

Centre de Referència en Economia Analítica

Barcelona Economics Working Paper Series

Working Paper n° 150

**Specificity Revised:
The Role of Cross-Investments**

Matthew Ellman

November, 2004

Specificity Revisited: The Role of Cross-Investments

Matthew Ellman*

Universitat Pompeu Fabra

November, 2004

Barcelona Economics WP n°150

Abstract

Previous analysis has shown that traders may opt for specific technologies with no joint productivity advantage as a way to commit themselves to trading jointly, but only when long-term contracting is infeasible. This paper proves that specificity can also be optimal (by relaxing the budget-balance constraint) in settings with long-term contracting. Traders will opt for specificity when one trader makes a cross-investment and either (1) this cross-investment has a direct externality on the other trader, (2) both parties invest, or (3) private information is present. The specificity (e.g. from non-salvageable investments, specific assets and technologies, narrow business strategies, and exclusivity restrictions) is equally effective regardless of which trader's alternative trade payoff is reduced. Specificity supports long-term contracts in a broad range of settings - both with and without renegotiation. The theory also offers a novel perspective on franchising and vertical integration.

Jel Classification numbers: D23, K40.

*I thank Benito Arruñada, Oliver Hart, Eric Maskin, Joel Shapiro, Andrei Shleifer, Tomas Sjostrom and seminar participants at Harvard University, the L.S.E., and ESEM 2004 for helpful comments. I gratefully acknowledge financial support from the Sloan Foundation during my thesis (this paper expands on chapter 2) and subsequent support from the European Commission (HPMF-CT-1999-00317) and the Spanish Ministry of Science and Technology (BEC 2003-00412). Email: Matthew.Ellman@UPF.edu.

1. Introduction

Specificity is usually viewed as an unavoidable consequence of technological specialization, but a substantial literature has built up evidence that traders sometimes seek out specificity for strategic reasons. Even at the expense of average productivity, these traders use technologies, investment strategies, ownership structures and restrictive contracts to reduce their productivity on separation. For instance, specific technologies lead to human capital that goes to waste on switching firms, and capabilities that go to waste on changing projects (Williamson, 1985); specific (non-salvageable) assets are wasted in alternative trading (Klein, 1980, and Williamson, 1983); firms forfeit flexibility by sinking investments prematurely in specific projects (Arya et al, 2003) and by adopting narrow business strategies (Rotemberg and Saloner, 1994 and 2000); franchisors often insist on holding the property lease of their franchisees (Dnes, 1992); joint ownership can lead to asset waste when traders separate in disagreement; and exclusivity restrictions, such as post-termination non-compete covenants, reduce the return to human and organizational capital in outside trading (Bercovitz, 2000). In each case, the claim is that traders *intentionally* increase relationship (or project) specificity as a strategic choice not dictated by technological necessity.

In this paper, I develop a theory that unifies these phenomena (whether specificity hurts buyer *or* seller, principal *or* agent) through their common characteristic of implying waste when their trading ends prematurely. Williamson's (1983) insightful explanation of specific assets potentially extends to other forms of intentional specificity. His argument is analogous to the parable of the advancing army that burns its own bridges to commit itself to fight on (Schelling, 1960). Traders choose specificity to lower their payoffs from not trading, because this partially commits them to joint trade. The specificity must distort upwards the probability of joint trade to mitigate 'holdup' of investments complementary to joint trade.¹ Specificity then acts as a substitute for an assumed impossible, advance legal commitment to trade. Williamson's theory is important because legal commitments to trade are sometimes infeasible or ineffective,² but this theory implicitly predicts that long-term contracting, when feasible, will dominate and therefore displace intentional specificity.

My goal is to explain intentional specificity in settings where long-term contracting *is* viable. This is important because most documented intentional specificity occurs in precisely these settings. For instance, traders in franchising, employment, or ongoing procurement can usually commit to future trade by agreeing to have past contracts (e.g., a company standard form, perhaps with quantitative variations) govern *future* trading. Long-term contracts leave

no role for intentional specificity in Williamson’s (1983) model, because his non-contractible investments are “self-investments”: the seller raises its *own* benefit from joint trade by investing in cost reduction. So (as established in Klein et al, 1978, and formalized in Grout, 1984), the seller fully internalizes its investment returns if the buyer commits to trade at a fixed price. Specificity is redundant. So in this paper, I move away from the self-investment setting, and show how specificity can support long-term contracting by substituting for budget-breaking ‘third-party fines.’

Long-term trade contracts need support when investments have externalities. For instance, a franchisee managing a restaurant can train the waiters to be friendlier. This effort has a direct benefit on the franchisor, because it raises trademark value and royalties. The franchisee’s training effort also raises the franchisor’s payoff from extending the franchise contract, so it is a “cross-investment.”³ The investor (seller) has no incentive to cross-invest unless the buyer has the option to terminate trade.⁴ The optimal contract gives the buyer the option to accept or reject trade at predefined terms. For instance, a franchisor can either accept continuing trade with a franchisee at existing contract terms, or terminate the franchise - even if termination is not “at will,” the penalties are usually low at contract renewal dates. Similar options are common in strategic alliances, procurement and employment contracting.

My contribution is to show how and when intentional specificity is used to reinforce these contracts. A typical option contract works as follows: the buyer has the option to accept or refuse trade at a pre-determined price. The price is chosen so that the buyer accepts trade only after high cross-investment (implying high quality trade) by the seller. The price must therefore lie between the buyer’s values for low and high quality trade. The seller then earns this price as a reward for high cross-investment and is otherwise punished by having to engage in an alternative trade. The difference between reward and punishment provides the seller with an incentive to exert high investment. I derive a role for specificity via its impact on reward credibility and punishment severity. First, when specificity reduces the buyer’s alternative trade payoff⁵ and hence opportunity cost of joint trade, the buyer can credibly raise the seller’s reward. Second, when specificity reduces the seller’s alternative trade payoff,⁶ it raises the severity of the punishment for low investment. These two effects combine: The traders can increase the seller’s incentive by the amount that specificity reduces the *sum* of their alternative trade payoffs. This permits a simple measure of specificity in terms of the *total* potential waste, and allows my theory to unify analyses of specificity affecting either buyers or sellers.

Using stipulated damages to increase the seller’s punishment fails to resolve the incentive problem, because this reduces reward credibility by as much as it increases punishment severity. Allocating the damages as fines to a third party would resolve the problem, but I make the standard assumption that budget-breaking is not viable.⁷ In essence, intentional specificity serves as a strategic solution to the budget-balance problem.

In the basic cross-investment model, option contracts permit the first-best without specificity so I need to identify a non-trivial demand for increased incentives. Two familiar reasons why option schemes might give insufficient incentives are the presence of stochastic variation in returns and two-sided (non-contractible) investments. I focus on a surprisingly neglected reason: cross-investments often have “direct externalities” in that they create “*direct benefits*” to the non-investing party.⁸ Recall how in the above example, the franchisee’s training efforts benefit the franchisor independent of future trading, as well as having the cross-investment benefit (contingent on continued trading). Similarly, in employment contracting, worker efforts that increase learning-by-doing usually benefit the employer directly, in addition to increasing future productivity. There is no way to make the seller internalize these direct benefits, so intentional specificity may be needed to raise effort incentives.

Renegotiation is not necessary for my specificity results. Prohibiting renegotiation allows incentives to be increased slightly, but does not preclude a need for specificity (see 5.1). In fact, renegotiation can be problematic. Che and Hausch (1999) argue that option schemes are not robust to renegotiation because the buyer would always reject trade to force subsequent renegotiation to a lower price. This problem does not arise when trade is decided at a sharp deadline, because subsequent renegotiation is then too late. Nor need it arise when trade opportunities disappear gradually, because if renegotiation is represented by finite horizon Rubinstein bargaining, then during each bargaining delay, trade can only occur at the option price (see Ellman, 2004). Infinite horizon bargaining on the other hand, is problematic, because it permits reputational equilibria in which the buyer strategically threatens to reject trade unless the trade price is lowered. Che and Hausch (1999) prove that short-term contracting is then optimal. The incentives are then lower and specificity can only help if it distorts the probability of joint trade upwards (as in Williamson, 1983).⁹ Fortunately, buyers can often escape this renegotiation problem by (directly or indirectly) building a reputation for not making strategic threats.¹⁰ Sometimes, they effectively commit against renegotiation altogether, as when a firm commits to improve its partners’ terms of trade over time (as in “up-or-out” employment), or a franchisor commits to homogeneity in all its franchise contracts (see section 6).

In summary, the main theoretical contribution of the paper is to prove that intentional specificity complements long-term contracting in the presence of a budget-balance constraint. Renegotiation is neither necessary nor sufficient, but the cross-investment effect is essential. Uncertainty, two-sided investment, or most commonly, direct externalities of the cross-investment, each offer a sufficient secondary condition.

The paper may also facilitate empirical analysis by justifying the measurement of intentional specificity as a single variable, based on total potential waste. To test my theory's predictions against competing alternatives, I focus on the case of franchising.¹¹ Franchising offers a good real-world example of how the theory works in practice. This is an important policy application, because of the regulatory controversy over arrangements increasing specificity, initially branded as "unfair practices" by law courts (see Klein, 1980, Williamson, 1985, and Dnes, 1992). Bercovitz (2000) has collected and analyzed a rich data set containing proxies for the externalities and levels of specificity central to my theory. So I use her data to test my theory's predictions against competing alternatives.

My results are related to the seminal work of Klein (1980) and Klein and Leffler (1981) who provide a theoretical foundation for specific capital investments (brandname advertising, production capital and 'illiquid contractual arrangements') in a market setting with reputation-building but no formal contracts (and no renegotiation). The biggest theoretical difference is that they are forced to assume an infinite horizon setting and trigger-strategy equilibrium,¹² because they have no cross-investment effect. By contrast, I derive my two-sided specificity result in finite-horizon settings with unique subgame perfect equilibria. So my cross-investment model provides a valuable contractual foundation for Klein's (1980) insightful argument that franchisors can "assure quality by requiring franchisee investments in specific (nonfully salvageable) production assets." Intentional specificity is equivalent to posting 'hostages' (see Williamson, 1983, above, and Raub and Keren, 1993) that, absent joint trade, go to waste - the specific returns in my model are 'held hostage' to continuation of joint trade. However, the hostage literature has not allowed for long-term contracting, nor identified the role of cross-investments. My two-sided definition of specificity is also novel and can clarify the literature. For instance, Rotemberg and Saloner's (1994 and 2000) analysis of employee innovations fits nicely into my specificity theory: the employee efforts to create a useful innovation are actually cross-investments (given their contractibility assumptions), and the firm increases specificity when, by narrowing its business strategy, it makes its own capabilities more specific to the desired employee innovations.

In a step towards creating novel applications, I also analyze the practical issue of imple-

menting specificity when not directly contractible. When specificity is controlled by only one party, it may be used for rent-seeking and will not necessarily enhance incentives. Nonetheless, organizational design may be used to allocate authority to parties with private incentives to implement efficiency-enhancing specificities (see 5.5). This often requires vertical integration, opening up an exciting and novel perspective on the question of integration.¹³

The paper follows a simple plan. Section 2 illustrates the key argument with a numerical example. Section 3 sets out the basic model. Section 4 derives the maximal incentives and identifies the role of intentional specificity. Section 5 analyzes robustness and generalizations, proving the centrality of budget-balance. Section 6 evaluates empirical evidence from franchising, and section 7 concludes.

2. Numerical Illustration

A brief numerical example demonstrates my main claims. I begin by assuming that traders can commit against renegotiating and then I establish robustness to renegotiation. A buyer, B, wants a seller, S, to supply a pre-contractible trade (a good or service) over two periods. For instance, B is a boss “buying” labor services from an employee, S. Before trade begins, B and S can choose to impose a non-compete restriction on S that has no impact on productivity within the relationship but reduces the external value of S’s human capital. This specificity reduces S’s period 2 alternative trade payoff from 9 to 0. I fix B’s period 2 alternative trade payoff at 0. Even without intentional specificity, joint trade is optimal in both periods but there is a moral hazard problem: the quality of each period’s joint trade depends on a single effort/investment chosen privately by S in the first period. Neither quality can ever be described in a contract.

There are two action periods (see figure 1 below for a generalized timeline): S chooses an effort in period 1 (where trade is always enforced); in period 2 (after B observes S’s period 1 effort), B and S either engage in the precontractible *joint trade* or take their alternative trading opportunities. Before period 1 commences, B and S can agree on an option contract that ensures trade in period 1 and prespecifies a transfer price, p , that B owes S if B opts for *joint trade* in period 2. (Note that p is the wage in the employment example.) High effort costs S 8 payoff units while low effort costs 0. High effort gives B a *direct benefit* of 5 in period 1. High effort also affects B’s value from taking (period 2) joint trade at price p but S’s cost of supplying trade is fixed at 0: if B retains S after low effort, B’s payoff is $10 - p$ and S gets p , and if B retains S after high effort, B’s payoff is $16 - p$ and S gets p . So high

effort has a *cross-investment effect* of 6. After rejection, B and S take their alternative trade payoffs which are 0 for B and 9 for S - I treat first the case without intentional specificity. High effort is optimal since its total benefits, 11 ($= 5 + 6$), exceed its cost of 8. B and S need a scheme for which S's *incentive* (for high effort) - the difference in S's continuation payoffs after choosing high and low efforts - is greater than 8 (his cost of effort).

The problem is non-trivial. Even when S has full bargaining power, short-term contracting (with p agreed *ex post*) leads to underinvestment: S negotiates $p = 16$ after high effort and $p = 10$ after low effort. His incentive is too low ($16 - 10 = 6 < 8$) - he internalizes the cross-investment **but** neglects the direct benefit. The option scheme just described can help (see figure 2 in Appendix C for a game tree representation with p set at $16 - \varepsilon$): before period 1, B and S set p at 16 as an exercise price at which B can take trade in period 2. After high effort, B is just willing to pay $p = 16$ since the net payoff of 0 ($= 16 - p$) equals her alternative trade payoff of 0. Whereas, after low effort, B rejects S since joint trade is then only worth 10 and $10 - 16 < 0$. This option scheme allows B to 'credibly promise' to reward high effort with the "carrot," $p = 16$, and to punish low effort by rejecting S. In subgame perfect equilibrium, S anticipates getting 16 from high effort and getting his alternative trade payoff, the "stick," of 9 from low effort. S now has an incentive (carrot less stick) of 7 (16 less 9) to choose high effort, but his cost is 8. So there is a demand for raising incentives further and this is where specificity enters the picture.

Selection of the specific technology prior to trading reduces S's alternative trade payoff from 9 to 0 - a potential waste of 9. B's effort-contingent reward strategy is clearly unchanged, so the carrot of 16 remains credible. Meanwhile, the stick becomes more severe - S now gets 0 not 9 in his alternative trade. S's incentive rises by 9 (the size of the potential waste) to 16. This easily overcomes S's cost of 8 and the problem is solved.¹⁴

Renegotiation robustness is clearest when period 2 has a sharp trade deadline. At the deadline it becomes too late to renegotiate so B exercises any valuable trade option. Hence, after high effort, S would reject B's renegotiation offers, anticipating that B will end up exercising her (open-ended) trade option. B cannot force S to renegotiate. So B pays out 16 for high effort as before. After low effort, B would reject the trade option, so B and S renegotiate (to joint trade) before the trade deadline. If B has all the bargaining power, S gets none of the renegotiation surplus and his punishment for low effort still equals his alternative payoff of 0. S's incentive is then 16, just as before. S's incentive falls as his bargaining power rises but the first-best remains possible so long as B has at least 20% of the bargaining power.

3. Basic Model

This section generalizes the numerical illustration and defines intentional specificity. There are two production periods and two risk-neutral actors, B and S.¹⁵ In the first production period, S chooses an investment represented by its private cost, $e \geq 0$. This investment is unavoidably specific to the joint trade with B and is non-contractible. By contrast, the intentional specificity decision (see below) is effectively contractible and only affects e through its strategic effect on incentives. In the second production period, S and B can choose whether to engage in joint trade, denoted $q = 1$, or engage in alternative trades, denoted $q = 0$.¹⁶ I adopt Hart and Moore's (1988) trade model with $q = q_S \cdot q_B$ where $q_S, q_B \in \{0, 1\}$ are S and B's private trading decisions in the sequential trading game, TG: TG(i) S chooses q_S ; TG(ii) B chooses q_B .¹⁷ Note that $q_B = 0$ represents B firing S in the numerical illustration, where I left implicit S's willingness to trade, $q_S = 1$.

Underlying Moral Hazard in e . Trade "quality" is never contractible. B's value from trade depends on S's investment, e , chosen privately in period 1. To facilitate a thorough mechanism design analysis, I model this as a binary choice: $e \in \{l, h\}$; I normalize l to 0 but write l for *low*; h is the *high* effort so $h > 0$. B's value from the period 2 trade contract is thus an increasing function $b_2(e)$ which I write as b_2^e : $b_2^h > b_2^l$. For simplicity, I assume here (see 5.2 for self effects) that e is a pure "*cross-investment*" - it affects the future trade payoff of the non-investing party, B and not S's own payoff - so I can normalize S's period 2 cost of supply to 0. B also receives a benefit, b_1^e , independent of whether trade occurs in period 2. This captures e 's *direct externality* on B. I assume $b_1^h \geq b_1^l$.¹⁸

Specificity Decision. When S and B do not trade jointly in period 2 ($q = 0$), they receive their *alternative trade payoffs*, r_S and r_B . These payoffs are independent of e but depend on a specificity decision determined before investment begins. The specificity decision may be directly contractible (as with exclusivity, joint ownership, purchase of specific assets, and sinking of contractible, specific investments), or it may be indirectly contractible (see 5.5 on authority transfer). I quantify specificity as the difference between the expected surplus from the specific joint trade (b_2^e), and the expected surplus from alternative trading ($r_{sum} \equiv r_B + r_S$). A decision is "intentionally specific" if it raises this measure of specificity beyond its value at the benchmark decision that maximizes B and S's average total surplus ($b_1^{\hat{e}} + \hat{q}b_2^{\hat{e}} + (1 - \hat{q})r_{sum}$) when e and q are fixed at their equilibrium average values, \hat{e} and \hat{q} . Until section 5, I assume specificity has no effect on b_1 and b_2 ,¹⁹ so the benchmark simply maximizes r_{sum} at $r_{sum}^0 \equiv \sup(r_{sum})$. So intentional specificity is captured by $W \equiv r_{sum}^0 -$

r_{sum} . Below, I compare the feasibility of incentive implementation using an intentionally specific technology ($W > 0$) and using a benchmark technology ($W = 0$).

Payoffs. I normalize the time-discount factor to unity and I denote the net cash payment from B to S by t , so that B's overall payoff is $b_1^e + \{q \cdot b_2^e + (1 - q) r_B\} - t$, and S's overall payoff is $-e + (1 - q) r_S + t$. In equilibrium, q and t will depend on the original contract, S's choice of e , contractual messages and renegotiation.

Contracts. B and S can contract on q before investing, but (as standard in incomplete contract theory) they cannot contract on the cross-investment, e , nor its payoff consequences (e , b_1^e and b_2^e). The basic contract dictates a trade-contingent transfer, $T(q)$, that B owes to S.²⁰ B owes $T(0)$ for sure and owes an additional amount, $T(1) - T(0)$, if trade occurs. I denote this price of trade by $p = T(1) - T(0)$. The contract offers B the option of trading at price, $p = T(1) - T(0)$, subject to S being willing to trade.²¹

For generality, I allow B and S to send verifiable messages. I can restrict to truthful revelation mechanisms using a **message game, MG**, where B and S send simultaneous messages, m_B and m_S from $M \equiv \text{support}(e) = \{l, h\}$.²² This message game must be played after B observes e and before trading. A **general contract** $(T(\underline{m}, q))_{\underline{m} \in M^2, q \in \{0,1\}}$ states how the transfer from B to S will depend on the message pair, \underline{m} , from MG and the trade outcome, q , of TG. Note that the *whole* transfer from B goes to S because I impose **Budget-Balance** as a constraint.

Here I assume that costless and efficient (information is symmetric except in 5.4) renegotiation can occur at any time (see 5.1 for renegotiation prohibitions) via the **Renegotiation Gameform, RG: (RGi)** nature selects B (probability, θ) or S (probability, $1 - \theta$); **(RGii)** the selected agent can offer a new contract; **(RGiii)** the other agent accepts or rejects. (Generalized Nash Bargaining gives identical results.) The bargaining power parameter, θ , is exogenous and (except in 5.1) I make,

Assumption 1: *B's exogenous bargaining power is non-trivial: $\theta \in (0, 1]$.*

Timing. B and S can renegotiate immediately after any message game so I place a renegotiation sequence RG (intervening) between MG and TG. This gives the time order, MG-RG-TG.²³ The original contract, $T(\underline{m}, q)$, and specificity decision are agreed before period 1 (to affect incentives for e). Until 5.5, I assume specificity is directly decided in the original contract negotiations. The overall **timing** is presented in figure 1.

stage 3, S gets a fraction $1 - \theta$ of the renegotiation gain, $b_2^e - r_{sum}$, and this *does* depend on e . Hence, S anticipates the continuation payoff of $r_S + (1 - \theta)(b_2^e - r_{sum})$. His incentive is therefore, $I = (1 - \theta)(b_2^h - b_2^l)$ and he chooses $e = h$ if and only if,

$$(1 - \theta)(b_2^h - b_2^l) \geq h \quad (4.1)$$

Even when S has all the bargaining power ($\theta = 0$), this short-term contracting may not permit the first-best. S internalizes B's period 2 benefits, $b_2^h - b_2^l$, but this is too little if $b_1^h - b_1^l > h - (b_2^h - b_2^l) > 0$. Underinvestment can occur because S does not internalize the direct externality, $b_1^h - b_1^l$.

The remaining case has $p \in (b_2^l - r_B, b_2^h - r_B]$. Here, (absent renegotiation) trade only occurs in the subgame following $e = h$: S always prefers to trade but B only wants to trade if $e = h$ (see figure 2, Appendix C, for game tree). B has a non-trivial option so this contract is called an 'option scheme.' After $e = h$, there is no gain from renegotiation, so S wins a continuation payoff of p . Meanwhile, after setting $e = l$, S correctly anticipates that B would reject trade (giving S a default payoff of r_S), so S renegotiates at stage 3 (earning a fraction, $1 - \theta$, of the renegotiation gain, $b_2^l - r_{sum}$). Adding gives S's continuation payoff: $r_S + (1 - \theta)(b_2^l - r_{sum})$. So S's incentive is,

$$I = p - r_S - (1 - \theta)(b_2^l - r_{sum}) \quad (4.2)$$

This option scheme generates a credible carrot of p after a high cross-investment, and a stick of r_S plus S's share of the renegotiation gains. Maximizing p at $b_2^h - r_B$ maximizes I at,

$$I^\theta \equiv b_2^h - \theta \cdot r_{sum} - (1 - \theta)b_2^l$$

This exceeds the incentive of $(1 - \theta)(b_2^h - b_2^l)$ from short-term contracting (see inequality 4.1) by the amount $\theta(b_2^h - r_{sum})$ (> 0 by Assumptions 1 and 3). I record this result as a lemma:

Lemma 1. *Within the class of option contracts, the "exercise price," $p = b_2^h - r_B$, maximizes S's incentive at $I = I^\theta = b_2^h - \theta \cdot r_{sum} - (1 - \theta)b_2^l + \theta \cdot W$.*

Note how I^θ increases with W (at the rate θ -to-1). This captures the essence of the specificity effect, but I first check that more complex, message game contracts cannot evade this upper bound.

Lemma 2. *General contracts cannot raise incentives beyond I^θ from the best option scheme.* The proof of lemma 2 is more involved (appendix A contains all proofs) but the intuition is simple. The highest reward (for $e = h$) that B can credibly promise to S depends on

her equilibrium service value (b_2^h) relative to her opportunity cost (r_B). The punishment (for $e = l$) is S's alternative (r_S) plus his gain from renegotiating after setting low effort ($(1 - \theta) \cdot (b_2^l - r_{sum})$). The *budget balance restriction drives the result*: to punish S by **fining** him for the no-trade outcome is frustrated because S's fines must then go to B, raising her opportunity cost and reducing reward credibility as fast as the punishment severity rises.

There is a demand for raising incentives whenever $h > I^\theta$. By assumption 3, $b_1^h - b_1^l > h - (b_2^h - b_2^l)$ so $h > I^\theta$ implies $b_1^h - b_1^l > \theta(b_2^l - r_{sum}) > 0$ (by assumption 2). So the direct externality, $b_1^h - b_1^l$, is crucial to my specificity result, until I allow for stochastics and two-sided investment (see 5.3). To summarize:

Proposition 1. *In the second-best with direct externalities, intentional specificity is strictly optimal if $r_{sum}^0 > \frac{b_2^h - (1-\theta)b_2^l - h}{\theta}$. Intentional specificity $W \geq r_{sum}^0 - \frac{b_2^h - (1-\theta)b_2^l - h}{\theta}$ is then necessary and sufficient to implement $e = h$.*

The benefit of specificity is that it relaxes the incentive implementation constraints. Here, its cost is zero, but in general, increasing specificity reduces average productivity (e.g., because separation is sometimes *ex post* optimal). There is then a simple tradeoff (see 5.4) between incentive benefits and productivity costs. The key observation is that the incentive benefit from specificity is increasing in the direct externality. So, after generalizing the model below, I use empirical estimates (in section 6) of this externality to predict specificity.

5. Generalizations and Robustness

This section assesses robustness to some natural generalizations and alternative assumptions. I show that my result depends on the restriction of budget-balance and not on renegotiation. Uncertainty, two-sided investment and limited advance contractibility of trade can weaken the incentive scheme, but I find that these generalizations can simultaneously strengthen my argument for specificity.

5.1. Budget-Balance and Renegotiation

To prove the critical role of budget-balance, I need to establish that: (a) there is no incentive problem when budget-breaking is feasible (so renegotiation is not sufficient for the specificity result), and (b) specificity may be needed even when renegotiation can be prohibited (so renegotiation is not necessary either). The first claim is evident from observing that a fine, F , owed by B or S to a budget-breaker after $q = 0$, is incentive-equivalent to specificity level, $W = F$ (and has no equilibrium cost).²⁵

Proposition 2. *Budget-breaking fines payable to third parties would permit the first-best and make specificity redundant (in spite of renegotiation).*

The second claim follows from proving, in Lemma 3, that preventing renegotiation permits an incentive increase from I^θ to I^1 ; substituting I^1 for I^θ in Proposition 1 reveals the need for specificity, formally stated as Proposition 3.²⁶

Lemma 3. *When renegotiation can be prohibited, S's maximal incentive from the class of all budget-balanced contractual mechanisms is $I^1 = b_2^h - r_{sum}$ (independent of θ).*

Proposition 3. *When renegotiation can be prohibited, intentional specificity $W \geq b_2^h - h - r_{sum}^0$ is strictly optimal if $b_2^h - h - r_{sum}^0 > 0$.*

Note the impact of changing bargaining power - **Observation 1:** *The maximal incentive, I^θ , decreases when S's bargaining power $1 - \theta$ rises.* This is a little surprising: incentives are greater when B has the bargaining power even though it is S that makes the noncontractible investment. Mathematically, I^θ rises with θ because $b_2^l > r_{sum}$. The intuition is that the surplus from renegotiation is only positive after S underinvests, thereby inducing B to reject the trade option. So S's bargaining power only raises S's payoff after low investment. Therefore it reduces S's incentives."

5.2. Uncertainty, Continuity and Additional Investment Problems

The model readily adapts to a continuum of effort choices, $e \in [0, \infty)$, and to uncertainty in the joint trade value, b_2^e . In the deterministic case, $p = b_2^{e^*} - r_B$ implements e^* provided that $e^* - e' \leq b_2^{e^*} - (1 - \theta) b_2^{e'} - \theta \cdot r_{sum} \forall e' < e^*$. So intentional specificity of W relaxes this implementation constraint by $\theta \cdot W$. In the stochastic case with symmetric information, an option price $p > r_S$ motivates S to raise e if this increases the reward probability, $\sigma(p, e) \equiv \Pr(b_2^e \geq p + r_B)$. This reward function can be maintained with higher incentives (up by $\theta \cdot W$) via intentional specificity, W (since $\sigma(p, e; r_B, r_S) \equiv \sigma(p + W_B, e; r_B - W_B, r_S - W_S)$). The effectiveness of the option scheme is nonetheless reduced by excessive uncertainty (see Nöldeke and Schmidt's, 1998, example with additive uncertainty and see 5.3).

When B makes a non-contractible and specific self-investment²⁷, i , the reward probability in the option contract, becomes $\sigma(p, e, i) \equiv \Pr(b_2^{e,i} \geq p + r_B)$. The intuitive concern is that B might decrease i so as to be able to threaten S with no trade (and bargain a lower price). Indeed, B can decrease σ by reducing i . However, B gains the full marginal return on i when terminating trade is non-credible and B gains a share θ of these returns when termination is credible. This pair of marginal effects can dominate the disincentive caused by the possibility of affecting termination credibility. The option scheme can therefore remain

a serious alternative (see 5.3).

When S's investment has self effects as well as cross effects - e.g., S's cost of performing the period 2 trade, c_2 , might decrease in e - the maximal incentive I^θ increases by $c_2^l - c_2^h$ (deterministic case). So the option scheme is robust.

When advance contracting on trade is entirely impossible, B and S cannot use the standard option scheme. Short-term contracting achieves $I = (1 - \theta)(b_2^h - b_2^l)$, but can be improved via an adapted option contract that gives S the right to define the terms and specifications of the trade option in period 2. S then extracts the full expected surplus from period 2 trade so $I = E(b_2^h - b_2^l)$ ($E(\cdot)$ denotes the common period 1 expectations). Specificity cannot help. Fortunately, zero advance contractibility is extremely rare. The option scheme only needs that e is a cross-investment with respect to *some* period 2 trade contract that can be written at stage 0 - i.e., a contract with payoffs, $b_2^h > b_2^l$.²⁸ The maximal credible option price falls to $p = b_2^h - r_B$ so I^θ falls by $b_2^h - b_2^l$, but remains dominant if intentional specificity is high enough.

5.3. Alternative Demands For Specificity

Specificity can be strictly advantageous even when the cross-investment has no direct externality (i.e., $b_1^h = b_1^l$). (a) Uncertainty can reduce the option contract incentives below the expected investment return, $E(b_2^h - b_2^l)$. This does not justify specificity if information is symmetric because the adapted option contract in 5.2 already provides an incentive of $E(b_2^h - b_2^l)$, but when B has private information, S cannot extract the full surplus and incentives may be insufficient. The option scheme from section 4 is then optimal with sufficient specificity to increase incentives (see Appendix B and Ellman, 1999 and 2004, for concrete examples). (b) With two-sided investment, the adapted option contract gives no incentive for B to invest at all, so the option contract solution can be optimal provided sufficient specificity is available.

5.4. The Costs and Reversibility of Specificity

In the basic model, specificity has no cost because joint productivity is unaffected and joint trade always occurs in equilibrium. In general, intentional specificity reduces average productivity (which by definition, it cannot increase). For instance, relaxing assumption 2 so r_{sum} can exceed b_2^e in equilibrium, specificity W causes average losses of $E(\min(W, \max(r_{sum}^0 - b_2^e, 0)))$. The marginal cost of specificity W is usually much below

unity,²⁹ but may outweigh its incentive benefit in less stable relationships, so the model predicts more specificity in more stable relationships (inverting the usual causal link from specificity to stability).

Exclusivity restrictions and ownership structures that reduce separation surplus generate a special type of intentional specificity, because the waste can often be avoided through agreement after separation. Restrictions can be removed, assets retraded or efficient uses agreed. This reversal of the specificity to escape the waste is a great advantage in unstable relationships, but the incentive effect from this “reversible” specificity is slightly less robust. In the gradual bargaining model, the threat of waste is not always credible. It depends on the nature of alternative trading. Alternative trades are called “inside options” if they do not change the value of returning to joint trade, and “outside options” if they effectively preclude switching back to joint trade (see Osborne and Rubinstein, 1990). If the relevant alternative trades (given any exclusivity restrictions and reigning ownership structure) are inside options, the threat of waste is credible, because (whenever joint trade is optimal) reversing exclusivity or adjusting ownership is irrelevant - it is not “on the bargaining table.”³⁰ However, in a subgame following exercise of an outside option, joint trade is ruled out, so B and S would try to renegotiate exclusivity and ownership if this can avoid the waste.³¹

5.5. The Contractibility of Specificity

So far, I have assumed that agreement to implement intentional specificity can be enforced with no difficulty. This is reasonable for specificity created by an exclusive contract or purchase of a specific asset, but in many examples specificity is determined by private actions of B and S that are not directly contractible. Provided S can observe the specificity before making its cross-investment, B may generate specificity to increase S’s incentives.³² However, the main factor determining private preferences over specificity is rent-seeking: each trader can strengthen its relative bargaining position by improving its own alternatives (e.g., S increasing r_S through general investments) and by damaging its partner’s alternatives (e.g., S reducing r_B by making itself indispensable). The first effect reduces specificity and the second effect increases specificity. The level of specificity depends on which effect dominates. This in turn depends on the authority structure. So specificity can be adjusted indirectly, by negotiating shifts in the authority structure. For instance, vertical integration that turns S from an independent partner into an employee of B, increases B’s authority and B may use her new discretion to raise the specificity of the skills S learns on the job. This integration is then a way to commit to higher specificity. More generally, B and S can adjust specificity

by negotiating transfer of (bundles of inseparable) control rights (see in Aghion et al, 2002). On the other hand, standard organizational theories (see e.g., Williamson, 1985, and Hart, 1995) may dominate the specificity motive in driving these decisions, as discussed below. These alternative theories may explain specificity as a side-effect (see below).

6. Empirical Application: Franchising

In this section, I use Bercovitz’s (2000) detailed empirical evidence from franchising to test my theory. The contractual and institutional details of franchising offer an encouraging real-world example of exactly how option contracting works. “Free-riding” is a significant problem in franchising: each franchise unit’s investments in service quality has a positive externality (via the reputation of the trademark) on other units and particularly on the franchisor (who receives royalties from all franchisees, up-front fees from new franchisees, and owns company units). As noted in the introduction, the externality is partly direct (trademark value rises independent of continued joint trade) and partly contingent on continued trade (i.e., a cross-investment, because the franchisee’s efforts generally increase its ability and motivation to supply high quality service in the future). Given the prominence of these two components of my theory, it is encouraging to find that franchising contracts generally fit well with the option contract solution, as I now show.

A typical franchising contract specifies detailed procedures that the franchisee must follow, and obliges the franchisor to train the franchisee, but it cannot enforce many important efforts. To cope with this two-sided moral hazard, the contract specifies a sharing rule whereby the franchisor has the right to a fraction of the franchisee revenues (the “royalty rate”).³³ Franchising contracts are usually several years long and the franchisor seeks to retain the right to terminate (with at most a moderate penalty).³⁴ This creates an option contract: the franchisee’s reward from continued trade with the franchisor is decreasing in the royalty rate and ongoing fees, so fixing these terms of trade is equivalent to fixing p in the basic model.³⁵

Even when the terms of trade are not explicitly fixed by contract, they may be effectively fixed by a franchise policy or legal obligation to treat all franchisees in the same way (see evidence on homogeneity restrictions in Lafontaine and Oxley, 2004) or to avoid downward shifts in any franchisee’s terms of trade. Both these restrictions (homogeneity and downward rigidity) may also prevent bilateral renegotiation. If the franchisee cooperates in advance (e.g., to create evidence that special factors require changes, or to self-enforce

hidden changes), the restrictions might be evaded, but by withholding cooperation, the franchisee can certainly refuse to renegotiate over a non-trivial time interval. Ellman (2004) shows that this can prevent the problematic effect of renegotiation suggested by Che and Hausch (1999).³⁶ Since proposition 3 predicts specificity even when renegotiation is totally prohibited, both cases support my theory.

So all the theory components fit. It only remains to test whether intentional specificity increases with the size of the free-riding externality. Klein (1980) and Williamson (1985) identified various franchising arrangements that had been criticized for infringing on the franchisee's alternative options, (e.g., requiring short-term land-leasing and specific asset purchase).³⁷ I focus on the specificity created by non-compete covenants on which Bercovitz (2000) has built a systematic data set. A post-termination non-compete covenant limits the franchisee's ability to operate any competitive business in a neighboring area for a defined duration of time after termination. The covenant generally decreases the franchisee's alternative trade payoff r_S by more than reduced local competition increases the franchisor's alternative payoff r_B . So it lowers r_{sum} and offers a contractible method of increasing specificity W . The expansiveness of a covenant, measured by multiplying its area and duration, is therefore a plausible proxy for W .³⁸

My theory predicts that this measure of specificity will correlate positively with the size of the franchisee's quality/free-riding externality. Bercovitz (2000) proxies this externality by interacting a proximity-based measure of reputational spillover with a proxy for brandname importance.³⁹ In her regression analysis, the free-riding proxy has a positive and highly significant effect on the expansiveness of non-compete covenants ($p < 0.002$).⁴⁰

This confirms the predictions of the specificity theory, but can alternative theories of exclusivity restrictions also explain this correlation? Segal and Whinston (2000) generate several alternative explanations of exclusivity⁴¹ that build on Areeda and Kaplow (1988), Marvel (1982) and Masten and Snyder (1993). The key idea is that when investments have "external effects," exclusivity restrictions can help by: (a) inducing holdup *of* the restricted party - e.g., to prevent retail franchisees from investing in promotion of other companies' products (see Areeda and Kaplow, 1988); (b) preventing holdup *by* the restricted party - e.g., a manufacturer, unable to trade with its major supplier's competitors, cannot expropriate knowledge learned from that supplier (see Masten and Snyder, 1993). Specificity could then be a side-effect of exclusivity restrictions aimed at changing the marginal default return to particular investments (as in the formal work summarized in Hart, 1995), rather than directly aimed at shifting the difference in the absolute surpluses from joint and separate trading (as

in my theory).

So could the non-compete covenants in Bercovitz's (2000) data be driven by the need to resolve problematic external effects of key investments? The answer is that there could be external effects, but an alternative explanation of the observed correlation requires evidence that these external effects are more important precisely when free-riding *within* the franchise is more severe.⁴² Since non-competes prevent a franchisee from exploiting its local reputation after a termination, they are likely to worsen the free-riding problem from the perspective of external effects. The most obvious application of this perspective therefore suggests the opposite correlation. By contrast, my intentional specificity theory permits a direct and convincing explanation. In conclusion, my theory receives strong support from this data.⁴³

7. Conclusion

Intentional specificity *can* be explained even in settings where long-term commitments to bilateral trade are feasible. The key condition is that one party makes cross-investments - i.e., long-term contracts cannot prevent investment externalities. In my theory, specificity reinforces the long-term contracts that motivate cross-investments. It serves as a substitute for 'third party budget-breakers' and is needed when the cross-investments have direct externalities - a widespread, though neglected, phenomenon. Specificity may also be needed when the contractual parties suffer from asymmetric information or both make non-contractible investments. Renegotiation is neither necessary or sufficient. Instead, the budget-balance restriction is central to my theory.

These results prove that empirical observation of long-term contracting does *not* preclude strategic explanations of specificity as a device for enhancing incentives. So my theory provides a strong foundation for the regulatory debate in franchising initiated by Klein (1980), Williamson (1983) and Frasco (1991). The logic of my theory is closest to that of Klein and Leffler (1981), though they study a market setting with neither contracting nor renegotiation. I have captured related intuitions in a contractual setting with a more robust analysis based on unique equilibria and not reliant on infinite-horizon, folk-theorem arguments.

My preliminary survey of the rich evidence on franchise contracting reveals that option contracts are standard, and suggests that franchisors successfully use them to deal with free-riding externalities, because they (like other parties with multiple partners or agents) can readily escape the renegotiation problems predicted by Che and Hausch (1999). Furthermore,

the evidence on non-compete covenants in Bercovitz's (2000) franchising data confirms the predictions of my theory, while competing organizational theories do not offer convincing alternative explanations.

In my theory, specificity is equally effective in raising cross-investment incentives, whether it reduces the alternative trade payoff of the cross-investor (raising the severity of the cross-investor's punishment) or of the other party (raising the credibility that this party rewards high investment). The simplicity of the resulting, surplus-based definition of intentional specificity should facilitate the measurement task needed for further testing of the theory. However, the specificity concept draws together a variety of apparently distinct phenomena (such as sunk costs, specific assets, exclusivity restrictions and wasteful ownership structures), so each case will require careful analysis of different alternative explanations.

The implications of this theory for authority transfers (e.g., via vertical integration) that affect specificity present promising avenues for further research. It will be vital to gather evidence and carefully sift through alternative arguments from the property rights and transactions cost literatures. The theory also promises to contribute to the topic of internal design in organizations. The options (e.g., from 'at will' contracts) at the core of my analysis offer powerful incentive devices when combined with designs that induce specificity. In this way, my approach to specificity fits into the broader project in which theorists are beginning to address the twin questions of incentives and control rights within a unified framework.

Appendix A

Proof of Lemma 2. For any implementation mechanism represented by contract $T(\cdot)$, renegotiation after effort level e and messages (m_B, m_S) induces trade with an adjusted transfer function $\bar{T}^e(m_B, m_S)$ given by,

$$\bar{T}^e(m_B, m_S) = \begin{cases} T(1, m_B, m_S), & \text{for } r_S \leq T(1, m_B, m_S) - T(0, m_B, m_S) \leq b_2^e - r_B \\ T(0, m_B, m_S) + \theta \cdot r_S + (1 - \theta)(b_2^e - r_B), & \text{otherwise} \end{cases}$$

In the first row case, B and S both want to trade (the first and second inequalities are S's and B's, respective, participation constraints) but in the second row of this box, renegotiation is needed for trade. Because renegotiation always leads to trade, S's period 2 payoff is exactly given by $\bar{T}^e(m_B, m_S)$.

I can restrict to truthful implementation mechanisms - see Maskin and Moore (1999). I define $\bar{t}^h \equiv \bar{T}^h(h, h)$ and $\bar{t}^l \equiv \bar{T}^l(l, l)$, S chooses e at stage 0, to maximize his anticipated overall payoff. In subgame perfect equilibrium, he sets $e = h$ if and only if $h \leq \bar{t}^h - \bar{t}^l$. So his incentive is $I = \bar{t}^h - \bar{t}^l$. I now use the message game incentive compatibility conditions to derive constraints on I - I must have: $\bar{t}^h \leq \bar{T}^h(l, h)$ else B would announce l at stage 2 after $e = h$; and $\bar{t}^l \geq \bar{T}^l(l, h)$ else S would announce h and not l after $e = l$. Now, setting $p = T(1, l, h) - T(0, l, h)$, I have,

$$\bar{T}^h(l, h) - \bar{T}^l(l, h) = \begin{cases} (1 - \theta)(b_2^h - b_2^l) & \text{for } r_S \leq p \leq b_2^l - r_B \\ p - \theta \cdot r_S - (1 - \theta)(b_2^l - r_B) & \text{for } p + r_B \in (b_2^l, b_2^h] \\ 0 & \text{if } p < r_S \text{ or } p > b_2^h - r_B \end{cases}$$

which is maximized at $p = b_2^h - r_B$ so $\bar{t}^h - \bar{t}^l \leq b_2^h - \theta(r_B + r_S) - (1 - \theta)b_2^l = I^\theta$. Thus the incentive is no higher than I^θ . As stated in Lemma 1, \hat{I} can be implemented by setting $T(1, \cdot, \cdot) \equiv p = b_2^h - r_B, T(0, \cdot, \cdot) \equiv 0$. ■

Proof of Proposition 1. $I^\theta(W) = b_2^h - \theta \cdot r_{sum}^0 - (1 - \theta)b_2^l + \theta \cdot W \geq h$ if and only if

$$W \geq r_{sum}^0 - \frac{b_2^h - (1 - \theta)b_2^l - h}{\theta}$$

This requires $W > 0$ whenever $r_{sum}^0 > \frac{b_2^h - (1 - \theta)b_2^l - h}{\theta}$, and such W is then strictly optimal since it raises the implementable effort level from $e = l$ to the efficient level $e = h$, and has no costs (as joint trade is always optimal). ■

Proof of Proposition 2. The option contract described in proposition 1 is used in conjunction with a fine F_B that is owed to a third party by B, and a fine F_S owed to the third party by S, whenever $q = 0$. The introduction of the fines is equivalent to reducing B and

S's alternative trade payoffs to $r_B^0 - F_B$ and $r_S^0 - F_S$. It follows that the fines are equivalent to introducing an intentional specificity, $W = F$. The optimized incentive is therefore given by $I = b_2^h - \theta \cdot r_{sum} - (1 - \theta) b_2^l + \theta \cdot F$ and can be increased arbitrarily high. Collusion with the third party is assumed impossible, but B and S can renegotiate so F is never paid out in equilibrium, even in stochastic extensions of the model. ■

Proof of Lemma 3. The Revelation Principle justifies reduction to truthful message games and full Nash implementation (see Maskin and Moore, 1999). Let ρ_{m_B, m_S}^e denote the probability of trade after effort level e and messages m_B, m_S , and let v_{m_B, m_S}^e denote B's equilibrium payoffs, and u_{m_B, m_S}^e denote S's equilibrium payoffs, from the continuation game beginning in period 2. By the Budget Balance restriction,

$$v_{h,h}^h + u_{h,h}^h = \rho_{h,h}^h \cdot b_2^h + (1 - \rho_{h,h}^h) \cdot r_{sum}; (Eq.1)$$

$$v_{l,h}^l + u_{l,h}^l = \rho_{l,h}^l \cdot b_2^l + (1 - \rho_{l,h}^l) \cdot r_{sum}; (Eq.2)$$

By incentive compatibility in the message game, $v_{l,h}^h \leq v_{h,h}^h$ (*Ineq.1*) and $u_{l,h}^l \leq u_{l,l}^l$ (*Ineq.2*). Using (*Ineq.2*),

$$I \equiv u_{h,h}^h - u_{l,l}^l \leq u_{h,h}^h - u_{l,h}^l$$

and by (*Eq.1*) and (*Eq.2*),

$$u_{h,h}^h - u_{l,h}^l = [\rho_{h,h}^h \cdot b_2^h + (1 - \rho_{h,h}^h) \cdot r_{sum} - v_{h,h}^h] - [\rho_{l,h}^l \cdot b_2^l + (1 - \rho_{l,h}^l) \cdot r_{sum} - v_{l,h}^l]$$

Combining and simplifying using (*Ineq.1*) yields,

$$\begin{aligned} I &\leq \rho_{h,h}^h (b_2^h - r_{sum}) - \rho_{l,h}^l (b_2^l - r_{sum}) \\ &\leq \rho_{h,h}^h (b_2^h - r_{sum}), \text{ as } \rho_{l,h}^l \geq 0 \text{ and } b_2^l \geq r_{sum} \\ &\leq (b_2^h - r_{sum}), \text{ as } \rho_{h,h}^h \leq 1 \text{ and } b_2^h \geq r_{sum} \end{aligned}$$

This maximum $b_2^h - r_{sum}$ equals I^1 as stated. That I^1 can indeed be achieved for any θ , is clear from observing that $I = b_2^h - r_{sum} = I^1$ from the option scheme, $p = b_2^h - r_B$ (setting $\theta = 1$ is equivalent to omitting S's renegotiation gains in the proof of Lemma 1). ■

Notes

¹Segal and Whinston (2000) does not contradict this result, because they assume a fixed probability of joint trade.

²Williamson (1983) emphasizes how limited wealth and liability may constrain enforcement of damages, and limited rationality constrains description of future trades.

³Also called a “cooperative investment” (Che and Hausch, 1999): one party raises the *other* party’s trade payoff.

⁴The buyer-seller terminology is not to be taken too literally: in the franchising example, the franchisee is *selling* restaurant-management services, but also buying the right to use the franchise trademark.

⁵E.g., specificity from the buyer agreeing to buy an unnecessarily specific asset.

⁶E.g., from the seller agreeing to sign a non-compete covenant (see section 6).

⁷See Holmström (1982) and Eswaran and Kotwal’s (1984) collusion arguments against budget-breaking.

⁸Nöldeke and Schmidt (1998) and Watson and Schwarz (2003) write “direct externality” to describe cross-investment, but that externality is trade-contingent.

⁹As noted in Segal and Whinston (2000, footnote 9), such distortions imply equilibrium waste. In section 6, I present alternative organizational theories generating specificity as a side-effect.

¹⁰Such a reputation (of B) can be enforced by outsiders if S alerts them when B tries to force S to renegotiate - Jolls (1997) reports that even judges restrict renegotiation induced through potentially “coercive” threats. See also Ellman (1999, §2.5, and 2004), Lyon and Rasmusen (2004), Nöldeke and Schmidt (1998) and Watson (2003) for the view that option schemes are robust to renegotiation.

¹¹Most other theories are context-dependent (and they require identification of the marginal specificity of particular investments, not just changes in the overall level of specificity).

¹²Similarly, firms follow trigger-strategies (firing workers for slacking) in Shapiro and Stiglitz (1984), and unemployment, which raises specificity by reducing employee alternatives, raises their efficiency wage incentive.

¹³For instance, a firm that needs its supplier to make cross-investments may want to raise relationship-specificity by imposing an idiosyncratic technology (or even by fostering a specific corporate culture) in the supply unit. In order to impose this technology or culture on the unit managers and workers, the firm may need to acquire it.

¹⁴The potential waste is crucial because if B gains the value (9) lost by S, then her opportunity cost of retaining S rises by 9 and the carrot must be cut by 9 to maintain credibility, exactly undoing the gain from increasing the stick.

¹⁵Principal-agent interpretations are equally valid, e.g. S could be an agent ‘selling’ labor services as in Kahn and Huberman (1988). Technically, I am adding direct externality, renegotiation, and an explicit model of trading to Demski and Sappington (1991).

¹⁶Usually $q = 0$ represents a separate trade but can be a default trade like non-promotion (see Prendergast, 1993).

¹⁷Simultaneity gives the same results, so I treat TG as instantaneous.

¹⁸Recall the franchising example and the idea that if S’s investment is from learning-by-doing, it rises with increased “doing” (or on-the-job effort). Similarly in consulting, a company (S) may prepare by asking questions that help the client company (B) to help itself, as well as raising the client’s future benefits from extending the consultant’s contract (via $q = 1$). Further, note that whenever a buyer B uses quality of a preliminary service as a proxy for future service quality, it must be that e raises B’s benefit in both periods.

¹⁹Even a specific investment simply lowers r_{sum} when B pays contractible investment costs in period 2.

²⁰Any contractual terms enforcing period 1 activity are left implicit.

²¹This constraint does not bind. In period 2, e is only payoff-relevant to B, so S’s trade preferences do not add implementation power. This is why separate verifiability of trading decisions - allowing $T(\cdot)$ to depend on q_B and q_S (as implicit in Demski and Sappington, 1991) instead of only q - does *not* resolve the implementation problem.

²²B and S both observe the non-verifiable event (S’s choice of e) and I assess weak Nash-implementability.

²³There is nothing to renegotiate after TG, as trade must occur at stage 4 (see Ellman, 2004, for the gradual trade deadlines model). Renegotiation prior to MG-RG-TG is superfluous as RG-TG guarantees ex post trade efficiency.

²⁴I assume cooperative behavior (high effort, trade, offer-acceptance) under indifference.

²⁵The budget-breaker must observe q or be alerted automatically (via a message game).

²⁶ I^1 is also feasible when the period 2 timing shifts from MG-RG-TG to RG-MG-TG. This is easily seen using MG': (i) B offers new contract; (ii) S accepts or rejects, because MG'-TG is identical to RG-TG with $\theta = 1$.

²⁷This is a natural assumption in my applications. Two-sided cross-investment is more problematic.

²⁸The optimal prior contract uses terms (nature and quantity of trade) that maximize b_2^h , though stochastic effects may motivate use of a menu contract (with non-linear pricing).

²⁹The separation probability is less than 1 (else $e^* = 0$) and intentional specificity may increase joint productivity (as only average productivity must weakly fall). Note also that my specificity result applies even when joint trade is always efficient, unlike Williamson (1983) where specificity must raise the probability of joint trade.

³⁰This claim is valid in standard bargaining games where offer-makers prefer to make efficient offer, and also by the axioms of the Nash Bargaining solution.

³¹Reversible specificity is ineffective in the extreme case where the relevant alternative trades are always outside options (see Ellman, 1999). Note by contrast that reversible specificity *never* helps in Williamson's (1983) theory (without liquidity constraints) because the trade probability is no longer distorted.

³²S might also create specificity when B observes it before signing the incentive contract. (See also Raub and Keren's, 1993, analysis of prisoners' dilemma games embedded in a hostage-posting game.)

³³See Bhattacharyya and Lafontaine (1995), Lafontaine (1992) and Rubin (1978).

³⁴In the U.S., some states restrict franchisors' ability to terminate by imposing a requirement of proving "good cause" (in spite of experts like Rubin, 1978, arguing for termination "at will"), but franchisors can often respond by writing shorter contracts (at renewal dates, they only need to show "lack of good standing" - see Klein, 1980, and Bercovitz, 2000, who finds that contracts are shorter when externalities are greater, confirming the theory). When the right to terminate is sufficiently constrained, the option contract is not feasible, the free-riding problem is very hard to solve and franchisors tend to avoid franchising in such states (as verified in Brickley et al., 1991).

³⁵Franchisee investments (including the quality investment) usually have a self-investment effect, because a franchisee with lower costs or better local reputation earns more profits under a fixed revenue sharing rule. This effect is unproblematic (see 5.2) and the reward implied by fixed trading terms increases with experience over time.

³⁶According to my theory, franchisors may voluntarily build a reputation for contractual homogeneity or downward rigidity to protect option scheme incentives in the face of renegotiation. C.f., Bhattacharyya and Lafontaine's (1995) alternative explanation of voluntary homogeneity.

³⁷Klein (1980) and Williamson (1983) reject the legal presumption that such arrangements are necessarily "unfair," because the practises might be agreed to increase efficiency. My theory provides a strong foundation for this argument despite the feasibility of long-term contractual commitments.

³⁸Law courts restrict the size of W , because they only uphold non-compete covenants that satisfy a test of reasonableness. As explained in Bercovitz (2000) (motivated by Brazener, 1998), the courts effectively tradeoff allowable time and geographical constraints (hence the multiplicative operationalization of the expansiveness variable). When upheld, non-compete covenants are usually not renegotiated owing to franchisor reputations for non-renegotiation and because multiple neighbouring franchisees are effected. So the reversibility issue of 5.4 does not arise here.

³⁹She measures the average distance between each outlet and its three closest outlets: Nearby outlets share customers, so brandname perceptions depending on one outlet's quality often affect another's revenue. She estimates brandname importance with recent advertising expenditures: customer sharing is more serious when the brandname's role is more important.

⁴⁰She uses the Seemingly Unrelated Regression (in case of contemporaneous correlation in error terms) to regress expansiveness of non-competes on the free-riding proxy, and on a proxy for franchisee specific investments. (The second proxy, based on franchisor estimates of franchisee set-up costs, as in Dnes, 1992, has no significant effect.) As a control, she measures the fraction of franchisees in states that restrict termination. The option scheme may become impossible in these states, so my argument for generating specificity disappears. The highly significant negative effect of this control on use of non-competes is therefore consistent with my theory (though perhaps, as argued by Bercovitz, 2000, non-competes are avoided because also restricted in such states). Verifying robustness of the free-riding proxy, Bercovitz (2000) also shows that increases in the spillover proxy raise non-compete expansiveness when brandname expenditure is fixed, and this relationship becomes steeper

at higher brandname expenditures.

⁴¹They reject Klein (1980) and Frasco's (1991) claim that exclusivity can improve the incentives for fully specific investments. In contrast, by allowing for cross-investments (excluded from Segal and Whinston, 2000) and option contracting, my analysis provides a supportive foundation.

⁴²Also the alternative theory must be consistent with the option contracts identified above.

⁴³Liquidity constraints present another possible explanation of specificity. For instance, a company might offer specific training to indirectly oblige its liquidity-constrained workers to 'pay' for the training through low future wages. However, liquidity constraints are not a pronounced problem for franchisees relative to franchisors, and are unlikely to have significant correlation with Bercovitz's (2000) free-riding proxy (given the unit is operating).

⁴⁴Even were $h < p \cdot Y$, S's bargaining power causes average waste of $(1 - p)y$ - avoided fully in the standard option contract solution (when $\theta = 1$). And even were $p \cdot Y < y$, S would only extract y which may also be too low for incentives to cover h .

References

- Aghion, Philippe, Matthias Dewatripont and Patrick Rey, 2002, "On Partial Contracting," *European Economic Review*, 46:745-53.
- Areeda, Philip and Louis Kaplow, 1998, "Antitrust Analysis: Problems, Text, Cases," 4th ed. Boston: Little, Brown & Co.
- Arya, Anil, Hans Frimor and Brian Mittendorf, 2003, "Pouring Money Down the Drain? How Sunk Investments and Signing Bonuses can Improve Employee Incentives," *Topics in Economic Analysis and Policy*, 3(1)6.
- Bercovitz, Janet, 2000, "An Analysis of Contract Provisions in Business Format Franchise Agreements," unpublished draft, University of Duke.
- Bhattacharyya, Sugato and Francine Lafontaine, 1995, "Double-Sided Moral Hazard and the Nature of Share Contracts," *Rand Journal of Economics*, 26:761-81.
- Brazener, Robert, 1998, "Validity and Construction of Restrictive Covenant Not to Compete Ancillary to Franchise Agreement," *American Law Reports*, 50:746.
- Brickley, James, Frederick Dark and Michael Weisbach, 1991, "The Economics Effects of Franchise Termination Laws," *Journal of Law and Economics*, 101-132.
- Che, Yeon-Koo and Donald Hausch, 1999, "Cooperative Investments and the Value of Contracting," *American Economic Review*, 89(1):125-47.
- Demski, Joel and David Sappington, 1991, "Resolving Double Moral Hazard Problems with Buyout Agreements," *Rand Journal of Economics*, 22,:232-240.
- Dnes, Anthony, 1992, "Unfair Contractual Practices and Hostages in Franchise Contracts," *Journal of Institutional and Theoretical Economics*, 148, 484-504.
- Edlin, Aaron and Benjamin Hermalin, 2000, "Contract Renegotiation and Options in Agency Problems," *Journal of Law Economics and Organization*, 16:395-423.
- Ellman, Matthew, 1999, "Essays on the Economics of Organization," Harvard University Doctoral Dissertation.
- Ellman, Matthew, 2004, "The Impact of Renegotiation on Option Schemes," unpublished draft, Universitat Pompeu Fabra.
- Eswaran, Mukesh and Ashok Kotwal, 1984, "The Moral Hazard of Budget-Breaking," *Rand Journal of Economics*, 84(15):578-81.
- Frasco, Gregg, 1991, "Exclusive Dealing: A Comprehensive Case Study," London: University Press of America.
- Grout, Paul, 1984, "Investment and Wages in the Absence of Binding Contracts: A Nash Bargaining Approach," *Econometrica*, 52:449-60.
- Hart, Oliver, 1995, "Firms, Contracts and Financial Structure," Oxford: Oxford University Press.

- Hart, Oliver and John Moore, 1988, "Incomplete Contracts and Renegotiation," *Econometrica*, 56(4):755-85.
- Holmström, Bengt, 1982, "Moral Hazard in Teams," *Bell Journal of Economics*, 13(2):324-40.
- Jolls, Christine, 1997, "Contracts as Bilateral Commitments: A New Perspective on Contract Modification," *Journal of Legal Studies* 26(1):203-37.
- Kahn, Charles and Gur Huberman, 1988, "Two-sided Uncertainty and 'Up-or-Out' Contracts," *Journal of Labor Economics*, 6(4):423-444.
- Klein, Benjamin, 1980, "Transaction Cost Determinants of 'Unfair' Contractual Relations," *American Economic Review Proceedings*, 70:356-62.
- Klein, Benjamin, Robert Crawford and Armen Alchian, 1978, "Vertical Integration, Appropriate Rents, and the Competitive Contracting Process," *Journal of Law and Economics*, 21:297-326.
- Klein, Benjamin and Kevin Leffler, 1981, "The Role of Market Forces in Assuring Contractual Performance," *Journal of Political Economy*, 89:615-41.
- Lafontaine, Francine, 1992, "Agency Theory and Franchising: Some Empirical Results," *Rand Journal of Economics*, 23(2):263-283.
- Lafontaine, Francine and Joanne Oxley, 2004, "International Franchising Practices in Mexico: Do Franchisors Customize their Contracts?" *Journal of Economics and Management Strategy*, 13(1):95-120.
- Lyon, Thomas and Eric Rasmusen, 2004, "Buyer-Option Contracts Restored: Renegotiation, Inefficient Threats, and the Hold-Up Problem," *Journal of Law, Economics, and Organization*, 20(1):148-69.
- MacLeod, Bentley and James Malcomson, 1993, "Investments, Holdup and the Form of Market Contracts," *American Economic Review*, 83(4):811-37.
- Marvel, Howard, "Exclusive Dealing," *Journal of Law and Economics*, 1982, 25(1):1-25.
- Maskin, Eric and John Moore, 1999, "Implementation with Renegotiation," *Review of Economic Studies*, 66:39-56.
- Masten, Scott and Edward Snyder, "United States v. United Shoe Machinery Corp.: On the Merits," *Journal of Law and Economics* 36:33-70.
- Nöldeke, Georg and Klaus Schmidt, 1998, "Sequential Investments and Options to Own," *Rand Journal of Economics*, 29(4):633-653.
- Osborne, Martin and Ariel Rubinstein, 1990, "Bargaining and Markets," San Diego, CA: Academic Press.
- Prendergast, Canice, 1993, "The Role of Promotion in Inducing Specific Human Capital Acquisition," *Quarterly Journal of Economics*, 108(2):523-34.
- Raub, Werner and Gideon Keren, 1993, "Hostages as a Commitment Device: A Game-Theoretic Model and an Empirical Test of Some Scenarios," *Journal of Economic Behaviour*

and Organization 21:43-67.

Rotemberg, Julio and Garth Saloner, 1994, "Benefits of Narrow Business Strategies," *American Economic Review*, 84(5):1330-49.

Rotemberg, Julio and Garth Saloner, 2000, "Visionaries, Managers, and Strategic Direction," *Rand Journal of Economics*, 31(4):693-716.

Rubin, Paul, 1978, "The Theory of the Firm and the Structure of the Franchise Contract," *The Journal of Law and Economics*, 21:223-233.

Schelling, Thomas, 1960, "The Strategy of Conflict," Harvard University Press.

Schwartz, Alan and Joel Watson, 2003, "The Law and Economics of Costly Contracting," *Journal of Law, Economics, and Organization*, forthcoming.

Segal, Ilya and Michael Whinston, 2000, "Exclusive Contracts and Protection of Investments," *Rand Journal of Economics*, 31(4):603-33.

Shapiro, Carl and Joseph Stiglitz, 1984, "Equilibrium Unemployment as a Worker Discipline Device," *American Economic Review*; 74(3):433-44.

Watson, Joel, 2003, "Contract, Mechanism Design, and Technological Detail," unpublished draft, University of San Diego.

Williamson, Oliver, 1983, "The use of hostages as a credible commitment in bilateral exchange," *American Economic Review*, 73(4):519-540.

Williamson, Oliver, 1985, "The Economic Institutions of Capitalism," The Free Press, New York.