (2000)), or agent’s willingness to work up to a substantial amount without incentives (Holmstrom and Milgrom (1991)), which do not seem to fit many of the industries where agency conflicts lead to vertical integration. For instance, in retail distribution, managers of riskier outlets receive greater portions of outlet profits (Norton (1988), Martin (1988), Lafontaine (1992)), which is contrary to what one would expect if these managers were risk-averse. Also, it seems unlikely that truck drivers, whose job is burdensome and unpleasant, would be happy to drive at the speed and pace required by the motor carrier, without explicit incentives to do so.

This paper provides a novel explanation for why agency conflicts lead to vertical integration, which, unlike the multi-tasking and transaction cost arguments discussed before, does not involve agent risk-aversion, willingness to work without incentives, or transfers of formal authority. The proposition advanced here is that a vertically integrated firm may be able to dictate conflictive decisions to managers (fiat) not by exerting formal authority, but, rather, by making the managers’ informal promise to act in the firm’s interest more credible.\(^4\) This point is illustrated through a simple agency model, in which two units — upstream and downstream — jointly produce a service, and surplus depends on a non-contractible action chosen by the downstream unit’s manager. In a one-shot

\(^4\) Van den Steen (2007) calls “interpersonal authority” a principal’s ability to make disagreeing agents obey her orders.
transaction, it is preferable to make the manager own the downstream unit, as that gives her stronger incentives to perform than vertical integration. This is no longer the case, however, when the manager and the upstream unit transact repeatedly and, therefore, can enter relational contracts, in which the manager promises to act to maximize the joint surplus, in exchange for future rents. Under both vertical integration and separation, if the manager reneges on her promise to perform, she benefits from greater free time. Under vertical separation, however, the shirking manager also appropriates any short-run increase in unit profits due to non-performance, because she is residual claimant of the downstream unit. Therefore, when the actions that maximize the upstream and the downstream unit’s profits differ substantially — that is, when spillovers between units are large — the manager’s promise to perform will be more credible under vertical integration than under separation.

The model also provides a novel explanation for the fact that a greater cost of monitoring downstream managers leads to less vertical integration (Brickley and Dark (1987), Lafontaine (1992), Lafontaine and Shaw (2005), Lafontaine and Slade (2007), Arruñada et al. (2008)). Since direct monitoring is essential to sustain relational contracts over the agent’s performance, and vertical integration is efficient only in the presence of relational contracts, the model predicts that greater monitoring costs will lead, ceteris paribus, to less vertical integration. This complements the model in Lafontaine and Slade (1996), according to which, in the presence of low monitoring costs, risk-averse agents should receive high commissions on direct measures of performance, and low commissions on indirect measures, such as sales. The difference is that the prediction in this model is directly applicable to the discrete choice between vertical integration and separation. Also,
it does not require agent’s risk aversion, whose importance, as mentioned before, seems
dubious, at least in retail contracting.

The work most closely related to this paper is probably Baker et al. (2002), which
interprets vertical integration as a means to facilitate relational incentive contracts. The
difference is that, in Baker et al. (2002), vertical integration is used to reduce the principal’s
and agent’s temptations to renege on the informally agreed incentives, by holding up each
other and renegotiating the distribution of surplus. As a consequence, their model cannot
yield the predictions of this paper on the relation between agency conflicts and vertical
integration. In particular, in a model a la Baker et al. (2002), spillovers from the agent’s
actions to the principal’s inalienable assets — for instance, from an outlet manager’s
service effort to the value of the franchisor’s brand — have the same effect on the parties’
payoffs from renegotiation under vertical integration and separation and, therefore, do not
affect the choice between these two governance structures.

The rest of this paper is organized as follows. Section 2 introduces the model’s
definitions and assumptions. Section 3 discusses the incentive properties of vertical
integration and separation in a spot environment. Section 4 discusses the choice between
vertical integration and separation in a relational environment. Section 5 derives some
comparative static predictions. Section 6 discusses empirical works that support these
predictions. Section 7 concludes.
2. The environment

Consider two specialized units, upstream and downstream, engaged in the joint production of a service, such as dining, banking or transportation. The service concept is developed by the upstream unit, run by manager U, and the service is delivered to consumers by the downstream unit, run by manager D. The joint surplus depends on D’s non-contractible action $d \in \mathbb{R}^+$, and is given by $B(d) + V(d) - C_a(d) - C_p(d)$, where $B(d)$ is the residual value of the upstream unit, $V(d)$ is the residual value of the downstream unit, and $C_p(d)$ is D’s personal cost of performing, which can be interpreted as her private valuation of leisure or, equivalently, as stress caused by the unpleasantness and difficulty of her task. For example, $d$ could be the degree of compliance of a McDonald’s restaurant with the brand’s outlet design and cleanliness standards, $B(d)$ the corresponding value of the McDonald’s brand, $V(d)$ the restaurant’s long-term revenues from serving customers according to the standards, $C_a(d)$ the restaurant’s forgone profit from following the standards, instead of offering a customized service, or a low quality service, and $C_p(d)$ the restaurant manager’s stress from coordinating standard implementation.\(^5\) I assume $B(d)$ and $V(d)$ are increasing in $d$ and concave, $C_a(d)$ and $C_p(d)$ are increasing in $d$ and convex, and $B(0) = V(0) = C_a(0) = C_p(0) = 0$.

\(^5\) The joint surplus may also depend on U’s effort and investments in developing the service concept. See Lutz (1995) for a complementary model that emphasizes upstream incentives.
In this model, the upstream unit is owned by manager U, whereas the downstream unit can be either owned by manager D (vertical separation) or manager U (vertical integration), in which case D runs the unit as U’s employee. I assume ownership of a unit conveys the right to appropriate its residual value (Holmstrom and Milgrom (1991, 1994), Baker et al. (2008a)), that the units’ residual values — $B(d)$ and $V(d) - C_a(d)$ — and D’s cost of performing $C_p(d)$ are all non-contractible, and that no contractible measures of performance are available. These assumptions are consistent with the fact that firms tend to appropriate most of the profits generated by the assets they own (Krueger (1991), Maness (1996)), and imply that D’s incentives to perform are completely determined by the allocation of ownership rights over the downstream unit, that is, by the choice between vertical integration and separation.

### 3. Spot governance

In this section, I assume U and D meet only once, with no opportunities to trade in the future. In this spot environment, trade occurs as follows. At stage 0, U and D assign ownership of the downstream unit; at stage 1, U makes an offer to D concerning any contract terms necessary to regulate their relationship; at stage 2, D chooses the action $d$ and incurs the personal cost $C_p(d)$; finally, at stage 3, the units’ residual values are realized as a function of $d$. 
3.2. First best

If effort was contractible, no matter who owns the downstream unit, U and D would agree, at stage 1, that D must choose \( d \) to maximize the joint surplus

\[
JS = B(d) + V(d) - C_a(d) - C_p(d),
\]

in exchange for a payment. The necessary and sufficient first order condition for this problem is

\[
B'(d) + V'(d) = C'_a(d) + C'_p(d) \tag{1}
\]

yielding effort \( d^{FB} > 0 \) and surplus \( JS^{FB} = B(d^{FB}) + V(d^{FB}) - C_a(d^{FB}) - C_p(d^{FB}) \).

However, since \( d \) is non-contractible, U and D cannot, in general, achieve the first best in a spot environment. To achieve the second best, they must choose, at stage 0, between assigning ownership of the downstream unit to U (vertical integration) or D (vertical separation).

3.3. Vertical integration

Under vertical integration, U receives, at stage 3, \( B(d) + V(d) - C_a(d) - s \) and D receives \( s \), where \( s \) is a fixed salary contracted at stage 1. Anticipating this, at stage 2, D chooses \( d \) to maximize \( s - C_p(d) \), which has a corner solution at \( d^{VI} = 0 \). Hence, at stage 1, U and D set \( s = 0 \), and the joint surplus under vertical integration is given by

\[
JS^{VI} = B(0) + V(0) - C_a(0) - C_p(0) = 0 < JS^{FB}. \tag{6}
\]

\[\]
3.4. Vertical separation

Under vertical separation, U receives, at stage 3, $B(d)$ and D receives $V(d) - C_a(d)$. Anticipating this, at stage 2, D chooses $d$ to maximize $V(d) - C_a(d) - C_p(d)$. The necessary and sufficient first order condition for this problem is

$$V'(d) = C'_a(d) + C'_p(d)$$

yielding effort $d^{VS}$ and joint surplus $JS^{VS} = B(d^{VS}) + V(d^{VS}) - C_a(d^{VS}) - C_p(d^{VS})$. We are now ready to state the following

**Proposition 1**: In a spot environment, the joint surplus under vertical separation is greater than under vertical integration.

**Proof**: Since $V(d)$ is increasing in $d$ and concave, $C_a(d)$ and $C_p(d)$ are increasing in $d$ and convex and $V(0) = C_a(0) = C_p(0) = 0$, it must be that $d^{VS} > d^{VI} = 0$. Given that $B(d)$ is increasing in $d$, the joint surplus under vertical separation is

$$JS^{VS} = JS(d^{VS}) = B(d^{VS}) + V(d^{VS}) - C_a(d^{VS}) - C_p(d^{VS}) > JS(0) = 0 = JS^{VI}.$$ QED.

This is the result one would expect from a standard agency model: if D spends more effort, in equilibrium, under vertical separation than under integration, but less than the first best under both governance structures, vertical separation should yield greater surplus than integration.
4. Relational governance

In this section, I assume that, while D’s action cannot be observed by a court, it can be observed by the upstream manager U (i.e., effort is *observable but not verifiable*). Hence, if U and D repeat their transaction forever (or, equivalently, if they do not know when their last transaction will occur), they can enter *relational contracts*, in which they use their mutual concern for future trade to enforce higher performance than can be achieved under spot governance. What determines the choice between vertical integration and separation in this relational contracting environment?

Suppose the relational contract requires D to take action $d^*$, such that $d^* \leq d^{FB}$. At stage 1 of every period $t$, U and D allocate ownership of the downstream unit, and U offers D a fixed payment $w_g$, where $g \in \{VI, VS\}$. At stage 2, D chooses $d$ and incurs the personal cost $C_p(d)$. At stage 3, the unit values $B(d)$ and $V(d) - C_a(d)$ are realized. If U fails to offer $w_g$ at stage 1, or D fails to accept it, the parties immediately terminate their relationship. If D fails to spend the promised effort $d^*$ at stage 2, the parties earn their spot payoffs under the present governance structure in period $t$, and then terminate the relationship.\(^7\)

\(^7\)The implicit assumption, here, is that terminating the relationship at stage 2, once production has already started, is too costly. For instance, U may need D to provide some ancillary, contractible cooperation after stage 2 and until the end of the period, and may find it difficult to immediately replace D for that purpose.
After termination, U and D renegotiate asset ownership, and the downstream unit is allocated to its best use outside the relationship. Without loss of generality, I assume U’s and D’s payoffs after termination are given by the Nash bargaining solution with equal bargaining power. Denote by $U_g^0$ and $D_g^0$ the post-termination payoffs the parties would earn in each period under allocation $g \in \{VI, VS\}$, and by $U_g^0$ and $D_g^0$ the post-termination payoffs they would earn under the optimal allocation. Then, U’s and D’s net per period payoffs after termination, when the initial governance structure is $g$, are given, respectively, by $U_g^T = \frac{1}{2} \left( U_g^0 + D_g^0 + U_g^0 - D_g^0 \right)$ and $D_g^T = \frac{1}{2} \left( U_g^0 + D_g^0 + D_g^0 - U_g^0 \right)$, where $JS_{VI}^T < U_{VI}^T + D_{VI}^T = U_{VI}^T + D_{VI}^T = JS_{VS}^T < JS_{FB}^T$.

Given the above assumptions and definitions, we can derive the conditions that make the relational contract self-enforcing. This will occur if, and only if each party’s present gains from reneging are smaller than the present value of her quasi-rents from future trade.

### 4.2. Vertical integration

With a slight abuse of notation, let me denote U’s and D’s per period profits from honoring the relational contract, gross of the fixed transfer $w_{VI}$, as

---

8 Other relational contracting models assume the initial governance structure is maintained after termination (Klein and Murphy (1997), Halonen (2002)). If that was the case, vertical integration, which concentrates asset ownership in the upstream unit, would have the additional advantage to reduce the downstream manager’s fall-back option after reneging on the relational contract. However, the assumption in this model, according to which the parties allocate asset ownership efficiently after termination, seems more realistic in many contexts. For instance, after termination, some franchisees sell their assets to the franchisor, whereas others keep the assets (except for those that carry the franchisor’s logo) and continue the activity under different brands.

9 See Bull (1987), MacLeod and Malcomson (1989), Levin (2003), and Baker et al. (2002, 2008b), for related models of relational contracts.
$U_{v_l}^* = B(d^*) + V(d^*) - C_u(d^*)$ and $D_{v_l}^* = -C_p(d^*)$, yielding joint surplus $JS^* = U_{v_l}^* + D_{v_l}^*$. Let $r$ be U’s and D’s common interest rate. Then, the relational contract is self-enforcing if, and only if

\[
\frac{1 + r}{r} (U_{v_l}^* - w_{v_l}) \geq \frac{1 + r}{r} U_{v_l}^T
\]

\[\frac{1 + r}{r} (D_{v_l}^* + w_{v_l}) \geq \frac{1 + r}{r} D_{v_l}^T\]

\[-C_p(d^*) + \frac{1}{r} (D_{v_l}^* + w_{v_l}) \geq \frac{1}{r} D_{v_l}^T\]

Conditions (3) and (4) are U and D’s participation constraints, whereas condition (5) is D’s dynamic incentive compatibility constraint. We are now ready to state the following

**Lemma 1**: Under vertical integration, the relational contract between U and D is self-enforcing if D’s reneging temptation, $C_p(d^*)$, is not greater than the present value of the parties’ aggregate quasi-rent stream, $\frac{1}{r} (JS^* - JS^T)$.

**Proof**: The largest $w_{v_l}$ satisfying both (3) and (4) is $\overline{w_{v_l}} = U_{v_l}^* - U_{v_l}^T$. Plugging $\overline{w_{v_l}}$ into (5) and rearranging yields the condition

\[C_p(d^*) \leq \frac{1}{r} (JS^* - JS^T)\]

Since $\overline{w_{v_l}}$ satisfies both (3) and (4), (6) is sufficient for self-enforcement. QED.
4.3. Vertical separation

Denote U and D’s per period profits from honoring the relational contract under vertical separation, gross of the transfer $w_{vy}$, as $U^*_v = B(d^*)$ and $D^*_v = V(d^*) - C_a(d^*) - C_p(d^*)$, yielding joint surplus $JS^* = U^* + D^* = U^*_v + D^*_v$, as under vertical integration. Also, denote D’s spot profit under vertical separation by $D^E = V(d^E) - C_a(d^E) - C_p(d^E)$. The relational contract is self-enforcing if, and only if

(7) \[
\frac{1+r}{r} (U^*_v - w_{vy}) \geq \frac{1+r}{r} U^*_v
\]

(8) \[
\frac{1+r}{r} (D^*_v + w_{vy}) \geq \frac{1+r}{r} D^*_v
\]

(9) \[
D^*_v + \frac{1}{r} (D^*_v + w_{vy}) \geq D^E + \frac{1}{r} D^*_v
\]

Paralleling the analysis of vertical integration, we can state the following

**Lemma 2:** Under vertical separation, the relational contract between U and D is self-enforcing if D’s reneging temptation $D^E - D^*$ is not greater than the present value of the parties’ aggregate quasi-rent stream $\frac{1}{r} (JS^* - JS^T)$.

**Proof:** The largest $w_{vy}$ satisfying both (7) and (8) is $w_{vy} = U^*_v - U^*_v$. Plugging $w_{vy}$ into (9) and rearranging yields the condition

(10) \[
D^E - D^* \leq \frac{1}{r} (JS^* - JS^T)
\]

Since $w_{vy}$ satisfies both (7) and (8), (10) is sufficient for self-enforcement. QED.
Given Lemmas 1 and 2, we can determine the governance structure that maximizes the “self-enforcing range” of a given relational contract (Klein (1996), Baker et al. (2002, 2008b)), that is, the structure that makes the relational contract self-enforcing for the largest range of interest rates. This is given by the following

**Proposition 2:** For a given relational contract, vertical integration maximizes the self-enforcing range if, and only if

\[
C_p \left(d^{VS}\right) \leq C_a \left(d^*\right) - C_a \left(d^{VS}\right) - \left(V \left(d^*\right) - V \left(d^{VS}\right)\right)
\]

**Proof:** Rearranging (6) and (10) in terms of the interest rate \(r\) yields

\[
r \leq \frac{JS^* - JS^0}{C_p \left(d^*\right)} \equiv r^{VI} \left(d^*\right)
\]

\[
r \leq \frac{JS^* - JS^0}{D^{VS} - D^{VIS}} \equiv r^{VS} \left(d^*\right)
\]

We know from the proof of Lemmas 1 and 2 that \(r^{VI} \left(d^*\right)\) and \(r^{VS} \left(d^*\right)\) are the largest possible values of \(r\) such that the relational contract is self-enforcing under vertical integration and separation, respectively. Since the numerators in \(r^{VI} \left(d^*\right)\) and \(r^{VS} \left(d^*\right)\) are identical, vertical integration maximizes the self-enforcing range if, and only if the denominator in \(r^{VS} \left(d^*\right)\) — D’s reneging temptation under vertical separation — is greater than the denominator in \(r^{VI} \left(d^*\right)\) — D’s reneging temptation under vertical integration — that is, if \(D^{VS} - D^{VIS} \geq C_p \left(d^*\right)\). Solving for \(D^{VS}\) and \(D^{VIS}\) yields condition (11). QED.
The message from Proposition 2 is that, while vertical integration mutes D’s incentives to perform in a spot environment (Proposition 1), it may strengthen them in a relational environment. Under both vertical integration and separation, D has a temptation to trade off effort for leisure. However, under vertical separation, D has an additional temptation to reduce performance in order to increase the downstream unit’s value, which is given by the term $C_v(d^*) - C_r(d^{VS})$. If this unit-level benefit from non-performance is large relative to the cost, given by $V(d^*) - V(d^{VS})$, D’s promise to perform will be more credible under vertical integration than under vertical separation.

As a last step, we can define the optimal governance structure. For a given governance $g \in \{VI, VS\}$, U and D will chose the best feasible relational contract such governance can support. This requires D to choose action $d^*_g$, which maximizes the joint surplus $JS(d)$, subject to the self-enforcement condition $r \leq r^*(d)$. Therefore, the optimal governance structure is given by $g^* = \arg \max_g \{JS(d^*_g)\}$. In particular, this implies that vertical integration is the optimal governance structure if, and only if $JS(d^*_VI) > JS(d^*_VS)$. 
5. Testable predictions

5.1. Costly relational contracts

According to the analysis in the previous section, absent relational contracts, vertical integration will not be observed, because it minimizes D’s incentives to perform. Conversely, in the presence of relational contracts, vertical integration will be observed whenever it supports the best feasible relational contract. In practice, relational contracts may be feasible but costly, as they require U to monitor D’s provision of effort, and both parties to develop reputational capital and communication mechanisms that enable informal coordination. As a tractable formalization of costly relational contracts, consider the following extension to the model. Assume that, before entering relational contracts, U and D must incur a fixed cost $m$. For instance, $m$ could be U’s cost of developing/purchasing a monitoring technology, or U’s cost of credibly signaling his intention to be in a long-term relationship with D. If the level of effort sustainable through the best feasible relational contract is $d^*$, U and D will enter such contract if, and only if the per period joint surplus it generates, net of the fixed cost $m$, is greater than the surplus that can be achieved without it, that is, if

$$JS^* - m \geq JS^T$$

We know from Proposition 1 that vertical integration will be observed only if (14) holds. Moreover, when (14) holds, vertical integration will be observed if it supports the best relational contract $d^*$. Assume, now, that the joint surpluses with and without
relational contracts are random variables, given by \( JS^* + \eta_s \) and \( JS^T + \eta_T \), respectively, where \( \eta_s \) and \( \eta_T \) are unobservable and independent, and have cumulative distribution function \( G(\cdot) \), such that \( G'(\cdot) > 0 \). Also, assume that whether vertical integration supports \( d^* \) or not is independent of \( \eta_s \) and \( \eta_T \), and, without loss of generality, denote as \( K \) the probability that it does. Then, the probability to observe vertical integration is

\[
\Pr(VI) = \Pr\left( JS^* + \eta_s - m \geq JS^T + \eta_T \right) K =
\]

\[
= \Pr\left( \eta_T - \eta_s \leq JS^* - JS^T - m \right) K =
\]

\[
= G\left( JS^* - JS^T - m \right) K
\]

This implies, in turn, that

\[
\frac{\partial \Pr(VI)}{\partial m} = -G\left( JS^* - JS^T - m \right) K < 0
\]

Intuitively, since an increase in the fixed cost \( m \) reduces the probability that U and D enter relational contracts, and vertical integration can only be observed in the presence of relational contracts, an increase in \( m \) should reduce, \textit{ceteris paribus}, the probability to observe vertical integration.

\section*{5.2. Managerial productivity and vertical integration}

In this section, I use a linear-quadratic version of the model to study how variations in the productivity of D for the upstream and downstream units affect the choice between vertical integration and separation. Let \( B(d) = bd \), \( V(d) = vd \), \( C_a(d) = \frac{c_a}{2}d^2 \) and
\[ C_p(d) = \frac{c_p.d^2}{2}, \text{ where } b, v, c_a, c_p > 0. \] Substituting these expressions into the first order conditions (1) and (2) yields \[ d^{FB} = \frac{b + v}{c_a + c_p} \text{ and } d^{VS} = \frac{v}{c_a + c_p}. \]

I focus on the case in which a relational contract requiring D to supply the first best relational contract \( d^{FB} \) is sustainable under some governance structure. This implies that vertical integration is optimal if, and only if it minimizes D’s reneging temptation (Proposition 2), which substantially simplifies the analysis. To derive testable comparative statics, assume D’s reneging temptations under vertical integration and separation are random variables given, respectively, by \( C_p\left(d^{FB}\right) + \varepsilon_{VI} \) and \( D^{VS} - D^{FB} + \varepsilon_{VS} \), where \( \varepsilon_{VI} \) and \( \varepsilon_{VS} \) are unobservable and independent, and have cumulative distribution function \( F(\cdot) \), such that \( F'(\cdot) > 0 \). Substituting for the values of \( d^{FB} \) and \( d^{VS} \) in the linear-quadratic model, the probability that vertical integration is optimal is given by

\[ \Pr(VI) = \Pr(\varepsilon_{VI} - \varepsilon_{VS} \leq D^{VS} - D^{FB} - C_p\left(d^{FB}\right)) = F\left(\frac{b^2c_a - c_p\left(v^2 + 2bv\right)}{2(c_a + c_p)^2}\right) \]

This implies that

---

\(^{10}\) The results in this section continue to hold for model specifications different from the linear-quadratic one. For example, they still hold if one assumes \( B(d) = b\sqrt{d}, V(d) = v\sqrt{d}, C_a(d) = c_ad, C_p(d) = c_pd \) or \( B(d) = b \ln d, V(d) = v \ln d, C_a(d) = \frac{c_a}{2}d^2, C_p(d) = \frac{c_p}{2}d^2. \)

\(^{11}\) The analysis for the case in which relational contracts can only sustain levels of effort lower than the first best yields qualitatively similar results, and is available from the author upon request.
An inspection of (18) indicates that the partial derivative for \( v \) is unambiguously negative, implying that vertical integration is more likely to be efficient the less productive \( D \) is for the downstream unit (smaller \( v \)). On the other hand, the sign of the partial derivative for \( b \) is ambiguous. In particular, it is positive if, and only if

\[
\frac{\partial \text{Pr}(VI)}{\partial b} = F'(\cdot) \left( \frac{b c_a - c_p v}{(c_a + c_p)^2} \right)
\]

In words, when the portion of \( D \)'s performance that spills over the upstream unit is large relative to the portion that benefits the downstream unit (i.e., \( \frac{b}{v} \) is large enough), \( D \)'s present benefit from reneging on the promised performance under vertical separation, given by the term \( C_a \left( d^{FB} \right) - C_a \left( d^{IS} \right) \) from (11), is large relative to the value reduction she bears, given by the term \( V \left( d^{FB} \right) - V \left( d^{IS} \right) \). In these conditions, an increase in the marginal spillover effect (greater \( b \)) makes \( D \)'s promise to perform more credible under vertical integration than under separation.
6. Evidence on the model’s predictions: a meta-analysis of the empirical literature

This section discusses several empirical works that support the model’s predictions, the most recent of which are summarized in Table 1. All of these works focus on long-term business relationships, such as the ones between franchisors and outlet managers, motor carriers and truck drivers, or major and regional airlines, where, consistent with the model, relational contracts are likely to play an important role.

Effect of the manager’s performance on the upstream unit (spillover)

Consistent with the idea that, when the manager’s actions generate greater spillovers on the upstream unit ($b$ grows) and spillovers are substantial ($\frac{b}{v} > \frac{c_D}{c_a}$), the downstream unit should be vertically integrated, Forbes and Lederman (2008) find that flight routes connecting US airports with worse weather conditions, and where the major carrier’s hub represents one of the endpoints, tend to be served by regional airlines that are owned by the major. In airports with bad weather conditions, it is crucial for a network’s reputation that regional carriers coordinate flight rescheduling with the major in order to avoid excessive delays and cancellations (greater $b$). Moreover, efficient rescheduling matters more to the network’s reputation in the major carrier’s hub, or in airports where the major operates more flights. Yeap (2006) finds that chain restaurants with in-house food production, dine-in service and high prices, where the quality of customer service and the restaurant’s cleanliness and comfort are more critical to the chain’s brand (greater $b$), are
more likely to be vertically integrated. Brickley and Dark (1987) find that vertical integration is more frequent than franchising in industries where retail outlets serve non-repeat customers and, therefore, the outlet manager’s effort increases the value of the common brand more than the individual outlet’s profits (greater $b$). Arruñada, Vázquez and Zanarone (2008) find that car dealerships in areas with greater outlet density, where a dealer’s effort to capture customers benefits neighboring dealers, tend to be vertically integrated. Lafontaine (1992) and Lafontaine and Shaw (2005) find that, in franchise networks, greater value of the common brand (greater $b$) leads to more vertical integration. Finally, Nickerson and Silverman (2003) find that motor carriers tend to own trucks when they provide less-than-truckload services, in which shipments must be coordinated and, therefore, breakdowns due to poor truck maintenance are detrimental to the carrier’s reputation, and when they invest more in advertising to build a brand name (greater $b$). All these works study industries where downstream firms operate under the upstream firm’s brand name and, therefore, spillovers are likely to be substantial (i.e., it is likely that $\frac{b}{v} > \frac{c_a}{c_a}$). According to the model, these are the types of industries in which a marginal increase in the impact of managerial performance on the brand (greater $b$) should lead to more vertical integration, which is consistent with the data.

---

12 An exception to these patterns is Brickley (1999), who finds no significant relationship between spillovers across retail outlets and the extent of vertical integration.

13 The positive empirical association between brand value and vertical integration can also be interpreted via two-sided agency models, in which integration gives upstream managers stronger incentives to spend brand-maintenance effort (Lutz (1995)).
The empirical literature also supports the model’s prediction that greater productivity of the manager’s performance for the downstream unit should decrease the extent of vertical integration. The evidence is consistent across industries, from gasoline distribution (Shepard (1993)) to footwear retailing (Woodruff (2002)) and banking (Brickley, Linck and Smith (2003)). Unlike the one on spillovers, however, this evidence can be easily reconciled with classic theories of vertical integration, such as the ones based on double-sided moral hazard (Lutz (1995), Maness (1996)). Therefore, I will not discuss it further here. The interested reader can find a detailed account in the excellent surveys of Lafontaine and Slade (1997, 2007).

The costs of relational contracting

The model predicts that vertical integration should be less frequently observed as the costs of entering relational contracts increase. A natural example of such costs is given by monitoring, since relational contracts require the upstream unit to monitor the effort of the downstream unit’s manager (although not in a court-verifiable way). The model would then predict that a greater cost of monitoring the downstream manager leads to less vertical integration. Consistent with that, several papers on franchising find that retail outlets that are more distant from the franchisor’s headquarters (Brickley and Dark (1987), Arruñada, Vázquez and Zanarone (2008)) or geographically dispersed (Lafontaine (1992), Lafontaine and Shaw (2005)) — and, therefore, more difficult to monitor — tend to be vertically separated.
Another possible interpretation for the costs of relational contracting is in terms of signaling and communication costs, which parties must incur in order to initiate a long-term relationship and reach an informal agreement on their implicit obligations. In particular, one could argue that parties who have transacted frequently in the past can implement relational contracts at lower signaling and communication costs. The model would then predict that, as the frequency of past transactions increases, vertical integration is more likely. Evidence supportive of this prediction is provided by Corts and Singh (2004), who find that oil companies and drillers that interact repeatedly and, therefore, rely more on relational contracts, allocate residual claims to the oil company—a solution that resembles vertical integration as defined in this paper.¹⁴

7. Conclusion

This paper has developed a relational contracting model to explain why, in the presence of interest conflicts and spillovers, units at different stages in the chain of production vertically integrate. The proposed explanation is that vertical integration may reduce the managers’ short-term gains from acting opportunistically, relative to their long-term rents from cooperating with the firm. If the manager of a vertically integrated unit reneges on her promise to perform, she benefits from greater free time, but does not appropriate the associated increase in unit profits. Therefore, when the actions that maximize the firm’s and the individual unit’s profits differ substantially—that is, when interest conflicts and spillovers between units are large—a manager’s promise to perform will be more credible.

¹⁴ See Corts (2007) for an alternative interpretation of this result.
under vertical integration than under separation. As shown in section 6, a rich body of empirical evidence from numerous industries is consistent with the model’s predictions.

References


Table 1. Recent evidence on the model’s predictions: the effect of spillovers ($b$) and monitoring costs ($m$) on vertical integration

<table>
<thead>
<tr>
<th>Empirical study</th>
<th>Year</th>
<th>Industry</th>
<th>Agent’s task</th>
<th>Observed variations</th>
<th>Change in model’s parameters</th>
<th>Spillover large?</th>
<th>Effect on vertical integration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nickerson &amp; Silverman</td>
<td>2003</td>
<td>Trucking</td>
<td>Drive safely</td>
<td>Less-than-truckload</td>
<td>$b$ $v$ $m$</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>Yeap</td>
<td>2006</td>
<td>Chain restaurants</td>
<td>Control service quality</td>
<td>In-house production; dine-in service; high price</td>
<td>$+$</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>Forbes &amp; Lederman</td>
<td>2008</td>
<td>Air transportation</td>
<td>Reschedule flights</td>
<td>Bad weather airport</td>
<td>$+$</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>Arruñada et al.</td>
<td>2008</td>
<td>Car distribution</td>
<td>Implement standards</td>
<td>Intrabrand competition</td>
<td>$+$</td>
<td>Yes</td>
<td>+</td>
</tr>
<tr>
<td>Lafontaine &amp; Shaw</td>
<td>2005</td>
<td>Retailing</td>
<td>Manage outlet</td>
<td>Franchisor operates in more states</td>
<td>$+$</td>
<td>Yes</td>
<td>–</td>
</tr>
</tbody>
</table>