MECS 475: The Economics of Organizations

Michael Powell

Last Updated: April 20, 2014

Abstract

These lecture notes were written for a second-year PhD course in Organizational Economics. They are a work in progress and may therefore contain errors or misunderstandings. Any comments or suggestions would be greatly appreciated.

Overview and Reading List

The course will meet ten times. The meeting dates, topics, and papers that will be covered in each lecture are listed below.

Lecture 1 April 2nd: Introduction to Organizational Economics; Incentives

Risk-Incentives Trade-Off
Holmstrom, B. (1979), "Moral Hazard and Observability," *The Bell Journal of Economics*, Vol. 1, No. 1, pp. 74-91

Limited Liability
Jewitt, I., Kadan, O., and Swinkels, J. (2008), "Moral Hazard with Bounded Payments," *Journal of Economic Theory*, Vol. 143, No. 1, pp. 59-82

Multiple Tasks and Misaligned Performance Measures
Holmstrom, B., and Milgrom, P. (1991) "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization*, Vol. 7, Special Issue, pp. 24-52

Lecture 2 April 9th: Incentives in Organizations

Contracts with Externalities
Segal, I. (1999), "Contracting with Externalities," *The Quarterly Journal of Economics*, Vol. 114, No. 2, pp. 337-388

Career Concerns
Holmstrom, B. (1999), "Managerial Incentive Problems–A Dynamic Perspective," *The Review of Economic Studies*, Vol. 66, No. 1, pp. 169-182

Gibbons, R. and Murphy, K. (1992), "Optimal Incentive Contracts in the Presence of Career Concerns: Theory and Evidence," *The Journal of Political Economy*, Vol. 100, No. 3, pp. 468-505

Relational Incentive Contracts
Malcomson, J. (2013), "Relational Incentive Contracts," in *The Handbook of Organizational Economics*, R. Gibbons and J. Roberts (eds.), Princeton, NJ: Princeton University Press

Lecture 3 April 16th: Decision Making in Organizations - Niko

Influence Activities

Milgrom, P. and Roberts, J. (1988), "An Economic Approach to Influence Activities in Organizations," *The American Journal of Sociology*, Vol. 94, pp. S154-S179
Delegation
Alonso, R. and Matouschek, N. (2008), "Optimal Delegation," *The Review of Economic Studies*, Vol. 75, No. 1, pp. 259-293

Authority
Aghion, P. and Tirole, J. (1997), "Formal and Real Authority in Organizations," *The Journal of Political Economy*, Vol. 105, No. 1, pp. 1-29

Dessein, W. (2002), "Authority and Communication in Organizations," *The Review of Economic Studies*, Vol. 69, pp. 811-838

Centralization vs. Decentralization
Alonso, R., Dessein, W., and Matouschek, N. (2008), "When Does Coordination Require Centralizations?" *The American Economic Review*, Vol. 98, No. 1, pp. 145-179

Lecture 4 April 23rd: Boundaries of the Firm

Property Rights
Hart, O. (1995), *Firms, Contracts, and Financial Structure.* New York: Oxford University Press (Chapter 2)

Incentive Systems
Holmstrom, B. and Milgrom, P. (1994), "The Firm as an Incentive System," *The American Economic Review*, Vol. 84, No. 4, pp. 972-991

Adaptation
Tadelis, S. and Williamson, O. (2013), "Transaction Cost Economics," in *The Handbook of Organizational Economics, R. Gibbons and J. Roberts (eds.), Princeton, NJ: Princeton University Press*

Influence Costs
Powell, M. (2013), "An Influence-Cost Model of Organizational Practices and Firm Boundaries," *Working Paper*

Lecture 5 April 30th: Boundaries of the Firm

Classic Evidence
Monteverde, K. and Teece, D. (1982), "Supplier Switching Costs and Vertical Integration in the Automobile Industry," *The Bell Journal of Economics*, Vol. 13, No. 1, pp. 206-213

Recent Evidence

Baker, G. and Hubbard, T. (2003), "Make versus Buy in Trucking: Asset Ownership, Job Design, and Information," *The American Economic Review*, Vol. 93, No. 3, pp. 551-572

Forbes, S. and Lederman, M. (2009), "Adaptation and Vertical Integration in the Airline Industry," *The American Economic Review*, Vol. 99, No. 5, pp. 1831-1849

Foundations
Maskin, E. and Tirole, J. (1999), "Unforeseen Contingencies and Incomplete Contracts," *The Review of Economic Studies*, Vol. 66, No. 1, pp. 83-114

Aghion, P., Fudenberg, D., Holden, R., Kunimoto, T., and Tercieux, O. (2012), "Subgame-Perfect Implementation Under Information Perturbations," *The Quarterly Journal of Economics*, Vol. 127, No. 4, pp. 1843-1884

Fehr, E., Powell, M., and Wilkening, T. (2013), "Handing Out Guns at a Knife Fight: Behavioral Limitations of Subgame-Perfect Implementation," *Working Paper*

Lecture 6 May 7th: Careers in Organizations - Jin

Internal Labor Markets
Baker, G., Gibbs, M., and Holmstrom, B. (1994), "The Internal Economics of the Firm: Evidence from Personnel Data," *The Quarterly Journal of Economics*, Vol. 109, No. 4, pp. 881-919

Gibbons, R. and Waldman, M. (1999), "A Theory of Wage and Promotion Dynamics inside Firms," *The Quarterly Journal of Economics*, Vol. 114, No. 4, pp 1321-1358

Lazear, E. (1979), "Why is there Mandatory Retirement?" *The Journal of Political Economy*, Vol. 87, No. 6, pp. 1261-1284

Ke, R., Li, J., and Powell, M. (2014), "Managing Careers in Organizations," *Working Paper*

Lecture 7 May 14th: Relational Incentive Contracts

Imperfect Public Monitoring
Levin, J. (2003), "Relational Incentive Contracts," *The American Economic Review*, Vol. 93, No. 3, pp. 835-857

Limited Transfers
Fong Y. and Li, J. (2013), "Relational Contracts, Efficiency Wages, and Employment Dynamics," *Working Paper*

Li, J. and Matouschek, N. (2013), "Managing Conflicts in Relational Contracts," *The American Economic Review*, Vol. 103, No. 6, pp. 2328-2351

Lecture 8 May 21st: Relational Incentive Contracts - Dan

Subjective Performance Measures
Fuchs, W. (2007), "Contracting with Repeated Moral Hazard and Private Evaluations," *The American Economic Review*, Vol. 97, No. 4, pp. 1432-1448

Persistent Private Information
Halac, M. (2012), "Relational Contracts and the Value of Relationships," *The American Economic Review*, Vol. 102, No. 2, pp. 750-759

Multiple Agents
Board, S. (2011), "Relational Contracts and the Value of Loyalty," *The American Economic Review*, Vol. 101, No. 7, pp.3349-3367

Imperfect Private Monitoring
Andrews, I. and Barron, D. (2013), "The Allocation of Future Business," *Working Paper*

Barron, D. and Powell, M. (2014), "Policy Commitments in Relational Contracts," *Working Paper*

Lecture 9 May 28th: Persistent Performance Differences

Overview
Syverson, C. (2011), "What Determines Productivity?" *Journal of Economic Literature*, Vol. 49, No. 2, pp. 326-365

Henderson, R., and Gibbons, R. (2013), "What Do Managers Do?" in *The Handbook of Organizational Economics*, R. Gibbons and J. Roberts (eds.), Princeton, NJ: Princeton University Press

Misallocation
Hsieh, C. and Klenow, P. (2009), "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, Vol. 124, No. 4, pp. 1403-1448

Management Practices
Bloom, N. and Van Reenen, J. (2007), "Measuring and Explaining Management Practices Across Firms and Countries," *The Quarterly Journal of Economics*, Vol. 122, No. 4, pp. 1351-1408

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013), "Does Management Matter? Evidence from India," *The Quarterly Journal of Economics*

Theories

Chassang, S. (2010), "Building Routines: Learning, Cooperation, and the Dynamics of Incomplete Relational Contracts," *The American Economic Review*, Vol. 11, pp. 448-465

Li, J., Matouschek, N., and Powell, M. (2014), "The Burden of Past Promises," *Working Paper*

Lecture 10 June 4th: Organizations in Market Equilibrium

Information
Gibbons, R., Holden, R., and Powell, M. (2012), "Organization and Information: Firms' Governance Choices in Rational-Expectations Equilibrium," *The Quarterly Journal of Economics*, Vol. 127, No. 4, pp. 1813-1841

Price Levels
Legros, P. and Newman, A. (2013), "A Price Theory of Vertical and Lateral Integration," *The Quarterly Journal of Economics*, Vol. 128, No. 2, pp. 725-770

Competitive Rents
Powell, M. (2013), "Productivity and Credibility in Industry Equilibrium," *Working Paper.*

1 Introduction (Updated: Apr 1 2014)

Neoclassical economics traditionally viewed a firm as a production set–a collection of feasible input and output vectors. Given market prices, the firm chooses a set of inputs to buy, turns them into outputs, and then sells those outputs on the market in order to maximize profits. This "black box" view of the firm captures many important aspects of what a firm does: a firm transforms inputs into outputs, it behaves optimally, and it responds to market prices. And for many of the issues that economists were focused on long ago (such as: what is a competitive equilibrium? do competitive equilibria exist? is there more than one? who gets what in a competitive equilibrium?), this was perhaps the ideal level of abstraction.

But this view is inadequate as a descriptive matter (what do managers do? why do firms often appear dysfunctional?), and it leads to the following result:

Representative Firm Theorem (Acemoglu, 2006) Let $\mathcal{F}$ be a countable set of firms, each with a convex production-possibilities set $Y^f \subset \mathbb{R}^N$. Let $p \in \mathbb{R}_+^N$ be the price vector in this economy and denote the profit-maximizing net supplies of firm $f \in \mathcal{F}$ by $\hat{Y}^f(p) \subset Y^f$. Then there exists a representative firm with production possibilities set $Y \subset \mathbb{R}^N$ and set of profit-maximizing net supplies $\hat{Y}(p)$ such that for any $p \in \mathbb{R}_+^N$, $\hat{y} \in \hat{Y}(p)$ if and only if $\hat{y}(p) = \sum_{f \in \mathcal{F}} \hat{y}^f$ for some $\hat{y}^f \in \hat{Y}^f(p)$ for each $f \in \mathcal{F}$.

That is, by abstracting from many of the interesting and complex things that happen within firms, we are also left with a simplistic perspective of the production side of the economy as a whole–namely, we can think of the entire production side as a single (price-taking) firm. This view therefore is also inadequate as a model of firm behavior for many of the questions economists are now pursuing (why do inefficient and efficient firms coexist? should we care about their coexistence? when two firms merge, should this be viewed as a bad thing?).

The purpose of this course is to move beyond the Neoclassical view of the firm and to provide you with a set of models that you can use as a first step when thinking about modern issues. In doing so, we will recognize the fact that organizations consist of many individuals who almost always have conflicting objectives, and we will see that this can lead production sets to be determined as an equilibrium object rather than as an exogenously specified set of technological constraints. In the first part of the course, we will think about how these incentive issues affect the set Y^f. That is, given what is technologically feasible, how do different sources of contracting frictions (limiting monetary transfers or transfers of control) affect what is actually feasible and what firms will actually do?

In the second part of the course, we will study theories of the boundary of the firm. We will revisit the representative-firm theorem and ask under what conditions is there a difference between treating two firms, say firm 1 and firm 2, separately or as a single firm. If we denote the characteristics of the environment as θ, and we look at the following object:

$$\Delta\left(\theta\right)=\max_{y\in Y^1+Y^2}\pi\left(y\right)-\left[\max_{y^1\in Y^1}\pi_1\left(y_1\right)+\max_{y^2\in Y^2}\pi_2\left(y_2\right)\right],$$

we will ask when it is the case that $\Delta\left(\theta\right)\geq 0$ or $\Delta\left(\theta\right)\leq 0$. The representative-firm theorem shows that under some conditions, $\Delta\left(\theta\right)=0$. Theories of firm boundaries based solely on technological factors necessarily run into what Oliver Williamson refers to as the "selective intervention puzzle"–why can't a single large firm do whatever a collection of two small firms could do and more (by taking into account whatever externalities these two small firms impose on each other)? That is, shouldn't it always be the case that $\Delta\left(\theta\right)\geq 0$? And theories of the firm based solely on the idea that "large organizations suffer from costs of bureaucracy" have to contend with the equally puzzling question–why can't two small firms contractually internalize whatever externalities they impose on each other and remain separate, thereby avoiding bureaucracy costs? That is, shouldn't it be the case that $\Delta\left(\theta\right)\leq 0$?

In the last part of the course, we will focus on the following widespread phenomenon. If we take any two firms i and j, we almost always see that $\pi_i^*>\pi_j^*$. Some firms are just more productive than others. This is true even within narrowly defined industries, and it is true not just at a point in time, but over time as well–the same firms that outperform their competitors today are also likely to outperform their competitors tomorrow. Understanding the underlying source of profitability is essentially the fundamental question of strategy, so we will spend some time on this question. Economists outside of strategy have also recently started to focus on the implications of these performance differences and have pointed to a number of mechanisms under which $\pi_i^* > \pi_j^*$ implies that $\pi_i^* + \pi_j^* < \max_{y \in Y^i + Y^j} \pi(y)$. That is, it may be the case that performance differences are indicative of misallocation of resources across different productive units within an economy, and there is some evidence that this may especially be the case in developing countries. The idea that resources may be misallocated in equilibrium has mouthwatering implications, since it suggests that it may be possible to improve living standards for people in a country simply by shifting around existing resources.

Because the literature has in no way settled on a "correct" model of the firm (for reasons that will become clear as the course progresses), much of our emphasis will be on understanding the individual elements that go into these models and the "art" of combining these elements together to create new insights. This will, I hope, provide you with an applied-theoretic tool kit that will be useful both for studying new phenomena within organizations as well as for studying issues in other fields. As such, the course will be primarily theoretical. But in the world of applied theory, a model is only as good as its empirical implications, so we will also spend a couple weeks confronting evidence both to see how our models stack up to the data and to get a sense for what features of reality our models do poorly at explaining.

2 Incentives in Organizations (Updated: Apr 1 2014)

In order to move away from the Neoclassical view of a firm as a single individual pursuing a single objective, different strands of the literature have proposed different approaches. The first is what is now known as "team theory" (going back to the 1972 work of Marschak and Radner). Team-theoretic models focus on issues that arise when all members of an organization have the same preferences–they typically impose constraints on information transmission between individuals and information processing by individuals and look at questions of task and attention allocation.

The alternative approach, which we will focus on in the majority of the course, asserts that different individuals within the organization have different preferences (that is, "People (i.e., individuals) have goals; collectivities of people do not." (Cyert and March, 1963: 30)) and explores the implications that these conflicts of interest have for firm behavior. In turn, this approach examines how limits to formal contracting restrict a firm's ability to resolve these conflicts of interest and how unresolved conflicts of interest determine how decisions are made. We will talk about several different sources of limits to formal contracts and the trade-offs they entail.

We will then think about how to motivate individuals in environments where formal contracts are either unavailable or they are so incomplete that they are of little use. Individuals can be motivated out of a desire to convince "the market" that they are intrinsically productive in the hopes that doing so will attract favorable outside offers in the future–that is, they are motivated by their own career concerns. Additionally, individuals may form long-term attachments with an organization. In such long-term relationships, goodwill can arise as an equilibrium phenomenon, and fear of shattering this goodwill can motivate individuals to perform well and to reward good performance.

2.1 Formal Incentive Contracts (Updated: Apr 1 2014)

We will look at several different sources of frictions that prevent individuals from writing contracts with each other that induce the same patterns of behavior they would choose if they were all acting as a single individual receiving all the payoffs. The first will be familiar from core microeconomics–individual actions chosen by an agent are not observed but determine the distribution of a verifiable performance measure. The agent is risk-averse, so writing a high-powered contract on that noisy performance measure subjects him to costly risk. As a result, there is a trade-off between incentive provision (and therefore the agent's effort choice) and risk costs. This is the famous **risk-incentives trade-off**.

The second contracting friction that might arise is that an agent is either liquidity-constrained or is subject to a limited-liability constraint. As a result, the principal is unable to extract all the surplus the agent generates. Offering the agent a higher-powered contract induces him to exert more effort and therefore increases the total size of the pie, but it also leaves the agent with a larger share of that pie. The principal then, in choosing a contract, chooses one that trades off the creation of surplus with her ability to extract that surplus. This is the **motivation-rent extraction trade-off.**

The third contracting friction that might arise is that the principal's objective simply cannot be written into a formal contract. Instead, the principal has to rely on imperfectly aligned performance measures. Increasing the strength of a formal contract that is based on imperfectly aligned performance measures may increase the agent's efforts toward the principal's objectives, but it may also induce the agent to exert costly effort towards objectives that either hurt the principal or at least do not help the principal. Since the principal ultimately has to compensate the agent for whatever effort costs he incurs in order to get him to sign a contract to begin with, even the latter proves costly for the principal. Failure to account for this is sometimes referred to as **the folly of rewarding A while hoping for B** (Kerr, 1975).

All three of these sources of contractual frictions lead to similar results–under the optimal contract, the agent chooses an action that is not jointly optimal from his and the principal's perspective. But in different applied settings, different assumptions regarding what is contractible and what is not are more or less plausible. As a result, it is useful to master at least elementary versions of models capturing these three sources of frictions, so that you are well-equipped to use them as building blocks.

In all three of these models, the effort level that the Principal would induce if there were no contractual frictions would solve:

$$\max_e pe - \frac{c}{2}e^2,$$

so that $e^{FB} = \frac{p}{c}$. All three of these models yield equilibrium effort levels $e^* < e^{FB}$.

2.1.1 Risk-Incentives Trade-off (Updated: Apr 1 2014)

The exposition of an economic model usually begins with a rough (but accurate and mostly complete) description of the players, their preferences, and what they do in the course of the game. It should also include a precise treatment of the timing, which includes spelling out who does what and when and on the basis of what information, and a description of the solution concept that will be used to derive predictions. Given the description of the economic environment, it is then useful to specify the program(s) that players are solving.

I will begin with a pretty general description of the standard principal-agent model, but I will shortly afterwards specialize the model quite a bit in order to focus on a single point–the risk-incentives trade-off.

Description There is a risk-neutral Principal (P) and a risk-averse Agent (A). The Agent chooses an effort level $e \in \mathbb{R}_+$ at a private cost of $c(e)$, with $c'', c' > 0$, and this effort level affects the distribution over outputs $y \in Y$, with y distributed according to cdf $F(\cdot| e)$. These outputs can be sold on the product market for price p. The Principal can write a contract $w \in W \subset \{w : Y \to \mathbb{R}\}$ that determines a transfer $w(y)$ that she is compelled to pay the Agent if output y is realized. The Agent has an outside option that provides utility $\bar{u}$ to the Agent and $\bar{\pi}$ to the Principal. If the outside option is not exercised, the Principal's and Agent's preferences are, respectively,

$$\begin{aligned}
\Pi(w,e) &= \int_{y\in Y}(py-w(y))\,dF(y|\,e) = E_y\left[py-w|\,e\right] \\
U(w,e) &= \int_{y\in Y}(u(w(y)-c(e)))\,dF(y|\,e) = E_y\left[u(w-c(e))|\,e\right].
\end{aligned}$$

Timing The timing of the game is:

1. P offers A a contract $w(y)$, which is commonly observed.

2. A accepts the contract $(d=1)$ or rejects it $(d=0)$ and receives $\bar{u}$ and the game ends. This decision is commonly observed.

3. If A accepts the contract, A chooses effort level e and incurs cost $c(e)$. e is only observed by A.

4. Output y is drawn from distribution with cdf $F(\cdot|\,e)$. y is commonly observed.

5. P pays A an amount $w(y)$. This payment is commonly observed.

A couple remarks are in order at this point. First, behind the scenes, there is the implicit assumption that there is a third-party contract enforcer (a judge or arbitrator) who can costlessly detect when agreements have been broken and costlessly exact harsh punishments on the offender. Second, it is not necessarily important that e is unobserved by the Principal–given that the Principal takes no actions after the contract has been offered, as long as the contract cannot be conditioned on effort, the outcome of the game will be the same whether or not the Principal observes e. Put differently, one way of viewing the underlying source of moral hazard problems is that contracts cannot be conditioned on relevant variables, not that the relevant variables are unobserved by the Principal. We will return to this issue when we discuss the Property Rights theory of the firm.

Equilibrium A **pure-strategy subgame-perfect equilibrium** is a contract $w^* \in W$, an acceptance decision $d^* : W \to \{0,1\}$, and an effort choice $e^* : W \times \{0,1\} \to \mathbb{R}_+$ such that given the contract w^*, the Agent optimally chooses d^* and e^*, and given d^* and e^*, the Principal optimally offers contract w^*. We will say that the optimal contract induces effort e^*.

The Program The principal offers a contract $w \in W$ and proposes an effort level e in order to solve

$$\max_{w \in W, e} \int_{y \in Y} (py - w(y))\, dF(y|\, e)$$

subject to two constraints. The first constraint is that the agent actually prefers to choose effort level e rather than any other effort level $\hat{e}$. This is the standard incentive-compatibility constraint:

$$e \in \operatorname*{argmax}_{\hat{e} \in \mathbb{R}_+} \int_{y \in Y} (u(w(y) - c(\hat{e})))\, dF(y|\, \hat{e}).$$

The second constraint is that given that the agent knows he will choose e if he accepts the contract, he prefers to accept the contract rather than to reject it and receive his outside utility. This is the standard individual-rationality or participation constraint:

$$\int_{y \in Y} (u(w(y) - c(e)))\, dF(y|\, e) \geq \bar{u}.$$

CARA-Normal Case In order to establish a straightforward version of the risk-incentives trade-off, we will make a number of simplifying assumptions.

Assumption 1. The Agent has CARA preferences over wealth and effort costs, which are quadratic:

$$u(w(y) - c(e)) = -\exp\left\{-r\left(w(y) - \frac{c}{2}e^2\right)\right\},$$

and his outside option yields utility $-\exp\{-r\bar{u}\}$.

Assumption 2. Effort shifts the mean of a normally distributed random variable. That is,

$y \sim N(e, \sigma^2)$.

Assumption 3. $W = \{w(y) = s + by\}$. That is, the contract space permits only affine contracts.

Discussion. In principle, there should be no exogenous restrictions on the function $w(y)$. Applications will often restrict attention to affine contracts: $w(y) = s + by$. In many environments, an optimal contract does not exist if the contracting space is sufficiently rich, and situations in which the agent chooses the first-best level of effort and the principal receives all the surplus can be arbitrarily approximated with a sequence of sufficiently perverse contracts (Mirrlees, 1974; Moroni and Swinkels, 2014). In contrast, the optimal affine contract often results in an effort choice that is lower than the first-best effort level, and the principal receives a lower payoff.

There are then at least three ways to view the exercise of solving for the optimal affine contract.

1. From an applied perspective, many performance-pay contracts in the world are affine in the relevant performance measure–franchisees pay a franchise fee and receive a constant fraction of the revenues their store generates, windshield installers receive a base wage and a constant piece rate, fruit pickers are paid per kilogram of fruit they pick. And so given that many practitioners seem to restrict attention to this class of contracts, why don't we just make sure they are doing what they do optimally? We shall brush aside global optimality on purely pragmatic grounds.

2. Many performance-pay contracts in the world are affine in the relevant performance measure. Our models are either too rich or not rich enough in a certain sense and therefore generate optimal contracts that are inconsistent with those we see in the world. Maybe the aspects that, in the world, lead practitioners to use affine contracts are orthogonal to the considerations we are focusing on, so that by restricting attention to the optimal affine contract, we can still say something about how real-world contracts ought to vary with changes in the underlying environment. This view presumes a more positive (as opposed to normative) role for the modeler and hopes that the theoretical equivalent of the omitted variables bias is not too severe.

3. Who cares about second-best when first-best can be attained? If our models are pushing us toward complicated, non-linear contracts, then maybe our models are wrong. Instead, we should be focusing on writing down models that generate affine contracts as the optimal contract, and therefore we should think harder about what gives rise to them. (And indeed, steps have been made in this direction–see Holmstrom and Milgrom (1987) and, more recently, Carroll (2013)) This perspective will come back later in the course when we discuss the Property Rights Theory of firm boundaries.

Given the assumptions, for any contract $w(y) = s + by$, the income stream the agent receives is normally distributed with mean $s + be$ and variance $b^2\sigma^2$. His expected utility over monetary compensation is therefore a moment-generating function for a normally distributed random variable, (recall that if $X \sim N(\mu, \sigma^2)$, then $E[\exp\{tX\}] = \exp\left\{\mu t + \frac{1}{2}\sigma^2 t^2\right\}$), so his preferences can be written as

$$E\left[-\exp\left\{-r\left(w\left(y\right)-c\left(e\right)\right)\right\}\right] = -\exp\left\{-r\left(s+be\right)+\frac{r^2}{2}b^2\sigma^2+r\frac{c}{2}e^2\right\}.$$

We can take a monotonic transformation of his utility function $(-\frac{1}{r}\log(-x))$ and represent his preferences as:

$$\begin{aligned} U(e,w) &= E\left[w\left(y\right)\right]-\frac{r}{2}Var\left(w\left(y\right)\right)-\frac{c}{2}e^2 \\ &= s+be-\frac{r}{2}b^2\sigma^2-\frac{c}{2}e^2. \end{aligned}$$

The Principal's program is then

$$\max_{s,b,e} pe-(s+be)$$

subject to incentive-compatibility

$$e \in \underset{\hat{e}}{\operatorname{argmax}}\, b\hat{e}-\frac{c}{2}\hat{e}^2$$

and individual-rationality

$$s+be-\frac{r}{2}b^2\sigma^2-\frac{c}{2}e^2 \geq \bar{u}.$$

Solving this problem is then relatively straightforward. Given an affine contract $s+be$, the agent will choose an effort level $e(b)$ that satisfies his first-order conditions

$$e(b) = \frac{b}{c},$$

and the Principal will choose the value s to ensure that the agent's individual-rationality constraint holds with equality (for if it did not hold with equality, the Principal could reduce s, making herself better off without affecting the Agent's incentive-compatibility constraint, while maintaining the Agent's individual-rationality constraint). That is,

$$s + be\left(b\right) = \frac{c}{2}e\left(b\right)^2 + \frac{r}{2}b^2\sigma^2 + \bar{u}.$$

In other words, the Principal has to ensure that the Agent's total expected monetary compensation, $s + be\left(b\right)$, fully compensates him for his effort costs, the risk costs he has to bear if he accepts this contract, and his opportunity cost. Indirectly, then, the Principal bears these costs when designing an optimal contract.

The Principal's remaining problem is to choose the incentive slope b to solve

$$\max_b pe\left(b\right) - \frac{c}{2}e\left(b\right)^2 - \frac{r}{2}b^2\sigma^2 - \bar{u}.$$

This is now an unconstrained problem with proper convexity assumptions, so the Principal's optimal choice of incentive slope solves her first-order conditions

$$\begin{aligned} 0 &= pe'\left(b^*\right) - ce^*\left(b^*\right)e'\left(b^*\right) - rb^*\sigma^2 \\ &= \frac{p}{c} - c\frac{b^*}{c}\frac{1}{c} - rb^*\sigma^2 \end{aligned}$$

and therefore

$$b^* = \frac{p}{1 + rc\sigma^2}.$$

Also, given b^* and the individual-rationality constraint, we can back out s^*.

$$s^* = \bar{u} + \frac{1}{2}\left(rc\sigma^2 - 1\right)\frac{\left(b^*\right)^2}{c}.$$

Depending on the parameters, it may be the case that $s^* < 0$. That is, the Agent would have to pay the Principal if he accepts the job and does not produce anything.

In this setting, if the Principal could contract directly on effort, she would choose a contract that ensures that the Agent's individual-rationality constraint binds and therefore would solve

$$\max_e pe - \frac{c}{2}e^2,$$

so that

$$e^{FB} = \frac{p}{c}.$$

If the Principal wanted to implement this same level of effort using a contract on output, y, she would choose $b = p$ (since the Agent would choose $\frac{b}{c} = \frac{p}{c}$).

Why, in this setting, does the Principal not choose such a contract? Let us go back to the Principal's problem of choosing the incentive slope b.

$$\max_b pe(b) - \frac{c}{2}e(b)^2 - \frac{r}{2}b^2\sigma^2 - \bar{u}$$

Many fundamental points in models in the Organizational Economics literature can be seen as a comparison of first-order losses or gains against second-order gains or losses. Suppose the Principal chooses $b = p$, and consider a marginal reduction in b away from this value. The change in the Principal's profits would be

$$\begin{aligned} & \frac{d}{db}\left(pe(b) - \frac{c}{2}e(b)^2 - \frac{r}{2}b^2\sigma^2\right)\Big|_{b=p} \\ = & \underbrace{\frac{d}{db}\left(pe(b) - \frac{c}{2}e(b)^2\right)\Big|_{b=p}}_{=0} - rp\sigma^2 < 0 \end{aligned}$$

This first term is zero, because $b = p$ in fact maximizes $pe(b) - \frac{c}{2}e(b)^2$ (since it induces the first-best level of effort). The second term is strictly negative. That is, relative to the contract that induces first-best effort, a reduction in the slope of the incentive contract yields a first-order gain resulting from a decrease in the risk costs the Agent bears, while it yields a second-order loss in terms of profits resulting from moving away from the effort level that maximizes revenues minus effort costs. The optimal contract balances the incentive benefits of higher-powered incentives with these risk costs.

This trade-off seems first-order in some settings (e.g., insurance contracts in health care markets, some types of sales contracts in industries in which individual sales are infrequent, large, and unpredictable) and for certain types of output. There are many other environments where contracts offer less than first-best incentives, but the first-order reasons for this seem completely different, and we will turn to these environments shortly.

Before doing so, it is worth pointing out that many models in this course will involve trade-offs that determine the optimal way of organizing. In many of the settings these models examine, results that take the form of "X is organized according to Y because A is more risk-averse than B" often seem intuitively unappealing. For example, suppose a model of hierarchies generated the result that less risk-averse individuals should be at the top of an organization, and more risk-averse individuals should be at the bottom. This sounds somewhat sensible–maybe richer individuals are better able to diversify their wealth, and they can therefore behave as if they are less risk averse with respect to the income stream they derive from a particular organization. But it perhaps sounds less appealing as a general rule for who should be assigned to what role in an organization–a model that predicts that more knowledgeable or more experienced workers should be assigned higher positions seems more consistent with experience.

- Production set?

- Technological possibilities Y^f–\$ paid to worker as input, y as output:

$$C = c(e) = \frac{c}{2}e^2 \Rightarrow y(C) = e = \left(\frac{2C}{c}\right)^{1/2}$$

- Technological possibilities set:

$$Y^f = \{(y, -C) : y \leq y(C)\}$$

- Contract-augmented set: look at firm's objective:

$$p\frac{b}{c}-\frac{c}{2}\left(\frac{b}{c}\right)^2-\frac{r}{2}\sigma^2b^2=p\frac{b}{c}-\frac{1}{2}\frac{1+rc\sigma^2}{c}b^2$$

- To produce $y=\frac{b}{c}$, costs

$$C=\frac{1}{2}\frac{1+rc\sigma^2}{c}b^2$$

- Expressing $y=\tilde{y}(C)$, solve above for $b:$

$$\begin{aligned} b &= \left(\frac{2Cc}{1+rc\sigma^2}\right)^{1/2}, \\ \tilde{y}(C) &= \left(\frac{2C}{c}\right)^{1/2}\left(\frac{1}{1+rc\sigma^2}\right)^{1/2}=\left(\frac{1}{1+rc\sigma^2}\right)^{1/2}y(C) \end{aligned}$$

- Contract-augmented possibilities set:

$$\tilde{Y}^f=\{(y,-C):y\leq\tilde{y}(C)\}$$

- Contractual frictions: $\tilde{Y}^f\subset Y^f$

2.1.2 Limited Liability (Updated: Apr 1 2014)

We saw in the previous model that the optimal contract sometimes involved upfront payments from the Agent to the Principal. To the extent that the Agent is unable to afford such payments (or legal restrictions prohibit such payments), the Principal will not be able to extract all the surplus that the Agent creates. Further, the amount that she can extract may depend on the total surplus created. As a result, the Principal may offer a contract that induces effort below the first-best.

Description Again, there is a risk-neutral Principal (P). There is also a **risk-neutral** Agent (A). The Agent chooses an effort level $e\in\mathbb{R}_+$ at a private cost of $c(e)$, with $c'',c'>0$, and this effort level affects the distribution over outputs $y \in Y$, with y distributed according to cdf $F(\cdot|\, e)$. These outputs can be sold on the product market for price p. The Principal can write a contract $w \in W \subset \{w : Y \to \mathbb{R}, w(y) \geq \underline{w}$ for all $y\}$ that determines a transfer $w(y)$ that she is compelled to pay the Agent if output y is realized. The Agent has an outside option that provides utility $\bar{u}$ to the Agent and $\bar{\pi}$ to the Principal. If the outside option is not exercised, the Principal's and Agent's preferences are, respectively,

$$\begin{aligned} \Pi(w, e) &= \int_{y \in Y} (py - w(y))\, dF(y|\, e) = E_y[py - w|\, e] \\ U(w, e) &= \int_{y \in Y} (w(y) - c(e))\, dF(y|\, e) = E_y[w - c(e)|\, e]. \end{aligned}$$

There are two differences between this model and the model in the previous subsection. The first is that the Agent is risk-neutral (so that absent any other changes, the equilibrium contract would induce first-best effort). The second is that the wage payment from the Principal to the Agent has to exceed, for each realization of ouptut, a value $\underline{w}$. Depending on the setting, this constraint is described as a liquidity constraint or a limited-liability constraint. In repeated settings, it is more naturally thought of as the latter–due to legal restrictions, the Agent cannot be legally compelled to make a transfer (larger than $-\underline{w}$) to the Principal. In static settings, either interpretation may be sensible depending on the particular application–if the Agent is a fruit picker, for instance, he may not have much liquid wealth that he can use to pay the Principal.

Timing The timing of the game is exactly the same as before.

1. P offers A a contract $w(y)$, which is commonly observed.

2. A accepts the contract $(d = 1)$ or rejects it $(d = 0)$ and receives $\bar{u}$ and the game ends. This decision is commonly observed.

3. If A accepts the contract, A chooses effort level e and incurs cost $c(e)$. e is only observed by A.

4. Output y is drawn from distribution with pdf $f(\cdot|\, e)$. y is commonly observed.

5. P pays A an amount $w(y)$. This payment is commonly observed.

Equilibrium The solution concept is the same as before. A **pure-strategy subgame-perfect equilibrium** is a contract $w^* \in W$, an acceptance decision $d^* : W \to \{0,1\}$, and an effort choice $e^* : W \times \{0,1\} \to \mathbb{R}_+$ such that given the contract w^*, the Agent optimally chooses d^* and e^*, and given d^* and e^*, the Principal optimally offers contract w^*. We will say that the optimal contract induces effort e^*.

The Program The principal offers a contract $w \in W$ and proposes an effort level e in order to solve

$$\max_{w \in W, e} \int_{y \in Y} (py - w(y))\, dF(y|\, e)$$

subject to three constraints: the incentive-compatibility constraint

$$e \in \operatorname*{argmax}_{\hat{e} \in \mathbb{R}_+} \int_{y \in Y} (u(w(y) - c(\hat{e})))\, dF(y|\, \hat{e}),$$

the individual-rationality constraint

$$\int_{y \in Y} (u(w(y) - c(e)))\, dF(y|\, e) \geq \bar{u},$$

and the limited-liability constraint

$$w(y) \geq \underline{w} \text{ for all } y.$$

Binary-Output Case Jewitt, Kadan, and Swinkels (2008) solves for the optimal contract in the general environment above (and even allows for agent risk aversion). Here, I will instead focus on an elementary case that highlights some of the main trade-offs.

Assumption 1. Output is $y \in \{0,1\}$, and given effort e, its distribution satisfies $\Pr[y = 1|\, e] =$

e.

Assumption 2. The agent's costs are quadratic: $c(e) = \frac{c}{2}e^2$. c is sufficiently high so that any equilibrium effort choice will be less than 1.

Finally, we can restrict attention to affine contracts

$$\begin{aligned} W &= \{w(y) = (1-y)w_0 + yw_1, w_0, w_1 \geq \underline{w}\} \\ &= \{w(y) = s + by, s \geq \underline{w}, b \geq 0\} \end{aligned}$$

When output is binary, this restriction is without loss of generality. For the purposes of this specific model, it can be useful to change the variables in the problem. Let $b = w_1 - w_0$ denote the wage differential if output is positive and $w = (1-e)w_0 + ew_1$ denote the average wage the Agent receives if he chooses effort e. The reason for this change of variables, we will see, is that b is the relevant part of compensation that affects the Agent's effort choice, and w is the relevant part of compensation that affects the Agent's participation decision. The Principal's problem can be expressed as

$$\max_{e,b,w} pe - w$$

subject to incentive-compatibility

$$e \in \operatorname*{argmax}_{\hat{e}} \hat{e}b - \frac{c}{2}\hat{e}^2,$$

individual-rationality

$$w - \frac{c}{2}e^2 \geq \bar{u}$$

and limited-liability

$$w \geq eb + \underline{w}$$

and $b \geq 0$. This latter constraint will always be satisfied as long as the equilibrium involves positive effort.

Motivation-Rent Extraction Trade-off Before solving the model in detail, I will first perform a partial characterization of the solution that highlights the motivation-rent extraction trade-off. First, note that first-best effort solves

$$\max_{\hat{e}} p\hat{e} - \frac{c}{2}\hat{e}^2,$$

so that $e^{FB} = \frac{p}{c}$. In order to implement first-best effort, the Principal can offer a contract with b^{FB} that solves $e^*\left(b^{FB}\right) = b^{FB}$, or $b^{FB} = p$. We will now perform an exercise similar to what we did above to establish the risk-incentives trade-off, showing that moving away from first-best incentives leads to second-order losses and first-order gains. In order to do so, we can think about the Principal's problem of choosing w and e while taking b as exogenous.

To do so, define

$$\tilde{\pi}(b) = \max_{e,w} pe - w$$

subject to incentive-compatibility

$$e \in \operatorname*{argmax}_{\hat{e}} \hat{e}b - \frac{c}{2}\hat{e}^2,$$

individual-rationality

$$w - \frac{c}{2}e^2 \geq \bar{u}$$

and limited-liability

$$w \geq eb + \underline{w}.$$

The first observation to make is that given b, e is uniquely determined by $e^*(b) = \frac{b}{c}$. The Principal's problem then becomes

$$\tilde{\pi}(b) = \max_{w} p\frac{b}{c} - w$$

subject to individual-rationality

$$
\begin{aligned}
w &\geq \frac{1}{2}\frac{b^2}{c} + \bar{u} \\
w &\geq \frac{b^2}{c} + \underline{w}.
\end{aligned}
$$

The Lagrangian for this problem is

$$
\begin{aligned}
\mathcal{L}(w, \lambda, \mu; b) &= p\frac{b}{c} - w + \lambda\left(w - \frac{1}{2}\frac{b^2}{c} - \bar{u}\right) + \mu\left(w - \frac{b^2}{c} - \underline{w}\right) \\
&= p\frac{b}{c} - \frac{1}{2}\frac{b^2}{c} - w + \frac{1}{2}\frac{b^2}{c}(1 - \lambda - 2\mu) + \lambda(w - \bar{u}) + \mu(w - \underline{w}).
\end{aligned}
$$

Taking first-order conditions with respect to w shows that at the optimum, it has to be the case that $\lambda^* + \mu^* = 1$. We can then use the envelope theorem:

$$
\left.\frac{d\tilde{\pi}}{db}\right|_{b=b^{FB}} = \left.\frac{\partial \mathcal{L}}{\partial b}\right|_{b=b^{FB}} = \underbrace{\frac{\partial}{\partial b}\left.\left(p\frac{b}{c} - \frac{1}{2}\frac{b^2}{c}\right)\right|_{b=b^{FB}}}_{=0} + \underbrace{\frac{b}{c}(1 - \lambda^* - \mu^*)}_{=0} - \frac{b}{c}\mu^* < 0,
$$

where the first expression is 0, because b^{FB} maximizes social surplus, and the second expression is 0 from the first-order conditions for the Principal's problem. In words, reducing b away from b^{FB} leads to a marginal reduction in the Agent's effort, but the total welfare loss associated with this reduction is second-order. The Principal effectively cares about total welfare at the margin, because she has to compensate the Agent for a marginal increase in his effort either through providing him more utility so that he participates in the contract (if the individual-rationality constraint binds) or through providing him more utility so that despite the existence of a limited-liability constraint, he is still willing to incur the desired effort level (if the limited-liability constraint binds). If at $b = b^{FB}$, the limited-liability constraint is optimally binding (i.e., $\mu^* > 0$), then reducing b yields a first-order gain. Put differently, the Principal has to provide the Agent with rents at the margin and will therefore optimally offer the Agent a contract that does not maximize total surplus.

Equilibrium Contracts and Effort While partial characterizations like the one we just carried out are useful for highlighting forces at play in an economic model (and can aid model design by providing the researcher with some guidance about which elements are important for generating those forces), in order to make use of these forces in a more complicated environment, it is useful to fully understand the solution to the workhorse model.

Given b, the Agent's incentive-compatibility constraint can be replaced with the Agent's first-order conditions

$$e(b) = \frac{b}{c}.$$

Let λ denote the Lagrange multiplier for the individual-rationality constraint and μ denote the Lagrange multiplier for the limited-liability constraint. The Principal's problem can be expressed as a Lagrangian

$$\mathcal{L}(b, w, \lambda, \mu) = pe(b) - w + \lambda\left(w - \frac{c}{2}e(b)^2 - \bar{u}\right) + \mu\left(w - e(b)\,b - \underline{w}\right).$$

The Kuhn-Tucker conditions are:

$$\begin{aligned}
(b) \quad &: \quad pe'(b^*) - \lambda^* ce(b^*)\,e'(b^*) - \mu^*\left(e(b^*) + b^* e'(b^*)\right) = 0 \\
(w) \quad &: \quad -1 + \lambda^* + \mu^* = 0
\end{aligned}$$

There are four cases to consider, only two of which will end up being important. At the optimum:

1. Neither the limited-liability constraint nor the individual-rationality constraint bind, so that $\lambda^* = 0, \mu^* = 0$

2. The limited-liability constraint does not bind, but the invidual-rationality constraint does, so that $\lambda^* > 0, \mu^* = 0$

3. The limited-liability constraint binds, but the individual-rationality constraint does

not, so that $\lambda^* = 0, \mu^* > 0$

4. Both constraints bind, so that $\lambda^* > 0, \mu^* > 0$.

The optimum can never be in the first case, for otherwise, the Principal could always reduce w, holding b fixed, and increase her profits without affecting the Agent's incentive-compatibility constraint. If the optimum is in the second case, the problem is rather trivial– the KT condition with respect to (w) gives $\lambda^* = 1$, and the KT condition with respect to (b) gives $e(b^*) = \frac{p}{c}$, which is the first-best effort level. This should not be surprising, given that the Agent is risk-neutral.

Now, let us consider the third case. When only the limited-liability constraint binds, the KT condition with respect to (w) gives us that $\mu^* = 1$, and the KT condition with respect to (b) gives us that

$$e(b^*) = pe'(b^*) - b^* e'(b^*) = \frac{1}{2}\frac{p}{c} = \frac{1}{2}e^{FB}$$

and $b^* = \frac{p}{2}$. Finally, let us consider the fourth case. When both constraints bind, we know that $w_0 = \underline{w}$, and we can directly use the individual-rationality constraint to solve for the equilibrium effort level:

$$e(b^*) = \frac{b^*}{c} = \sqrt{\frac{2(\bar{u} - \underline{w})}{c}},$$

and $b^* = \sqrt{2c(\bar{u} - \underline{w})}$

The more delicate part of the analysis involves figuring out which constraints bind at the optimum.

If the IR is not binding, then $e = \frac{1}{2}\frac{p}{c}$, $b = \frac{p}{2}$, and therefore, the IR is indeed not binding as long as

$$\begin{aligned} \bar{u} &< e(b^*) b^* + \underline{w} - \frac{c}{2} e(b^*)^2 \\ \bar{u} - \underline{w} &< \frac{1}{8}\frac{p^2}{c} \end{aligned}$$

If the LL is not binding, then $e = \frac{p}{c}$, $b = p$. We can solve the IR constraint for $w_0 = \bar{u} - \frac{p^2}{2c}$, and therefore the LL is indeed not binding as long as

$$\bar{u} - \underline{w} > \frac{1}{2}\frac{p^2}{c}.$$

So there are three cases:

$$\begin{aligned}
\bar{u} - \underline{w} &> \frac{1}{2}\frac{p^2}{c} \Rightarrow \text{LL not binding, IR binding} \\
\frac{1}{2}\frac{p^2}{c} &\geq \bar{u} - \underline{w} \geq \frac{1}{8}\frac{p^2}{c} \Rightarrow \text{LL binding, IR binding} \\
\bar{u} - \underline{w} &< \frac{1}{8}\frac{p^2}{c} \Rightarrow \text{LL binding, IR not binding}
\end{aligned}$$

It is worth noting that the solution is continuous across the boundaries.

Unfortunately, even in this elemental model of liquidity constraints in agency problems, the optimal contract is somewhat messy. There are several cases to consider, each of which yields a different functional form for the optimal contract, and which case prevails and depends on $\bar{u} - \underline{w}$, which is not exactly something that is easily measured in the data.

I think liquidity constraints are extremely important and are probably one of the main reasons for why many jobs do not involve first-best incentives. The Vickrey-Clarke-Groves logic that first-best outcomes can be obtained if the firm transfers the entire profit stream to each of its members in exchange for a large up-front payment seems simultaneously compelling, trivial, and obviously impracticable. In for-profit firms, in order to make it worthwhile to transfer a large enough share of the profit stream to an individual worker to significantly affect his incentives, the firm would require a large up-front transfer that most workers cannot afford to pay. It is therefore not surprising that we do not see most workers' compensation tied directly to the firm's overall profits in a meaningful way. One implication of this is that firms have to find alternative instruments to use as performance measures, which we will turn to next. In principle, models in which firms do not motivate their workers by writing contracts directly on profits should include assumptions in which the firm optimally chooses not to write contracts directly on profits, but they almost never do.

Exercise 1. *Holmstrom (1979) shows that in the risk-incentives model in the previous subsection, if there is a costless additional performance measure m that is informative about e, then an optimal formal contract should always put some weight on m unless y is a sufficient statistic for y and m. A rough intuition for this result is the following. Start from a contract that puts zero weight on both y and m. Locally, the agent is risk-neutral with respect to contracts that put a tiny amount of weight on y, m, or both. If they are both informative about e, then neither is a perfect substitute for the other, so a move in the direction of a contract that puts weight on both will yield a higher increase in profits than one that puts weight only on one. This is known as Holmstrom's "informativeness principle" and suggests that optimal contracts should always be extremely sensitive to the details of the environment the contract is written in. Suppose instead that the agent is risk-neutral but liquidity-constrained, and suppose there is a performance measure $m \in \{0,1\}$ such that $\Pr[m=1|\,e]=e$ and conditional on e, m and y are independent. Suppose contracts of the form $w(y,m)=s+b_y y+b_m m+b_{ym} ym$ can be written but must satisfy $w(y,m)\geq \bar{w}$ for each realization of (y,m). Is it again always the case that $b_m \neq 0$ and/or $b_{ym}\neq 0$?*

2.1.3 Multiple Tasks and Misaligned Performance Measures (Updated: Apr 2 2014)

In the previous two models, what the Principal cared about was output, and output, though a noisy measure of effort, was perfectly measurable. This assumption seems sensible when we think about overall firm profits (ignoring basically everything that accountants think about every day), but as we alluded to in the discussion above, overall firm profits are generally too blunt of an instrument to use to motivate individual workers within the firm if they are liquidity-constrained. As a result, firms often try to motivate workers using more specific performance measures, but while these performance measures are informative about what actions workers are taking, they may be less useful as a description of how the workers' actions affect the objectives the firm cares about. And paying workers for what is measured may not get them to take actions that the firm cares about. This observation underpins the title of the famous 1975 paper by Steve Kerr called "On the Folly of Rewarding A, while Hoping for B."

Description Again, there is a risk-neutral Principal (P) and a risk-neutral Agent (A). The Agent chooses an effort vector $e = (e_1, e_2) \in \mathbb{R}^2_+$ at a private cost of $\frac{c}{2}\left(e_1^2 + e_2^2\right)$. This effort vector affects the distribution of output $y \in Y = \{0, 1\}$ and a performance measure $m \in M = \{0, 1\}$ as follows:

$$
\begin{aligned}
\Pr\left[y = 1 \mid e\right] &= f_1 e_1 + f_2 e_2 \\
\Pr\left[m = 1 \mid e\right] &= g_1 e_1 + g_2 e_2,
\end{aligned}
$$

where it may be the case that $f = (f_1, f_2) \neq (g_1, g_2) = g$. Assume that $f_1^2 + f_2^2 = g_1^2 + g_2^2 = 1$ (i.e., the norms of the f and g vectors are unity). The output can be sold on the product market for price p. The Principal can write a contract $w \in W \subset \{w : M \to \mathbb{R}\}$ that determines a transfer $w(m)$ that she is compelled to pay the Agent if performance measure m is realized. Since the performance measure is binary, contracts take the form $w = s + bm$. The Agent has an outside option that provides utility $\bar{u}$ to the Agent and $\bar{\pi}$ to the Principal. If the outside option is not exercised, the Principal's and Agent's preferences are, respectively,

$$
\begin{aligned}
\Pi(w, e) &= f_1 e_1 + f_2 e_2 - E\left[w(m) \mid e\right] \\
U(w, e) &= s + b\left(g_1 e_1 + g_2 e_2\right) - \frac{c}{2}\left(e_1^2 + e_2^2\right).
\end{aligned}
$$

Timing The timing of the game is exactly the same as before.

1. P offers A a contract $w(m)$, which is commonly observed.

2. A accepts the contract ($d = 1$) or rejects it ($d = 0$) and receives $\bar{u}$ and the game ends. This decision is commonly observed.

3. If A accepts the contract, A chooses effort vector e and incurs cost $c(e)$. e is only observed by A.

4. Performance measure m is drawn from distribution with pdf $f(\cdot \mid e)$ and output y is

drawn from distribution with pdf $g(\cdot | e)$. m is commonly observed.

5. P pays A an amount $w(m)$. This payment is commonly observed.

Equilibrium The solution concept is the same as before. A **pure-strategy subgame-perfect equilibrium** is a contract $w^* \in W$, an acceptance decision $d^* : W \to \{0, 1\}$, and an effort choice $e^* : W \times \{0, 1\} \to \mathbb{R}_+^2$ such that given the contract w^*, the Agent optimally chooses d^* and e^*, and given d^* and e^*, the Principal optimally offers contract w^*. We will say that the optimal contract induces effort e^*.

The Program The principal offers a contract $w = s + bm$ and proposes an effort level e in order to solve

$$\max_{s,b,e} p(f_1 e_1 + f_2 e_2) - (s + b(g_1 e_1 + g_2 e_2))$$

subject to the incentive-compatibility constraint

$$e \in \operatorname*{argmax}_{\hat{e} \in \mathbb{R}_+} s + b(g_1 \hat{e}_1 + g_2 \hat{e}_2) - \frac{c}{2}\left(\hat{e}_1^2 + \hat{e}_2^2\right)$$

and the individual-rationality constraint

$$s + b(g_1 e_1 + g_2 e_2) - \frac{c}{2}\left(e_1^2 + e_2^2\right) \geq \bar{u}.$$

Equilibrium Contracts and Effort Given a contract $s + bm$, the Agent will choose efforts

$$\begin{aligned} e_1^*(b) &= \frac{b}{c} g_1 \\ e_2^*(b) &= \frac{b}{c} g_2. \end{aligned}$$

The Principal will choose s so that the individual-rationality constraint holds with equality

$$s + b\left(g_1 e_1^*(b) + g_2 e_2^*(b)\right) = \bar{u} + \frac{c}{2}\left(e_1^*(b)^2 + e_2^*(b)^2\right).$$

Since contracts send the Agent off in the "wrong direction" relative to what maximizes total surplus, providing the Agent with higher-powered incentives by increasing b sends the agent farther of in the wrong direction. This is costly for the Principal, because in order to get the Agent to accept the contract, she has to compensate him for his effort costs, even if they are in the wrong direction.

The Principal's unconstrained problem is therefore

$$\max_b p\left(f_1 e_1^*(b) + f_2 e_2^*(b)\right) - \frac{c}{2}\left(e_1^*(b)^2 + e_2^*(b)^2\right) - \bar{u}.$$

Taking first-order conditions,

$$pf_1 \frac{\partial e_1^*}{\partial b} + pf_2 \frac{\partial e_2^*}{\partial b} = ce_1^*(b^*)\frac{\partial e_1^*}{\partial b} + ce_2^*(b^*)\frac{\partial e_2^*}{\partial b},$$

or

$$\begin{aligned} pf_1 g_1 + pf_2 g_2 &= b^* g_1 g_1 + b^* g_2 g_2 \\ b^* &= p\frac{f_1 g_1 + f_2 g_2}{g_1^2 + g_2^2} = p\frac{f \cdot g}{g \cdot g} = p\frac{||f||}{||g||}\cos\theta = p\cos\theta, \end{aligned}$$

where $\cos\theta$ is the angle between the vectors f and g. That is, the optimal incentive slope depends on the relative magnitudes of the f and g vectors (which in this model were assumed to be the same, but in a richer model this need not be the case) as well as how well-aligned they are. If m is a perfect measure of what the firm cares about, then g is a linear transformation of f and therefore the angle between f and g would be zero, so that $\cos\theta = 1$. If m is completely uninformative about what the firm cares about, then f and g

are orthogonal, and therefore $\cos\theta = 0$. As a result, this model is often referred to as the **cosine of theta model**.

Another way to view this model is as follows. Since formal contracts allow for unrestricted lump-sum transfers between the Principal and the Agent, the Principal would optimally like efforts to be chosen in such a way that they maximize total surplus:

$$\max_{e} p\left(f_1 e_1 + f_2 e_2\right) - \frac{c}{2}\left(e_1^2 + e_2^2\right),$$

or $e_1^* = \frac{p}{c} f_1$ and $e_2^* = \frac{p}{c} f_2$. That is, the Principal would like to choose a vector of efforts that is collinear with the vector f :

$$\left(e_1^*, e_2^*\right) = \frac{p}{c} \cdot \left(f_1, f_2\right).$$

Since contracts can only depend on m and not directly on y, the Principal has only limited control over the actions that the Agent chooses. That is, given a contract specifying incentive slope b, the Agent chooses $e_1^*(b) = \frac{b}{c} g_1$ and $e_2^*(b) = \frac{b}{c} g_2$. Therefore, the Principal can only (indirectly) choose a vector of efforts that is collinear with the vector g:

$$\left(e_1^*(b), e_2^*(b)\right) = \frac{b}{c} \cdot \left(g_1, g_2\right).$$

The question is then: which such vector maximizes total surplus (which the Principal will extract with an ex-ante lump-sum transfer)? That is, which point along the $k \cdot (g_1, g_2)$ ray minimizes the mean-squared error distance to $\frac{p}{c} \cdot (f_1, f_2)$?

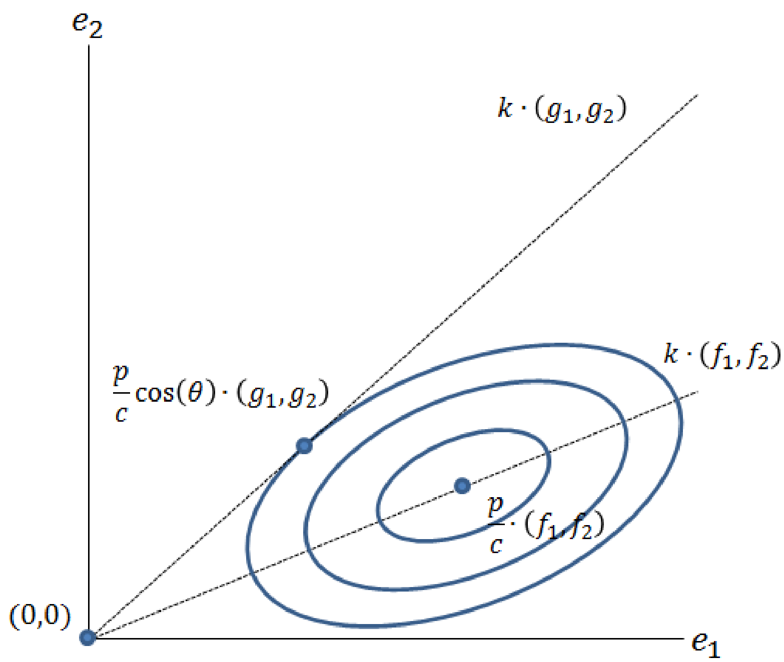


Figure 1: Cosine of theta.

This is a more explicit "incomplete contracts" model of motivation. That is, we are explicitly restricting the set of contracts that the Principal can offer the Agent in a way that directly determines a subset of the effort space that the Principal can induce the Agent to choose among. And it is founded not on the idea that certain measures (in particular, y) are unobservable, but rather that they cannot be contracted upon.

Exercise 2. *Suppose there are N tasks rather than 2 (i.e., $e = (e_1, \ldots, e_N)$ and $c(e) = \frac{1}{2}(e_1^2 + \cdots + e_N^2)$) and $M < N$ linearly independent performance measures rather than 1 (i.e., $m_j = g_{1j}a_1 + \cdots + g_{Nj}a_N$ for $j = 1, \ldots, M$). Show that the optimal incentive slope vector is equal to the regression coefficient that would be obtained if one ran the regression $f_i = \alpha + \beta_1 g_{i1} + \cdots + \beta_M g_{iM} + \varepsilon_i$.*

2.1.4 Contracts with Externalities (Updated: Apr 2 2014)

Before moving on to consider environments in which no formal contracts are available, we will briefly examine another source of contractual frictions that can arise and prevent parties from taking first-best actions. So far, we have considered what happens when certain states of nature or actions were impossible to contract upon or where there were legal or practical restrictions on the form of the contract. Here, we will consider limits on the number of parties that can be part of the same contract. We refer to these situations as "contracts with externalities," following Segal (1999). We will highlight, in the context of two separate models, some of the problems that can arise when there are multiple principals offering contracts to a single Agent.

In the first model, I show that when there are otherwise no contracting frictions, so that if the Principals could jointly offer a single contract to the Agent, they would be optimally choose a contract that induces first-best effort, there may be **coordination failures**. There are equilibria in which the Principals offer contracts that do not induce first-best effort, and there are equilibria in which they offer contracts that do induce first-best effort. In the second model, I show that when there are direct costs associated with higher-powered incentives (as is the case when the Agent is risk-averse or when Principals have to incur a setup cost to put in place higher-powered incentive schemes, as in Battigalli and Maggi (2002)). In this setting, if the Principals could jointly offer a single contract to the Agent, they would optimally choose a contract that induces an effort level e^C lower than the first-best effort level, because of a contracting costs-incentives trade-off (analogous to the risk-incentives trade-off). If they cannot jointly offer a single contract, there will be a unique equilibrium in which the Principals offer contracts that induce effort $e^* < e^C$.

Description of Coordination-Failure Version There is a risk-neutral Agent (A) and two risk-neutral Principals (P_1 and P_2). The Agent chooses an effort $e \in \{0, 1\}$ at cost ce. This effort determines outputs $y_1 = e$ and $y_2 = e$ that accrue to the Principals. These outputs can be sold on the product market for prices p_1 and p_2, respectively, and let $p \equiv p_1 + p_2$. Principals simultaneously offer contracts $w_1, w_2 \in W = \{w : \{0, 1\} \to \mathbb{R}\}$ to the Agent. Denote Principal i's contract offer by $w_i = s_i + b_i e$. The Agent has an outside option that yields utility $\bar{u}$ to the Agent and 0 to each Principal. If the outside option is not exercised, players' payoffs are:

$$\begin{aligned}\Pi_1(w_1, w_2, e) &= p_1 e - w_1 \\ \Pi_2(w_1, w_2, e) &= p_2 e - w_2 \\ U(w_1, w_2, e) &= w_1 + w_2 - ce.\end{aligned}$$

Timing The timing of the game is exactly the same as before.

1. P_1 and P_2 simutaneously offer contracts w_1, w_2 to A. Offers are commonly observed.
2. A accepts both contracts $(d = 1)$ or rejects both contracts $(d = 0)$ and receives $\bar{u}$ and the game ends. This decision is commonly observed.
3. If A accepts the contracts, A chooses effort $e \in \{0, 1\}$ at cost ce. e is commonly observed.
4. P_1 and P_2 pay A amounts $w_1(e), w_2(e)$. These payments are commonly observed.

Equilibrium A **pure-strategy subgame-perfect equilibrium** is a pair of contracts $w_1^*, w_2^* \in W$, an acceptance strategy $d^* : W^2 \to \{0, 1\}$, an effort strategy $e^* : W^2 \to \{0, 1\}$ such that given contracts w_1^*, w_2^*, A optimally chooses $d^*(w_1^*, w_2^*)$ and $e^*(w_1^*, w_2^*)$. Given d^*, e^* and w_2^*, P_1 optimally chooses w_1^*, and given d^*, e^*, and w_1^*, P_2 optimally chooses w_2^*.

The Program Given contracts w_1 and w_2 specifying (s_1, b_1) and (s_2, b_2), if the Agent accepts these contracts, he will choose $e = 1$ if $b_1 + b_2 \geq c$, and he will choose $e = 0$ if $b_1 + b_2 \leq c$. Define

$$e(b_1, b_2) = \begin{cases} 1 & b_1 + b_2 \geq c \\ 0 & b_1 + b_2 \leq c, \end{cases}$$

and he will accept these contracts if

$$s_1 + b_1 e(b_1, b_2) + s_2 + b_2 e(b_1, b_2) - ce(b_1, b_2) \geq \bar{u}.$$

Suppose P_1 believes P_2 will offer contract (s_2, b_2). Then P_1 will choose $\hat{s}_1, \hat{b}_1$ to solve

$$\max_{\hat{s}_1, \hat{b}_1} p_1 e\left(\hat{b}_1, b_2\right) - \left(\hat{s}_1 + \hat{b}_1 e\left(\hat{b}_1, b_2\right)\right)$$

subject to the Agent's individual-rationality constraint

$$\hat{s}_1 + \hat{b}_1 e\left(\hat{b}_1, b_2\right) + s_2 + b_2 e\left(\hat{b}_1, b_2\right) - ce\left(\hat{b}_1, b_2\right) \geq \bar{u}.$$

P_1 will choose $\hat{s}_1$ so that this individual-rationality constraint holds with equality:

$$\hat{s}_1 + \hat{b}_1 e\left(\hat{b}_1, b_2\right) = \bar{u} - s_2 - b_2 e\left(\hat{b}_1, b_2\right) + ce\left(\hat{b}_1, b_2\right).$$

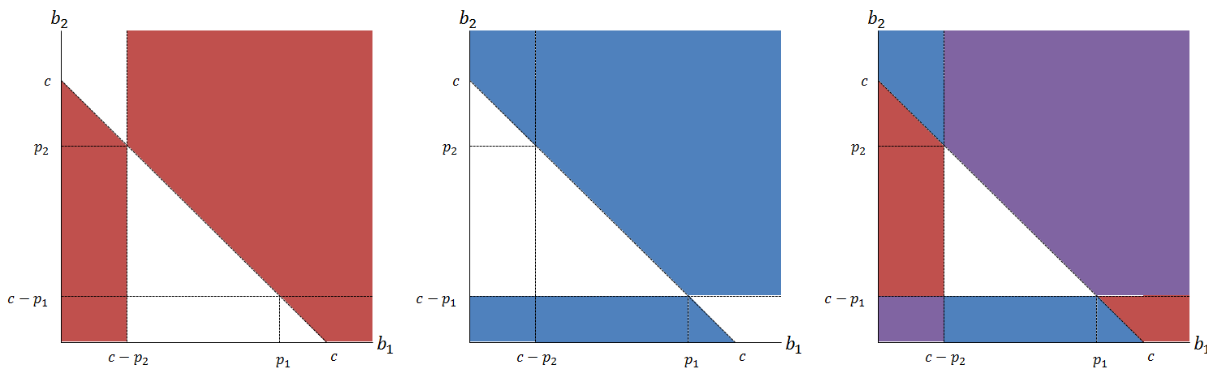
P_1's unconstrained problem is then

$$\max_{\hat{b}_1} p_1 e\left(\hat{b}_1, b_2\right) + b_2 e\left(\hat{b}_1, b_2\right) - ce\left(\hat{b}_1, b_2\right) - \bar{u} + s_2.$$

Since $\hat{b}_1$ only affects P_1's payoff inasmuch as it affects the Agent's effort choice, and since given any b_2 and any effort choice e, P_1 can choose a $\hat{b}_1$ so that the Agent will choose that effort choice. In other words, we can view Principal 1's problem as:

$$\max_{e \in \{0,1\}} (p_1 + b_2 - c)\, e.$$

When $p_1 + b_2 \geq c$, P_1 will choose b_1 to ensure that $e^*(b_1, b_2) = 1$. That is, when $b_2 \geq c - p_1$, it is a best response for P_1 to offer any contract (s_1, b_1) with $b_1 \geq c - b_2$ (and with s_1 such that the Agent's individual-rationality constraint holds with equality). When $p_1 + b_2 \leq c$, b_1 will be chosen to ensure that $e^*(b_1, b_2) = 0$. That is, when $b_2 \leq c - p_1$, it is a best response for P_1 to offer any contract (s_1, b_2) with $b_1 < c - b_2$. The following figures show best-response correspondences in this (b_1, b_2) space. In the first figure, the red regions represent the optimal choices of b_2 given a choice of b_1. In the second figure, the blue regions represent the optimal choices of b_1 given a choice of b_2. The third figure puts these together–the purple regions represent equilibrium contracts. Those equilibrium contracts in the upper-right region induce $e^* = 1$, while those in the lower-left region do not.



The set of equilibrium contracts is therefore any (s_1, b_1) and (s_2, b_2) such that either:

1. $b_1 \geq c - p_1$, $b_2 \geq c - p_2$ and $b_1 + b_2 \geq c$.

2. $0 \leq b_1 < c - p_1$ and $0 \leq b_2 < c - p_2$.

The first set of equilibrium contracts implement $e^* = 1$, while the second set of equilibrium contracts implement $e^* = 0$. Equilibrium contracts with $\{0 \leq b_i \leq c - p_i\}$ therefore represent a **coordination failure**.

Description of Free-Rider Version There is a risk-neutral Agent (A) and two risk-neutral Principals (P_1 and P_2). The Agent chooses an effort $e \in [0, 1]$ at cost $\frac{c}{2}e^2$. Output is $y \in Y = \{0, 1\}$ with $\Pr[y = 1|\, e] = e$. Principals 1 and 2 receive revenues $p_1 y$ and $p_2 y$, respectively. The Principals simultaneously offer contracts $w_1, w_2 \in W = \{w : Y \to \mathbb{R}\}$. Denote Principal i's contract offer by $w_i = s_i + b_i y$. If the Agent accepts a pair of contracts with total incentives $b = b_1 + b_2$, he incurs an additional cost $k \cdot b$. These costs are reduced-form, but we can think of them either as risk costs associated with higher-powered incentives or, if they were instead borne by the Principals, we could think of them as setup costs associated with writing higher-powered contracts (as in Battigalli and Maggi, 2002). The analysis would be similar in this latter case. The Agent has an outside option that yields utility $\bar{u}$ to the Agent and 0 to each Principal. If the outside option is not exercised, players' expected payoffs are:

$$
\begin{aligned}
\Pi_1 (w_1, w_2, e) &= p_1 e - w_1 \\
\Pi_2 (w_1, w_2, e) &= p_2 e - w_2 \\
U (w_1, w_2, e) &= w_1 + w_2 - \frac{c}{2} e^2 - k \cdot (b_1 + b_2)
\end{aligned}
$$

Timing The timing of the game is exactly the same as before.

1. P_1 and P_2 simutaneously offer contracts w_1, w_2 to A. Offers are commonly observed.

2. A accepts both contracts $(d = 1)$ or rejects both contracts $(d = 0)$ and receives $\bar{u}$ and the game ends. This decision is commonly observed.

3. If A accepts the contracts, he incurs cost $k \cdot (b_1 + b_2)$ and then chooses effort $e \in [0, 1]$ at cost $\frac{c}{2} e^2$. e is commonly observed.

4. Output $y \in Y$ is realized with $\Pr [y = 1 | e] = e$. Output is commonly observed.

5. P_1 and P_2 pay A amounts $w_1 (y), w_2 (y)$. These payments are commonly observed.

Equilibrium A **pure-strategy subgame-perfect equilibrium** is a pair of contracts $w_1^*, w_2^* \in W$, an acceptance strategy $d^* : W^2 \to \{0, 1\}$, an effort strategy $e^* : W^2 \to \{0, 1\}$ such that given contracts w_1^*, w_2^*, A optimally chooses $d^* (w_1^*, w_2^*)$ and $e^* (w_1^*, w_2^*)$. Given d^*, e^* and w_2^*, P_1 optimally chooses w_1^*, and given d^*, e^*, and w_1^*, P_2 optimally chooses w_2^*.

The Program Given total incentives $b = b_1 + b_2$, A chooses effort e to solve

$$
\max_e be - \frac{c}{2} e^2,
$$

or $e^*(b) = \frac{b}{c}$. Suppose P_1 believes P_2 will offer contract (s_2, b_2). Then P_1's problem is to

$$\max_{\hat{b}_1, \hat{s}_1} p_1 e^*\left(\hat{b}_1 + b_2\right) - \hat{s}_1 - \hat{b}_1 e^*\left(\hat{b}_1 + b_2\right)$$

subject to A's individual-rationality constraint

$$\hat{s}_1 + \hat{b}_1 e^*\left(\hat{b}_1 + b_2\right) + s_2 + b_2 e^*\left(\hat{b}_1 + b_2\right) - \frac{c}{2} e^*\left(\hat{b}_1 + b_2\right)^2 - k \cdot \left(\hat{b}_1 + b_2\right) \geq \bar{u}.$$

As in the previous models, P_1 will choose s_1 so that this constraint holds with equality:

$$\hat{s}_1 + \hat{b}_1 e^*\left(\hat{b}_1 + b_2\right) = \bar{u} + \frac{c}{2} e^*\left(\hat{b}_1 + b_2\right)^2 + k \cdot \left(\hat{b}_1 + b_2\right) - s_2 - b_2 e^*\left(\hat{b}_1 + b_2\right).$$

P_1's unconstrained problem is then to

$$\max_{\hat{b}_1} p_1 e^*\left(\hat{b}_1 + b_2\right) + b_2 e^*\left(\hat{b}_1 + b_2\right) - \frac{c}{2} e^*\left(\hat{b}_1 + b_2\right)^2 - k \cdot \left(\hat{b}_1 + b_2\right),$$

which yields first-order conditions

$$\begin{aligned} 0 &= p_1 \frac{\partial e^*}{\partial \hat{b}_1} + b_2 \frac{\partial e^*}{\partial \hat{b}_1} - c e^*\left(b_1 + b_2\right) \frac{\partial e^*}{\partial \hat{b}_1} - k \\ &= \left(p_1 + b_2 - \left(b_1^* + b_2\right)\right) \frac{1}{c} - k \end{aligned}$$

so that $b_1^* = p_1 - ck$. This choice of b_1 is independent of b_2. Analogously, P_2 will choose a contract with $b_2^* = p_2 - ck$. The Agent's equilibrium effort will satisfy

$$e^*\left(b_1^* + b_2^*\right) = \frac{p}{c} - 2k.$$

If the two Principals could collude and offer a single contract $w = s + by$ to the agent, they would offer a contract that solves:

$$\max_{b} pe(b) - \frac{c}{2} e(b)^2 - kb,$$

where $e(b) = \frac{b}{c}$. The associated first-order conditions are

$$\frac{p}{c} - \frac{b^C}{c} = k$$

or

$$b^C = p - ck$$

and therefore equilibrium effort would be

$$e^*\left(b^C\right) = \frac{p}{c} - k.$$

In particular, $e^*\left(b^C\right) = e^*(b^*) + k > e^*(b^*)$. This effect is often referred to as the free-rider effect in common-agency models.

2.2 No Contracts (Updated: Apr 1 2014)

In many environments, contractible measures of performance may be so bad as to render them useless. Yet, aspects of performance that are relevant for the firm's objectives may be observable, but for whatever reason, they cannot be written into a formal contract. These aspects of performance may then form the basis for informal reward schemes. We will discuss two classes of models that build off this insight. We will then examine an environment in which formal contracts cannot be made contingent on the appropriateness of decisions, but the Principal is able to restrict the set of decisions that the Agent is allowed to make.

2.2.1 Career Concerns (Updated: Apr 1 2014)

An Agent's performance within a firm may be observable to outside market participants–for example, fund managers' returns are published in prospectuses, academics post their papers online publicly, a CEO's performance is partly announced in quarterly earnings reports. Holmstrom (1982/1999) developed a model to show that in such an environment, even when formal performance-contingent contracts are impossible to write, workers may be motivated to work hard out of a desire to convince "the market" that they are intrinsically productive in the hopes that doing so will attract favorable outside offers in the future–that is, they are motivated by their own career concerns.

Description There are two risk-neutral Principals, whom we will denote by P_1 and P_2, and a risk-neutral Agent (A) who interact in periods $t = 1, 2$. The Agent has ability θ, which is drawn from a normal distribution, $\theta \sim N\left(m_0, h_0^{-1}\right)$. θ is unobservable by all players, but all players know the distribution from which it is drawn. In each period, the Agent chooses an effort level $e_t \in E$ at cost $c\left(e_t\right)$ (with $c\left(0\right) = c'\left(0\right) = 0 < c', c''$) that, together with his ability and luck (denoted by ε_t), determine his output $y_t \in Y$ as follows:

$$y_t = \theta + e_t + \varepsilon_t.$$

Luck is also normally distributed, $\varepsilon_t \sim N\left(0, h_\varepsilon^{-1}\right)$ and is independent across periods and independent from θ. This output accrues to whichever Principal employs the Agent in period t. At the beginning of each period, each Principal i offers the Agent a short-term contract $w_i \in W \subset \{w_i : M \to \mathbb{R}\}$, where M is the set of outcomes of a performance measure. The Agent has to accept one of the contracts, and if he accepts Principal i's contract in period t, then Principal $j \neq i$ receives 0 in period t. For now, we will assume that there are no available performance measures, so short-term contracts can only take the form of a constant wage.

Comment on Assumption. Do you think the assumption that the Agent does not know more about his own productivity than the Principals do is sensible?

If Principal P_i employs the Agent in period t, the agent chooses effort e_t, and output y_t is realized, payoffs are given by

$$\begin{aligned}\pi_i(w_{it}, e_t, y_t) &= y_t - w_{it} \\ \pi_j(w_{it}, e_t, y_t) &= 0 \\ u_i(w_{it}, e_t, y_t) &= w_{it} - c(e_t).\end{aligned}$$

Players share a common discount factor of $\delta < 1$.

Timing There are two periods $t = 1, 2$. In each period, the following stage game is played:

1. P_1 and P_2 propose contracts w_{1t} and w_{2t}. These contracts are commonly observed
2. A chooses one of the two contracts. The Agent's choice is commonly observed. If A chooses contract offered by P_i, denote his choice by $d_t = i$. The set of choices is denoted by $D = \{1, 2\}$.
3. A receives transfer w_{it}. This transfer is commonly observed.
4. A chooses effort e_t and incurs cost $c(e_t)$. e_t is only observed by A.
5. Output y_t is realized and accrues to P_i. y_t is commonly observed.

Equilibrium The solution concept is Perfect-Bayesian Equilibrium. A **Perfect-Bayesian Equilibrium** of this game consists of a strategy profile $\sigma^* = \left(\sigma^*_{P_1}, \sigma^*_{P_2}, \sigma^*_A\right)$ and a belief profile μ^* (defining beliefs of each player about the distribution of θ at each information set) such that σ^* is sequentially rational for each player given his beliefs (i.e., each player plays the best response at each information set given his beliefs) and μ^* is derived from σ^* using Bayes's rule whenever possible.

It is worth spelling out in more detail what the strategy space is. By doing so, we can get an appreciation for how complicated this seemingly simple environment is, and how different assumptions of the model contribute to simplifying the solution. Further, by understanding the role of the different assumptions, we will be able to get a sense for what directions the model could be extended in without introducing great complexity.

Each Principal i chooses a pair of contract-offer strategies $w_{i1}^* : \Delta(\Theta) \to \mathbb{R}$ and $w_{i2}^* : W \times D \times Y \times \Delta(\Theta) \to \mathbb{R}$. The first-period offers depend only on each Principal's beliefs about the Agent's type (as well as their equilibrium conjectures about what the Agent will do). The second-period offer can also be conditioned on the first-period contract offerings, the Agent's first-period contract choice, and the Agent's first-period output. In equilibrium, it will be the case that these variables determine the second-period contract offers only inasmuch as they determine each Principal's beliefs about the Agent's type.

The Agent chooses a set of acceptance strategies in each period, $d_1 : W^2 \times \Delta(\Theta) \to \{1, 2\}$ and $d_2 : W^4 \times D \times E \times Y \times \Delta(\Theta) \to \{1, 2\}$ and a set of effort strategies $e_1 : W^2 \times D \times \Delta(\Theta) \to \mathbb{R}_+$ and $e_2 : W^4 \times D \times E \times Y \times \Delta(\Theta) \to \mathbb{R}_+$. In the first period, the agent chooses which contract to accept based on which ones are offered as well as his beliefs about his own type. In the present model, the contract space is not very rich (since it is only the set of scalars), so it will turn out that the Agent does not want to condition his acceptance decision on his beliefs about his own ability. This is not necessarily the case in richer models in which Principals are allowed to offer contracts involving performance-contingent payments. The Agent then chooses effort on the basis of which contracts were available, which one he chose, and his beliefs about his type. In the second period, his acceptance decision and effort choice can also be conditioned on events that occurred in the first period.

It will in fact be the case that this game has a unique Perfect-Bayesian Equilibrium, and in this Perfect-Bayesian equilibrium, both the Principals and the Agent will use **public** strategies in which $w_{i1}^* : \Delta(\Theta) \to \mathbb{R}$, $w_{i2}^* : \Delta(\Theta) \to \mathbb{R}$, $d_1 : W^2 \to \{1, 2\}$, $d_2 : W^2 \to \{1, 2\}$, $e_1 \in \mathbb{R}_+$ and $e_2 \in \mathbb{R}_+$.

The Program Sequential rationality implies that the Agent will choose $e_2^* = 0$ in the second period, no matter what happened in previous periods. This is because no further actions or payments that the Agent will receive are affected by the Agent's effort choice in the second period. Given that the agent knows his effort choice will be the same no matter which contract he chooses, he will choose whichever contract offers him a higher payment.

In turn, the Principals will each offer a contract in which they earn zero expected profits. This is because they have the same beliefs about the Agent's ability. This is the case since they have the same prior and have seen the same public history, and in equilibrium, they have the same conjectures about the Agent's strategy and therefore infer the same information about the Agent's ability. As a result, if one Principal offers a contract that will yield him positive expected profits, the other Principal will offer a contract that pays the Agent slightly more, and the Agent will accept the latter contract. The second-period contracts offered will therefore be

$$w_{12}^* \left(\hat{\theta}\left(y_1\right)\right) = w_{22}^* \left(\hat{\theta}\left(y_1\right)\right) = w_2^* \left(\hat{\theta}\left(y_1\right)\right) = E\left[y_2 | \, y_1, \sigma^*\right] = E\left[\theta | \, y_1, \sigma^*\right],$$

where $\hat{\theta}\left(y_1\right)$ is the equilibrium conditional distribution of θ given realized output y_1.

If the agent chooses e_1 in period 1, first-period output will be $y_1 = \theta + e_1 + \varepsilon_1$. Given conjectured effort e_1^*, the Principals' beliefs about the Agent's ability will be based on two signals: their prior, and the signal $y_1 - e_1^* = \theta + \varepsilon_1$, which is also normally distributed with mean m_0 and variance $h_0^{-1} + h_\varepsilon^{-1}$. The joint distribution is therefore

$$\begin{bmatrix} \theta \\ \theta + \varepsilon_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \theta \\ \varepsilon_1 \end{bmatrix} \sim N\left(\begin{bmatrix} m_0 \\ m_0 \end{bmatrix}, \begin{bmatrix} h_0^{-1} & h_0^{-1} \\ h_0^{-1} & h_0^{-1} + h_\varepsilon^{-1} \end{bmatrix} \right)$$

Their beliefs about θ conditional on these signals will therefore be normally distributed:

$$\theta | \, y_1 \sim N\left(\varphi y_1 + \left(1 - \varphi\right) m_0, \frac{1}{h_\varepsilon + h_0}\right),$$

where $\varphi = \frac{h_\varepsilon}{h_0+h_\varepsilon}$ is the signal-to-noise ratio. Here, we used the normal updating formula, which just to jog your memory is stated as follows. If X is a $K \times 1$ random vector and Y is an $N - K$ random vector, then if

$$\begin{bmatrix} X \\ Y \end{bmatrix} \sim N\left(\begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix}, \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma'_{XY} & \Sigma_{YY} \end{bmatrix}\right),$$

then

$$X|\,Y = y \sim N\left(\mu_X + \Sigma_{XY}\Sigma_{YY}^{-1}\left(y - \mu_Y\right), \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma'_{XY}\right).$$

Therefore, given output y_1, the Agent's second-period wage will be

$$w_2^*\left(\hat{\theta}\left(y_1\right)\right) = \varphi\left(y_1 - e_1^*\right) + \left(1 - \varphi\right) m_0 = \varphi\left(\theta + e_1 + \varepsilon_1 - e_1^*\right) + \left(1 - \varphi\right) m_0.$$

In the first period, the Agent chooses a non-zero effort level, even though his first-period contract does not provide him with performance-based compensation. He chooses a non-zero effort level, because it affects the distribution of output, which the Principals use in the second period to infer his ability. In equilibrium, of course, they are not fooled by his effort choice.

Given an arbitrary belief about his effort choice, $\hat{e}_1$, the signal the Principals use to update their beliefs about the Agent's type is $y_1 - \hat{e}_1 = \theta + \varepsilon_1 + e_1 - \hat{e}_1$. The agent's incentives to exert effort in the first period to shift the distribution of output are therefore the same no matter what the Principals conjecture his effort choice to be. He will therefore choose effort e_1^* in the first period to solve

$$\max_{e_1} -c\left(e_1\right) + \delta E_{y_1}\left[w_2^*\left(\hat{\theta}\left(y_1\right)\right)\middle| e_1\right] = \max_{e_1} -c\left(e_1\right) + \delta\left(\varphi\left(\theta + e_1 - e_1^*\right) + \left(1 - \varphi\right) m_0\right),$$

so that he will choose

$$c'\left(e_1^*\right) = \delta\frac{h_\varepsilon}{h_0 + h_\varepsilon}.$$

This second-period effort choice is, of course, less than first-best, since first-best effort satisfies $c'\left(e_1^{FB}\right) = 1$. He will choose a higher effort level in the first period the less he discounts the future (δ larger), the more prior uncertainty there is about his type (h_0 small), and the more informative output is about his ability (h_ε large). Finally, given that the Agent will choose e_1^*, the first-period wages will be

$$w_{11}^* = w_{21}^* = E\left[y_1\right] = m_0 + e_1^*.$$

This model has a number of nice features. First, despite the fact that the Agent receives no formal incentives, he still chooses a positive effort level, at least in the first period. Second, he does not choose first-best effort (indeed, in versions of the model with three or more periods, he may initially choose excessively high effort), even though there is perfect competition in the labor market for his services. When he accepts an offer, he cannot commit to choose a particular effort level, so competition does not necessarily generate efficiency when there are contracting frictions.

The model is remarkably tractable, despite being quite complicated. This is largely due to the fact that this is a symmetric-information game, so players neither infer nor communicate information about the agent's type when making choices. The functional-form choices are also aimed at ensuring that it not only starts out as a symmetric information game, but it also remains one as it progresses. At the end of the first period, if one of the Principals (say the one that the Agent worked for in the first period) learned more about the Agent's type than the other Principal did, then there would be asymmetric information at the wage-offering stage in the second period.

This model extends nicely to three or more periods. In such an extension, however, if the Agent's effort affected the variance of output, he would have more information about his type at the beginning of the second period than the Principals would. This is because he would have more information about the conditional variance of his own ability, because he knows what effort he chose. In turn, his choice of contract in the second period would be informative about what effort level he would be likely to choose in the second period, which would in turn influence the contract offerings. If ability and effort interact, and their interaction cannot be separated out from the noise with a simple transformation (e.g., if $y_t = \theta e_t + \varepsilon_t$), then the Agent would acquire private information about his marginal returns to effort, which would have a similar effect. For these reasons, the model has seen very little application to environments with more than two periods, except in a couple special cases (see Bonatti and Horner (2013) for a recent example with public all-or-nothing learning).

Exercise 3. *Can the above model be extended in a straightforward way to environments with more than* 3 *periods if the Agent has imperfect recall regarding the effort level he chose in past periods?*

2.2.2 Relational Incentive Contracts (Updated: Apr 1 2014)

If an Agent's performance is commonly observed only by other members of his organization, or if the market is sure about his intrinsic productivity, then the career concerns motives above cannot serve as motivation. However, individuals may form long-term attachments with an organization. In such long-term relationships, goodwill can arise as an equilibrium phenomenon, and fear of shattering this goodwill can motivate individuals to perform well and to reward good performance. This intuition is captured in models of relational contracts (informal contracts enforced by relationships). An entire section of this course will be devoted to studying many of the issues that arise in such models, but for now we will look at the workhorse model in the literature to get some of the more general insights.

The workhorse model is an infinitely repeated Principal-Agent game with publicly observed actions. We will characterize the "optimal relational contract" as the equilibrium of the repeated game that either maximizes the Principal's equilibrium payoffs or the Principal and Agent's joint equilibrium payoffs. A couple comments are in order at this point. First, these are applied models of repeated games and therefore tend to focus on situations where the discount factor is not close to 1, asking questions like "how much effort can be sustained in equilibrium?"

Second, such models often have many equilibria, and therefore we will be taking a stance on equilibrium selection in their analysis. The criticism that such models have no predictive power is, as Kandori puts it "... misplaced if we regard the theory of repeated games as a theory of informal contracts. Just as anything can be enforced when the party agrees to sign a binding contract, in repeated games [many outcomes can be] sustained if players agree on an equilibrium. Enforceability of a wide range of outcomes is the essential property of effective contracts, formal or informal." (Kandori, 2008, p. 7) Put slightly differently, focusing on optimal contracts when discussing formal contract design is analogous to focusing on optimal relational contracts when discussing repeated principal-agent models.

Description A risk-neutral Principal and risk-neutral Agent interact repeatedly in periods $t = 0, 1, 2, \dots$. In period t, the Agent chooses an effort level $e_t \in E$ at cost $c(e_t) = \frac{c}{2}e_t^2$ that determines output $y_t = e_t \in Y$, which accrues to the Principal. The output can be sold on the product market for price p. At the beginning of date t, the Principal proposes a compensation package to the agent. This compensation consists of a fixed salary s_t and a contingent payment $b_t : E \to \mathbb{R}$ (with positive values denoting a transfer from the Principal to the Agent and negative values denoting a transfer from the Agent to the Principal), which can depend on the Agent's effort choice. The Agent can accept the proposal (which we denote by $d_t = 1$) or reject it (which we denote by $d_t = 0$) in favor of an outside option that yields per-period utility $\bar{u}$ for the Agent and $\bar{\pi}$ for the Principal. If the Agent accepts the proposal, the Principal is legally compelled to pay the transfer s_t, but she is not legally compelled to pay the contingent payment b_t.

Timing The stage game has the following five stages

1. P makes A a proposal (b_t, s_t)

2. A accepts or rejects in favor of outside opportunity yielding $\bar{u}$ to A and $\bar{\pi}$ to P.

3. P pays A an amount s_t

4. A chooses effort $\hat{e}_t$ at cost $c(\hat{e}_t)$, which is commonly observed

5. P pays A a transfer $\hat{b}_t$

Equilibrium The Principal is not legally required to make the promised payment b_t, so in a one-shot game, she would always choose $\hat{b}_t = 0$ (or analogously, if $b_t < 0$, the Agent is not legally required to pay b_t, so he would choose $\hat{b}_t = 0$). However, since the players are engaged in a long-term relationship and can therefore condition future play on this transfer, nonzero transfers can potentially be sustained as part of an equilibrium.

Whenever we consider repeated games, we will always try to spell out explicitly the variables that players can condition their behavior on. This exercise is tedious but important. Let $h_0^t = \left\{s_0, d_0, \hat{e}_0, \hat{b}_0, \ldots, s_{t-1}, d_{t-1}, \hat{e}_{t-1}, \hat{b}_{t-1}\right\}$ denote the history up to the beginning of date t. In this game, all variables are commonly observed, so the history up to date t is a public history. We will also adopt the notation $h_s^t = h^t \cup \{s_t\}$, $h_d^t = h_s^t \cup \{d_t\}$, and $h_e^t = h_d^t \cup \{\hat{e}_t\}$. (If we analogously defined h_b^t, it would be the same as h_0^{t+1}, so we will refrain from doing so.) Finally, let $\mathcal{H}_0^t, \mathcal{H}_s^t, \mathcal{H}_d^t$, and $\mathcal{H}_e^t$ denote, respectively, the sets of such histories.

Following Levin (2003), we define a **relational contract** to be a complete plan for the relationship. It describes (1) the salary that the Principal should offer the Agent ($h_0^t \mapsto s_t$), (2) whether the Agent should accept the offer ($h_s^t \mapsto d_t$), (3) what effort level the Agent should choose ($h_d^t \mapsto \hat{e}_t$), and (4) what bonus payment the Principal should make $\left(h_e^t \mapsto \hat{b}_t\right)$. A relational contract is **self-enforcing** if it describes a subgame-perfect equilibrium of the repeated game. An **optimal relational contract** is a self-enforcing relational contract that yields higher equilibrium payoffs for the Principal than any other self-enforcing relational contract. It is important to note that a relational contract describes behavior on and off the equilibrium path.

Comment. *Early papers in the relational-contracting literature (Bull, 1987; MacLeod and*

Malcomson, 1989; Baker, Gibbons, and Murphy, 1994) referred to the equilibrium of the game instead as an implicit (as opposed to relational) contract. More recent papers eschew the term implicit, because the term "implicit contracts" has a connotation that seems to emphasize whether agreements are common knowledge, whereas the term "relational contracts" more clearly focuses on whether agreements are enforced formally or must be self-enforcing.

The Program Though the stage game is relatively simple, and the game has a straightforward repeated structure, solving for the optimal relational contract should in principle seem like a daunting task. There are tons of things that the Principal and Agent can do in this game (the strategy space is quite rich), many of which are consistent with equilibrium play–there are lots of equilibria, some of which may have complicated dynamics. Our objective is to pick out, among all these equilibria, those that maximize the Principal's equilibrium payoffs.

Thankfully, there are several nice results (many of which are contained in Levin (2003) but have origins in the preceding literature) that make this task achievable. We will proceed in the following steps:

1. We will argue, along the lines of Abreu (1988), that the unique stage game SPNE is an optimal punishment.

2. We will show that optimal reward schedules are "forcing." That is, they pay the Agent a certain amount if he chooses a particular effort level, and they revert to punishment otherwise. An optimal relational contract will involve an optimal reward scheme.

3. We will then show that distribution and efficiency can be separated out in the stage game. Ex-ante transfers have to satisfy participation constraints, but they otherwise do not affect incentives or whether continuation payoffs are self-enforcing.

4. We will show that an optimal relational contract is sequentially optimal on the equilibrium path. Increasing future surplus is good for ex-ante surplus, which can be divided in any way, according to (3), and it improves the scope for incentives in the current period. Total future surplus is always maximized in an optimal relational contract, and

since the game is a repeated game, this implies that total future surplus is therefore constant in an optimal relational contract.

5. We will then argue that we can restrict attention to stationary relational contracts. By (4), the total future surplus is constant in every period. Contemporaneous payments and the split of continuation payoffs are perfect substitutes for motivating effort provision and bonus payments and for participation. Therefore, we can restrict attention to agreements that "settle up" contemporaneously rather than reward and punish with continuation payoffs.

6. We will then solve for the set of stationary relational contracts, which is not so complicated. This set will contain an optimal relational contract.

In my view, while the restriction to stationary relational contracts is helpful for being able to tractably characterize the optimal relational contract, the important economic insights are actually that the relational contract is sequentially optimal and how this result depends on the separation of distribution and efficiency. The separation of distribution and efficiency in turn depends on several assumptions: risk-neutrality, unrestricted and costless transfers, and a simple information structure. Later in the course, we will return to these issues and think about settings where one or more of these assumptions is not satisfied.

Step 1 is straightforward. In the unique SPNE of the stage game, the Principal never pays a positive bonus, the Agent exerts zero effort, and he rejects any offer the Principal makes. The associated payoffs are $\bar{u}$ for the Agent and $\bar{\pi}$ for the Principal. It is also straightforward to show that these are also the Agent's and Principal's maxmin payoffs, and therefore they constitute an optimal penal code (Abreu, 1988). Define $\bar{s} = \bar{u} + \bar{\pi}$ to be the outside surplus.

Next, consider a relational contract that specifies, in the initial period, payments w and $b(\hat{e})$, an effort level e, and continuation payoffs $u(\hat{e})$ and $\pi(\hat{e})$. The equilibrium payoffs of this relational contract, if accepted are:

$$\begin{aligned} u &= (1-\delta)(w - c(e) + b(e)) + \delta u(e) \\ \pi &= (1-\delta)(p \cdot e - w - b(e)) + \delta \pi(e). \end{aligned}$$

Let $s = u + \pi$ be the equilibrium contract surplus. This relational contract is self-enforcing if the following four conditions are satisfied.

1. Participation:

$$u \geq \bar{u}, \pi \geq \bar{\pi}$$

2. Effort-IC:

$$e \in \underset{\hat{e}}{\operatorname{argmax}} \{(1-\delta)(-c(\hat{e}) + b(\hat{e})) + \delta u(\hat{e})\}$$

3. Payment:

$$\begin{aligned} (1-\delta)(-b(e)) + \delta \pi(e) &\geq \delta \bar{\pi} \\ (1-\delta) b(e) + \delta u(e) &\geq \delta \bar{u} \end{aligned}$$

4. Self-enforcing continuation contract: $u(e)$ and $\pi(e)$ correspond to a self-enforcing relational contract that will be initiated in the next period.

Step 2: Define the Agent's **reward schedule** under this relational contract by

$$R(\hat{e}) = b(\hat{e}) + \frac{\delta}{1-\delta} u(\hat{e}).$$

The Agent's no-reneging constraint implies that $R(\hat{e}) \geq \frac{\delta}{1-\delta}\bar{u}$ for all $\hat{e}$. Given a proposed effort level e, suppose there is some other effort level $\hat{e}$ such that $R(\hat{e}) > \frac{\delta}{1-\delta}\bar{u}$. Then we can define an alternative relational contract in which everything else is the same, but $\tilde{R}(\hat{e}) = R(\hat{e}) - \varepsilon$ for some $\varepsilon > 0$. The payment constraints remain satisfied, and the effort-IC constraint becomes easier to satisfy. Therefore, such a change makes it possible to weakly improve at least one player's equilibrium payoff. Therefore, it has to be that $R(\hat{e}) = \frac{\delta}{1-\delta}\bar{u}$ for all $\hat{e} \neq e$.

Step 3: Consider an alternative relational contract in which everything else is the same, but $\tilde{w} = w - \varepsilon$ for some $\varepsilon \neq 0$. This changes the equilibrium payoffs u, π to $\tilde{u}, \tilde{\pi}$ but not the joint surplus s. Further, it does not affect the effort-IC, the payment, or the self-enforcing continuation contract conditions. As long as $\tilde{u} \geq \bar{u}$ and $\tilde{\pi} \geq \bar{\pi}$, then the proposed relational contract is still self-enforcing.

Define the value s^* to be the maximum total surplus generated by any self-enforcing relational contract. The set of possible payoffs under a self-enforcing relational contract is then $\{(u, \pi) : u \geq \bar{u}, \pi \geq \bar{\pi}, u + \pi \leq s^*\}$. For a given relational contract to satisfy the self-enforcing continuation contract condition, it then has to be the case that for any equilibrium effort e,

$$(u(e), \pi(e)) \in \{(u, \pi) : u \geq \bar{u}, \pi \geq \bar{\pi}, u + \pi \leq s^*\}.$$

Step 4: Suppose the continuation relational contract satisfies $u(e) + \pi(e) < s^*$. Then $\pi(e)$ can be increased in a self-enforcing relational contract, holding everything else the same. Increasing $\pi(e)$ does not affect the effort-IC constraint, it relaxes both the Principal's participation and payment constraints, and it increases equilibrium surplus. The original relational contract is then not optimal. Therefore, any optimal relational contract has to satisfy $s(e) = u(e) + \pi(e) = s^*$.

Step 5: Suppose the proposed relational contract is optimal and generates surplus $s(e)$. By the previous step, it has to be the case that $s(e) = e - c(e) = s^*$. This in turn implies that optimal relational contracts involve the same effort choice, e^*, in each period. Now we want to construct an optimal relational contract that provides the same incentives for the agent to exert effort, for both players to pay promised bonus payments, and also yields continuation payoffs that are equal to equilibrium payoffs (i.e., not only is the action that is chosen the same in each period, but so are equilibrium payoffs). To do so, suppose an optimal relational contract involves reward scheme $R(\hat{e}) = \frac{\delta}{1-\delta}\bar{u}$ for $\hat{e} \neq e^*$ and

$$R(e^*) = b(e^*) + \frac{\delta}{1-\delta} u(e^*).$$

Now, consider an alternative reward scheme $\tilde{R}(e^*)$ that provides the same incentives to the agent but leaves him with a continuation payoff of u^*:

$$\tilde{R}(e^*) = \tilde{b}(e^*) + \frac{\delta}{1-\delta} u^* = R(e^*).$$

This reward scheme also leaves him with an equilibrium utility of u^*

$$\begin{aligned} u^* &= (1-\delta)(w - c(e^*) + b(e^*)) + \delta u(e^*) = (1-\delta)(w - c(e^*) + R(e^*)) \\ &= (1-\delta)\left(w - c(e^*) + \tilde{R}(e^*)\right) = (1-\delta)\left(w - c(e^*) + \tilde{b}(e^*)\right) + \delta u^*. \end{aligned}$$

Since $\bar{u} \leq u^* \leq s^* - \bar{\pi}$, this alternative relational contract also satisfies the participation constraints.

Further, this alternative relational contract also satisfies all payment constraints, since by construction,

$$\tilde{b}(e^*) + \frac{\delta}{1-\delta} u^* = b(e^*) + \frac{\delta}{1-\delta} u(e^*),$$

and this equality also implies the analogous equality for the Principal (since $s^* = u^* + \pi^*$ and $s^* = u(e^*) + \pi(e^*)$):

$$-\tilde{b}(e^*) + \frac{\delta}{1-\delta} \pi^* = -b(e^*) + \frac{\delta}{1-\delta} (\pi(e^*)).$$

Finally, the continuation payoffs are (u^*, π^*), which can themselves be part of this exact same self-enforcing relational contract initiated the following period.

Step 6: The last step allows us to set up a program that we can solve to find an optimal relational contract. A stationary effort level e generates total surplus $s = e - c(e)$. The Agent is be willing to choose effort level e if he expected to be paid a bonus b satisfying

$$b + \frac{\delta}{1-\delta}(u - \bar{u}) \geq c(e).$$

That is, as long as his effort costs are less than the bonus b and the change in his continuation payoff that he would experience if he did not choose effort level e. Similarly, the Principal is willing to pay a bonus b if

$$\frac{\delta}{1-\delta}(\pi - \bar{\pi}) \geq b.$$

A necessary condition for both of these inequalities to be satisfied is that

$$\frac{\delta}{1-\delta}(s - \bar{s}) \geq c(e).$$

This condition is also sufficient for an effort level e to be sustainable in a stationary relational contract, since if it is satisfied, there is a b such that the preceding two inequalities are satisfied. This pooled inequality is referred to as the **dynamic-enforcement constraint.**

The Program: Putting all this together, then, an optimal relational contract will involve an effort level that solves

$$\max_e pe - \frac{c}{2}e^2$$

subject to the dynamic-enforcement constraint:

$$\frac{\delta}{1-\delta}\left(pe - \frac{c}{2}e^2 - \bar{s}\right) \geq \frac{c}{2}e^2.$$

The first-best effort level $e^{FB} = \frac{p}{c}$ solves this problem as long as

$$\frac{\delta}{1-\delta}\left(pe^{FB} - \frac{c}{2}\left(e^{FB}\right)^2 - \bar{s}\right) \geq \frac{c}{2}\left(e^{FB}\right)^2,$$

or

$$\delta \geq \frac{p^2}{2p^2 - 2c\bar{s}}.$$

Otherwise, the optimal effort level e^* is the larger solution to the dynamic-enforcement constraint, when it holds with equality:

$$e^* = \frac{p}{c}\left(\delta + \sqrt{\frac{p^2\delta^2 - 2\delta\bar{s}c}{p^2}}\right).$$

For all $\delta < \frac{p^2}{2p^2 - 2c\bar{s}}$, $\delta + \sqrt{\frac{p^2\delta^2 - 2\delta\bar{s}c}{p^2}} < 1$, so $e^* < e^{FB}$.

Comment. *People not familiar or comfortable with these models often try to come up with ways to artificially generate commitment. For example, they might propose something along the lines of, "If the problem is that the Principal doesn't have the incentives to pay a large bonus when required to, why doesn't the Principal leave a pot of money with a third-party enforcer that she will lose if she doesn't pay the bonus?" This proposal seems somewhat compelling, except for the fact that it would only solve the problem if the third-party enforcer could withhold that pot of money from the Principal if and only if the Principal breaks her promise to the Agent. Of course, this would require that the third-party enforcer condition its behavior on whether the Principal and the Agent cooperate. If the third-party enforcer could do this, then the third-party enforcer could presumably also enforce a contract that conditions on these events as well, which would imply that cooperation is contractible. On the other hand, if the third-party enforcer cannot conditionally withhold the money from the Principal, then the Principal's reneging temptation will consist of the joint temptation to* (a) *not pay the bonus she promised the agent and* (b) *recover the pot of money from the third-party enforcer.*

There may indeed be an inherent tension among the assumptions that (1) *cooperation is common knowledge between the Principal and the Agent,* (2) *the Principal and the Agent can make non-contingent transfers that are costlessly compelled by a third-party enforcer (i.e., they are contractible), and* (3) *they cannot be compelled to make transfers that are contingent upon whether cooperation occurs. We will (hopefully) return to this issue when we discuss the foundations of incomplete contracting.*

2.3 Delegation

Alonso and Matouschek (2008)

3 Decision Making in Organizations (Updated: Apr 1 2014)

In the previous section, we considered environments in which one party (the Agent) chose actions that affected the payoffs of another party (the Principal), and we asked how the Principal could motivate the Agent to choose actions that were more favorable for the Principal. We considered the use of formal incentive contracts, and we also looked at environments in which such contracts were unavailable. Throughout, however, we took as given that only the agent was able to make those decisions. In other words, **decision rights** were **inalienable**. In this section, we will allow for the possibility that control can be transferred from one party to the other, and we will ask when one party or the other should possess these decision rights.

If in principle, important decisions could be made by the Principal, why would the Principal ever want to delegate such decisions to an Agent? The models we will examine will focus on three reasons. First, the Agent may be better-informed about what decisions are available or what their payoff consequences are, and delegation may provide the Agent with the opportunity to act upon this information. The trade-off is then between a loss of control under delegation and a loss of information under centralization. In his book on the design of bureaucracies, James Q. Wilson concludes that "In general, authority [decision rights] should be placed at the lowest level at which all essential elements of information are available." We will see that there is a cost of moving decision rights downward, so that things are not as straightforward as Wilson's conclusions.

Another reason that delegation might be beneficial for the firm is that an agent who is able to act upon information values having information and therefore has the incentives to acquire it. The trade-off in these models is then between a loss of control under delegation and reduced information acquisition under centralization. Finally, when players are able to communicate strategically, the allocation of control may interact with their incentives to communicate truthfully.

3.1 Influence Activities

Milgrom and Roberts (1988)

3.2 Authority

3.2.1 Informed Agent

Simon (1951)

Dessein (2002)

3.2.2 Endogenous Information Acquisition

Aghion and Tirole (1997)

3.3 Centralization vs Decentralization

Alonso, Dessein, and Matouschek (2008)

4 Boundaries of the Firm (Updated: April 21 2014)

The central question goes back to Ronald Coase (1937): if markets are so great at coordinating productive activity, why is productive activity carried out within firms rather than by self-employed individuals who transact on a spot market? And indeed it is, as Herbert Simon (1991) vividly illustrated:

> A mythical visitor from Mars... approaches Earth from space, equipped with a telescope that reveals social structures. The firms reveal themselves, say, as solid green areas with faint interior contours marking out divisions and departments. Market transactions show as red lines connecting firms, forming a network in the

> spaces between them. Within firms (and perhaps even between them) the approaching visitor also sees pale blue lines, the lines of authority connecting bosses with various levels of workers... No matter whether our visitor approached the United States or the Soviet Union, urban China or the European Community, the greater part of the space below it would be within the green areas, for almost all inhabitants would be employees, hence inside the firm boundaries. Organizations would be the dominant feature of the landscape. A message sent back home, describing the scene, would speak of "large green areas interconnected by red lines." It would not likely speak of "a network of red lines connecting green spots." ...When our visitor came to know that the green masses were organizations and the red lines connecting them were market transactions, it might be surprised to hear the structure called a market economy. "Wouldn't 'organizational economy' be the more appropriate term?" it might ask.

It is obviously difficult to put actual numbers on the relative importance of trade within and between firms, since, I would venture to say, most transactions within firms are not recorded. From dropping by a colleague's office to ask for help finding a reference, transferring a shaped piece of glass down the assembly line for installation into a mirror, getting an order of fries from the fry cook to deliver to the customer, most economic transactions are difficult even to define as such, let alone track. But we do have some numbers. In what I think is one of the best opening sentences of a job-market paper, Pol Antras provides a lower bound: "Roughly one-third of world trade is intrafirm trade."

Of course, it could conceivably be the case that boundaries don't really matter–that the nature of a particular transaction and the overall volume of transactions is the same whether boundaries are in place or not. And indeed, this would exactly be the case if there were no costs of carrying out transactions: Coase (1960)'s famous theorem suggests, roughly, that in such a situation, outcomes would be the same no matter how transactions were organized. But clearly this is not the case–in 1997, to pick a random year, the volume of corporate mergers and acquisitions was $1.7 trillion dollars (Holmstrom and Roberts, 1998). It is wholly implausible that this would be the case if boundaries were irrelevant, as even the associated legal fees have to ring up in the billions of dollars.

And so, in a sense, the premise of the Coase Theorem's contrapositive is clearly true. Therefore, there must be transaction costs. And understanding the nature of these transaction costs will hopefully shed some light on the patterns we see. And as D.H. Robertson also vividly illustrated, there are indeed patterns to what we see. Firms are "islands of conscious power in this ocean of unconscious co-operation like lumps of butter coagulating in a pail of buttermilk." So the question becomes: what transaction costs are important, and how are they important? How, in a sense, can they help make sense out of the pattern of butter and buttermilk?

The field was basically dormant for the next forty years until the early 1970s, largely because "transaction costs" came to represent essentially "a name for the residual"–any pattern in the data could trivially be attributed to some story about transaction costs. The empirical content of the theory was therefore zero. That is, until Oliver Williamson arrived on the scene.

Williamson put structure on the theory by identifying specific factors that composed these transaction costs. And importantly, the specific factors he identified had implications about economic objects that at least could, in principle, be contained in a data set. Therefore his causal claims could be, and were, tested. (As a conceptual matter, it is important to note that even if Williamson's causal claims were refuted, this would not invalidate the underlying claim that "transaction costs are important," since as discussed earlier, this more general claim is essentially untestable, because it is impossible to measure, or even conceive of, *all* transaction costs associated with *all* different forms of organization.) The gist of his theory, which we will describe in more detail shortly, is that when contracts are incomplete and parties have disagreements, they may waste resources "haggling" over the appropriate course of action if they transact in a market, whereas if they transact within a firm, these disagreements can be settled by "fiat" by a mediator. Integration is therefore more appealing when haggling costs are higher, which is the case in situations in which contracts are relatively more incomplete and parties disagree more.

But there was a sense in which his theory (and the related work by Klein, Crawford, and Alchian (1978)) was silent on many foundational questions. After all, why does moving the transaction from the market into the firm imply that parties no longer haggle–that is, what is integration? Further, if settling transactions by fiat is more efficient than by haggling, why aren't all transactions carried out within a single firm? Williamson's and others' response was that there are bureaucratic costs ("accounting contrivances," "weakened incentives," and others) associated with putting more transactions within the firm. But surely those costs are also higher when contracts are more incomplete and when there is more disagreement between parties. Put differently, Williamson identified particular costs associated with transacting in the market and other costs associated with transacting within the firm and made assertions about the rates at which these costs vary with the underlying environment. The resulting empirical implications were consistent with evidence, but the theory still lacked convincing foundations, because it treated these latter costs as essentially exogenous and orthogonal. We will discuss the Transaction-Cost Economics (TCE) approach in the first subsection.

The Property Rights Theory, initiated by Grossman and Hart (1986) and expanded upon in Hart and Moore (1990), proposed a theory which (a) explicitly answered the question of "what is integration?" and (b) treated the costs and benefits of integration symmetrically. Related to the first point is an observation by Alchian and Demsetz that

> It is common to see the firm characterized by the power to settle issues by fiat, by authority, or by disciplinary action superior to that available in the conventional market. This is delusion. The firm does not own all its inputs. It has no power of fiat, no authority, no disciplinary action any different in the slightest degree from ordinary market contracting between any two people. I can "punish" you only by withholding future business or by seeking redress in the courts for any

> failure to honor our exchange agreement. This is exactly all that any employer can do. He can fire or sue, just as I can fire my grocer by stopping purchases from him or sue him for delivering faulty products.

What, then, is the difference between me "telling my grocer what to do" and me "telling my employee what to do"? In either case, refusal would potentially cause the relationship to break down. The key difference, according to Grossman and Hart's theory, is in what happens after the relationship breaks down. If I stop buying goods from my grocer, I no longer have access to his store and all its associated benefits. He simply loses access to a particular customer. If I stop employing a worker, on the other hand, the worker loses access to all the assets associated with my firm. I simply lose access to that particular worker.

Grossman and Hart (1986)'s key insight is that property rights determine who can do what in the event that a relationship breaks down–property rights determine what they refer to as the residual rights of control. And allocating these property rights to one party or another may change their incentives to take actions that affect the value of this particular relationship. This logic leads to what is often interpreted as Grossman and Hart's main result: property rights (which define whether a particular transaction is carried out "within" a firm or "between" firms) should be allocated to whichever party is responsible for making more important investments in the relationship. We will discuss the Property Rights Theory (PRT) approach in the second subsection.

From a theoretical foundations perspective, Grossman and Hart was a huge step forward– their theory treats the costs of integration and the costs of non-integration symmetrically and systematically analyzes how different factors drive these two costs in a single unified framework. From a conceptual perspective, however, all the action in the theory is related to how organization affects parties' incentives to make relationship-specific investments. As we will see, their theory assumes that conditional on relationship-specific investments, transactions are always carried out efficiently. A manager never wastes time and resources arguing with an employee. An employee never wastes time and resources trying to convince the boss to let him do a different, more desirable task.

In contrast, in Transaction-Cost Economics, all the action takes place ex-post, during the time in which decisions are made. Integration is chosen, precisely because it avoids inefficient haggling costs. We will look at two implications of this observation in the context of two models. The first, which we will examine in the third subsection, will be the adaptation model of Tadelis and Williamson. The second, which we will examine in the fourth subsection, will be a model based on influence activities. In the fifth subsection, we will look at a model in which asset ownership conveys payoff rights, rather than control rights.

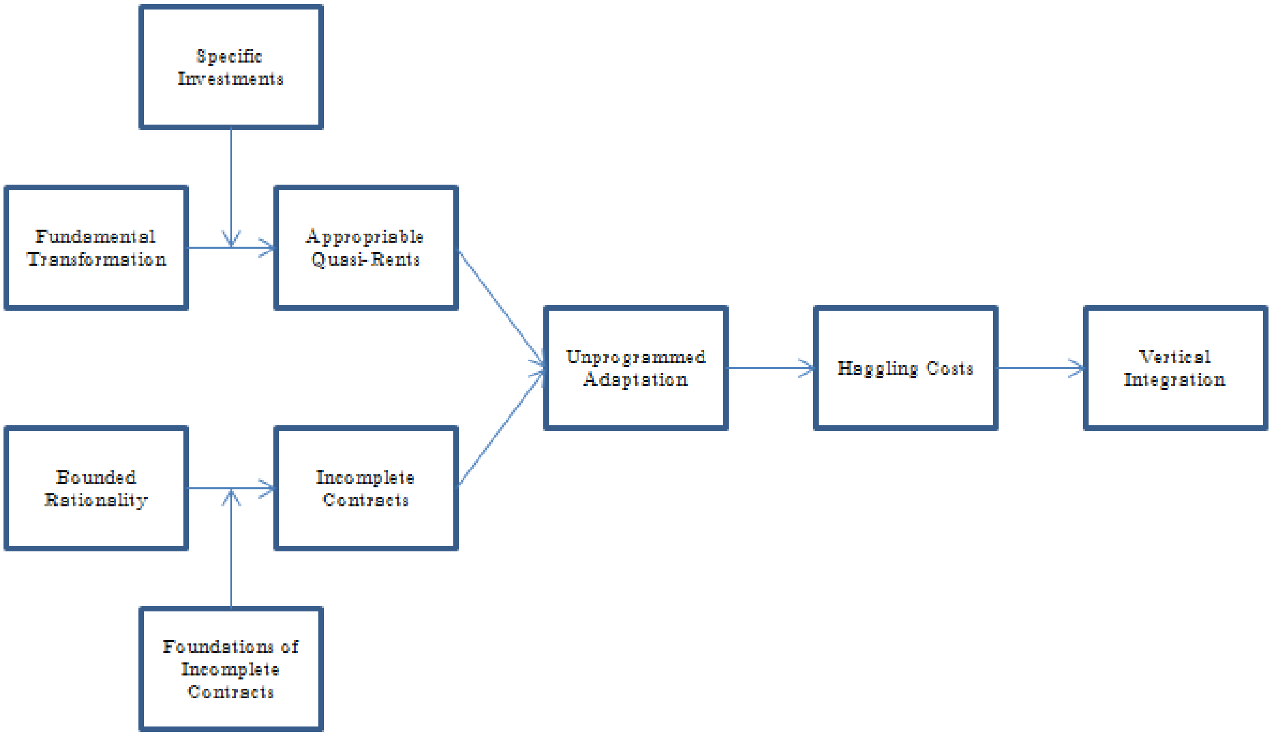
Finally, even the Property Rights Theory does not stand on firm theoretical grounds, since the theory considers only a limited set of institutions the players can put in place to manage their relationship. That is, they focus only on the allocation of control, ignoring the possibility that individuals may write contracts or put in place other types of mechanisms that could potentially do better. In particular, they rule out revelation mechanisms that, in principle, should induce first-best investment. We will address this issue in the sixth subsection.

As a research topic, the theory of the firm can be a bit overwhelming. In contrast to many applied-theory topics, the theory of the firm takes a lot of non-standard variables as endogenous. Further, there is often ambiguity about what variables should be taken as endogenous and what methodology should be used, so the "playing field" is not well-specified. But ultimately, I think that developing a stronger understanding of what determines firm boundaries is important, since it simultaneously tells us what the limitations of markets are. I will try to outline some of the "ground rules" that I have been able to discern from spending some time studying these issues.

4.1 Transaction-Cost Economics (Updated: April 21 2014)

The literature on the boundaries of the firm introduces many new concepts and even more new terms. So we will spend a little bit of time sorting out the terminology before proceeding.

The following figure introduces most of the new terms that we will talk about in this section and in the following sections.



As an overview, the basic argument of the Transaction-Cost Economics approach is the following.

Consider a transaction between an upstream manufacturer and a downstream distributor. Should the distributor **buy** from the manufacturer or should it buy the manufacturer and **make** goods itself? The **fundamental transformation** of ex-ante perfect competition among manufacturers for the customer's business to ex-post small numbers results *from* **specific investments** and results *in* **appropriable quasi-rents**–ex-post rents that parties can potentially fight over, because they are locked in with each other. If the parties have written **incomplete contracts** (the **foundations** for which are that **bounded rationality** limits their ability to foresee all relevant contingencies), then they might find themselves in situations that call for **unprogrammed adaptation**. At this point, they may fight over the appropriate course of action, incurring **haggling costs**. These haggling costs can be reduced or eliminated if either the manufacturer purchases the distributor or the distributor purchases the manufacturer, and they become **vertically integrated**.

It is worth discussing essentially every link in this figure. The **fundamental transformation** is, in my view, the most important economic idea to emerge from this theory. We had known since at least Edgeworth that under bilateral monopoly, many problems were possible (Edgeworth focused on indeterminacy) and perhaps inevitable (e.g., the Myerson-Satterthwaite theorem), but that competition among large numbers of potential trading partners would generally (with some exceptions–perfect complementarities between, say, right-shoe manufacturers and left-shoe manufacturers could persist even if the economy became arbitrarily large) push economies towards efficient allocations. Perfect competition is equivalent to the no-surplus condition (Ostroy, 1980; Makowski and Ostroy, 1995)–under perfect competition, you fully appropriate whatever surplus you generate, so everyone else in the economy as a whole is indifferent toward what you do. As a result, your incentives to maximize your own well-being do not come into conflict with others, so this leads to efficient allocations and nothing worth incurring costs to fight over. The underlying intuition for why large numbers of trading partners leads to efficient allocations is that a buyer can always play a seller and her competitors off each other, and in the limit, the next-best seller is just as good as the current one (and symmetrically for buyers). Williamson's observation was that after a trading relationship has been initiated, the buyer and the sellers develop ties to each other (**quasi-rents**), so that one's current trading partner is always discretely better than the next best alternative. In other words, the beneficial forces of perfect competition almost never hold.

Of course, if during the ex-ante competition phase of the relationship, potential trading partners competed with each other by offering enforceable, complete, long-term trading contracts, then the fact that ex-post, parties are locked in to each other would be irrelevant. Parties would compete in the market by offering each other utility streams that they are contractually obligated to fulfill, and perfect competition ex-ante would lead to maximized long-term gains from trade.

This is where **incomplete contracts** comes into the picture. Such contracts are impossible to write, because they would require parties to be able to conceive of and enumerate all possible contingencies. Because parties are **boundedly rational**, they will only be able to do so for a subset of the possible states. As a result, ex-ante competition will lead parties to agree to incomplete contracts for which the parties will need to fill in the details as they go along. In other words, they will occasionally need to make **unprogrammed adaptations**. As an example, a legacy airline (say, AA) and a regional carrier may agree on a flight schedule for flights connecting two cities. But when weather-related disruptions occur, the ideal way of shifting around staff and equipment depends to a large extent on where both carrier's existing staff and equipment are, and there are simply too many different potential configurations for this. As a result, airlines typically do not contractually specify what will happen in the event that there are weather-related disruptions, and they therefore have to figure it out on the spot.

The need to make unprogrammed adaptations would also not be a problem if the parties could simply agree to bargain efficiently ex-post after an event occurs that the parties had not planned for. (And indeed, if there was perfect competition ex-post, they would not even need to bargain ex-post.) However, under the TCE view, ex-post bargaining is rarely if ever efficient. The legacy airline will insist that its own staff and equipment are unable to make it, so everything would be better if the regional carrier made concessions, and conversely. Such ex-post bargaining inevitably leads either to bad ex-post decisions (the carrier with the easier-to-access equipment and staff is not the one who ends up putting it in place) or results in other types of **rent-seeking** costs (time and resources are wasted in the bargaining process). These **haggling costs** could be eliminated if both parties were under the direction of a common headquarters that could issue commands and easily resolve these types of conflicts. This involves setting up a **vertically integrated** organization.

Further, vertically integrated organizations involve **bureaucratic costs**. Reorganization involves setup costs. Incentives are usually low-powered inside organizations. Division managers engage in **accounting contrivances** in order to alter decision making of other divisions or the headquarters. Finally, the contract law that governs decisions made by one division that affect another division differs from the contract law that governs decisions made by one firm that affect another–in essence, the latter types of contracts are enforceable, whereas the former types of contracts are not. This difference in contract law is referred to as **forbearance**.

When would we be more likely to see vertical integration? When the environment surrounding a particular transaction is especially complex, contracts are more likely to be incomplete, or they are likely to be more incomplete. As a result, the need for unprogrammed adaptations and their associated haggling costs will be greater. When parties are more locked in to each other, their ability to access the outside market either to look for alternatives or to use alternatives to discipline their bargaining process is lessened. As a result, there is more to fight over when unprogrammed adaptations are required, and their associated haggling costs will be greater. Additionally, integration involves setup costs, and these setup costs are only worth incurring if the parties expect to interact with each other often. Finally, integration itself involves other bureaucratic costs, and so vertical integration is more appealing if these costs are low. Put differently, the integration decision involves a trade-off between haggling costs under non-integration and bureaucratic costs under integration. To summarize, the main empirical predictions of the TCE theory are:

1. Vertical integration is more likely for transactions that are more complicated.

2. Vertical integration is more likely when there is more specificity.

3. Vertical integration is more likely when the players interact more frequently.

4. Vertical integration is more likely when bureaucratic costs are low.

This discussion of TCE has been informal, because the theory itself is informal. There are at least two aspects of this informal argument that can be independently formalized.

When unprogrammed adaptation is required, the associated costs can either come from costly haggling (rent-seeking) or from inefficient ex-post decision making (adaptation). We now describe a model that captures this latter source of ex-post inefficiency. After we discuss the property rights theory, we will return to the idea that the source of ex-post costs are rent-seeking costs.

- Coase meets Heckman
- "private ordering" - resolving conflict need not involve government intervention
- Hayek vs Barnard - adaptation facilitated by market mechanism vs adaptation facilitated by entrepreneur
- Differences between NI and I:
  - Incentive intensity
  - Administrative control
  - Contract law regime
- More than just NI and I:
  - Spot market transactions
  - Short-run contracting across firm boundaries
  - Long-term contracting across firm boundaries
  - Integration

Description There is a risk-neutral upstream manufacturer U of an intermediate good and a risk-neutral downstream producer, D, who can costlessly transform a unit of the intermediate good into a final good that is then sold into the market at price p. Production of the intermediate good involves a cost of $C(e, g) = \bar{C} - eg \in \mathcal{C}$, where e is an effort choice by

U and involves a private cost of $c(e) = \frac{c}{2}e^2$ being borne by U. $g \in G$ denotes the governance structure, which we will describe shortly. There is a state of the world $\theta \in \Theta = \Theta_C \cup \Theta_{NC}$, with $\Theta_C \cap \Theta_{NC} = \emptyset$, where $\theta \in \Theta_C$ is a contractible state and $\theta \in \Theta_{NC}$ is a noncontractible state. The parties can either be integrated ($g = I$) or non-integrated ($g = NI$), and they can also sign a contract $w \in W = \left\{ w : \mathcal{C} \times \times \to \{s + (1-b)\, C\}_{b \in \{0,1\}} \right\}$, which compels D to make a transfer of $s + (1-b)\, C$ to U in any state $\theta \in \Theta$. Additionally, if $\theta \in \Theta_{NC}$, the contract has to be renegotiated. In this case, D incurs **adaptation costs** of $k(b, g)$, which depends on whether or not the parties are integrated as well as on the cost-sharing characteristics of the contract. The contract is always successfully renegotiated so that trade still occurs, and the same cost-sharing rule as specified in the original contract is obtained. The probability that adaptation is required is $\Pr[\theta \in \Theta_{NC}] = \sigma$.

Timing The timing is as follwos:

1. D makes an offer of a governance structure g and a contract w to U. (g, w) is publicly observed.

2. U can accept the contract ($d = 1$) or reject it ($d = 0$) in favor of an outside option that yields utility 0.

3. If $d = 1$, then U chooses effort e at cost $c(e) = \frac{c}{2}e^2$. e is commonly observed.

4. $\theta \in \Theta$ is realized and is commonly observed.

5. If $\theta \in \Theta_{NC}$, parties have to adjust the contract, in which case D incurs adaptation costs $k(b, g)$. Trade occurs, and the final good is sold at price p.

Equilibrium A **subgame-perfect equilibrium** is a governance structure g^*, a contract w^*, an acceptance decision strategy $d^* : G \times W \to \{0, 1\}$, and an effort choice strategy $e^* : G \times W \times D \to \mathbb{R}_+$ such that given g^* and w^*, U optimally chooses $d^*(g^*, w^*)$ and $e^*(g^*, w^*, d^*)$, and D optimally offers governance structure g^* and contract w^*.

The Program The downstream producer makes an offer of a governance structure g and a contract $w = s + (1-b)\,C$ as well as a proposed effort level e to maximize his profits:

$$\max_{g\in\{I,NI\},b\in\{0,1\},e,s} p - s - (1-b)\,C\,(e,g) - \sigma k\,(b,g)$$

subject to U's incentive-compatibility constraint

$$e \in \underset{\hat{e}}{\operatorname{argmax}}\, s + (1-b)\,C\,(\hat{e},g) - C\,(\hat{e},g) - c\,(\hat{e})$$

and her individual-rationality constraint

$$s + (1-b)\,C\,(e,g) - C\,(e,g) - c\,(e) \geq 0.$$

Since we are restricting attention to linear contracts, U's incentive-compatibility constraint can be replaced by her first-order condition:

$$\begin{aligned} c'\,(e\,(b,g)) &= -b\frac{\partial C\,(e,g)}{\partial e} \\ e\,(b,g) &= \frac{b}{c}g, \end{aligned}$$

and any optimal contract offer by D will ensure that U's individual-rationality constraint holds with equality. D's problem then becomes

$$\max_{g\in\{I,NI\},b\in\{0,1\}} p - C\,(e\,(b,g)\,,g) - c\,(e\,(b,g)) - \sigma k\,(b,g)\,.$$

That is, D chooses a governance structure $g \in \{I, NI\}$ and an incentive intensity $b \in \{0,1\}$ in order to maximize total ex-ante expected equilibrium surplus. Let

$$W\,(g,b;\sigma) = p - C\,(e\,(b,g)\,,g) - c\,(e\,(b,g)) - \sigma k\,(b,g)\,.$$

We will refer to the following problem as the **Coasian Program**, the solution to which is the optimal governance structure (g^*, b^*):

$$W^*(\sigma) = \max_{g \in \{I, NI\}, b \in \{0,1\}} W(g, b; \sigma).$$

Assumptions Several assumptions will be important for the main results of this model:

1. Supplier effort is more effective under non-integration than under integration (i.e., $\frac{\partial C(e,NI)}{\partial e} < \frac{\partial C(e,I)}{\partial e}$)

2. Adaptation costs are lower under integration than under non-integration (i.e., $k(b, NI) > k(b, I)$)

3. Adaptation costs are lower when cost incentives are weaker (i.e., $\frac{\partial k(b,g)}{\partial b} > 0$)

4. Reducing adaptation costs by weakening incentives is more effective under integration than under non-integration (i.e., $\frac{\partial k(b,NI)}{\partial b} > \frac{\partial k(b,I)}{\partial b}$).

Tadelis and Williamson (2013) outline many ways to justify several of these assumptions, but at the end of the day, these assumptions are quite reduced-form. However, they map nicely into the main results of the model, so at the very least, we can get a clear picture of what a more structured model ought to satisfy in order to get these results.

Solution To solve this model, we will use some straightforward monotone comparative statics results. Recall that if $F(x, \theta)$ is a function of choice variables $x \in X$ and parameters $\theta \in \Theta$, then if $F(x, \theta)$ is supermodular in (x, θ), $x^*(\theta)$ is increasing in θ, where

$$x^*(\theta) = \operatorname*{argmax}_{x \in X} F(x, \theta).$$

Once the Coasian program has been expressed as an unconstrained maximization problem, the key comparative statics are very easy to obtain if the objective function is supermodular. This model's assumptions are purposefully made in order to ensure that the objective function is supermodular.

To see this, let

$$\begin{aligned} W(g,b;\sigma) &= p - C(e(b,g),g) - c(e(b,g)) - \sigma k(b,g) \\ &= p - \bar{C} + e(b,g)\cdot g - \frac{c}{2}e(b,g)^2 - \sigma k(b,g) \end{aligned}$$

We can easily check supermodularity by taking some first-order derivatives and looking at second-order differences:

$$\begin{aligned} \frac{\partial W}{\partial b} &= \frac{\partial e(b,g)}{\partial b}\cdot g - ce(b,g)\frac{\partial e(b,g)}{\partial b} - \sigma\frac{\partial k(b,g)}{\partial b} \\ &= (1-b)\frac{g^2}{c} - \sigma\frac{\partial k(b,g)}{\partial b} \\ \frac{\partial W}{\partial \sigma} &= -k(b,g) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial^2 W}{\partial b\partial\sigma} &= -\frac{\partial k(b,g)}{\partial b} \\ \frac{\partial W(NI,b;\sigma)}{\partial b} - \frac{\partial W(I,b;\sigma)}{\partial b} &= \left((1-b)\frac{(NI)^2}{c} - \sigma\frac{\partial k(b,NI)}{\partial b}\right) - \left((1-b)\frac{I^2}{c} - \sigma\frac{\partial k(b,I)}{\partial b}\right) \\ &= \frac{1-b}{c}\left((NI)^2 - I^2\right) + \sigma\left(\frac{\partial k(b,I)}{\partial b} - \frac{\partial k(b,NI)}{\partial b}\right) \\ \frac{\partial W(NI,b;\sigma)}{\partial \sigma} - \frac{\partial W(I,b;\sigma)}{\partial \sigma} &= k(b,I) - k(b,NI). \end{aligned}$$

By assumption 1, $I < NI$. By assumption 2, $\frac{\partial W(NI,b;\sigma)}{\partial\sigma} - \frac{\partial W(I,b;\sigma)}{\partial\sigma} < 0$. By assumption 3, $\frac{\partial^2 W}{\partial b\partial\sigma} < 0$. By assumptions 1 and 4, $\frac{\partial W(NI,b;\sigma)}{\partial b} - \frac{\partial W(I,b;\sigma)}{\partial b} > 0$. Putting these together, we have that $W(g,b;\sigma)$ is supermodular in $(g,b;-\sigma)$, where we adopt the order that $g = NI$ is greater than $g = I$. We then have some straightforward comparative statics:

1. g^* is increasing in $-\sigma$. That is, when there is more uncertainty, NI becomes more

desirable relative to I.

2. b^* is increasing in $-\sigma$. That is, when there is more uncertainty, low-powered incentives become relatively more desirable.

3. (g^*, b^*) are complementary. Incentives are more high-powered under non-integration than under integration.

4.2 Property Rights (Updated: April 19 2014)

Essentially the main result of TCE is the observation that when haggling costs are high under non-integration, then integration is optimal. This result is unsatisfying in at least two senses. First, TCE does not tell us what exactly is the mechanism through which haggling costs are reduced under integration, and second, it does not tell us what the associated costs of integration are, and it therefore does not tell us when we would expect such costs to be high. In principle, in environments in which haggling costs are high under non-integration, then the within-firm equivalent of haggling costs should also be high.

Grossman and Hart (1986) and Hart and Moore (1990) set aside the "make or buy" question and instead begin with the more fundamental question, "What is a firm?" In some sense, nothing short of an answer to *this* question will consistently provide an answer to the questions that TCE leaves unanswered. Framing the question slightly different, what do I get if I buy a firm from someone else? The answer is typically that I become the owner of the firm's non-human assets.

Why, though, does it matter who owns non-human assets? If contracts are complete, it does not matter. The parties to a transaction will, ex-ante, specify a detailed action plan. One such action plan will be optimal. That action plan will be optimal regardless of who owns the assets that support the transaction, and it will be feasible regardless of who owns the assets. If contracts are incomplete, however, not all contingencies will be specified. The key insight of the PRT is that ownership endows the asset's owner with the right to decide what to do with the assets in these contingencies. That is, ownership confers **residual control rights**. When unprogrammed adaptations become necessary, the party with residual control rights has **power** in the relationship and is protected from expropriation by the other party. That is, control over non-human assets leads to control over human assets, since they provide leverage over the person who lacks the assets. Since she cannot be expropriated, she therefore has incentives to make investments that are specific to the relationship.

Firm boundaries are tantamount to asset ownership, so detailing the costs and benefits of different ownership arrangements provides a complete account of the costs and benefits of different firm-boundary arrangements. Asset ownership, and therefore firm boundaries, determine who possesses power in a relationship, and power determines investment incentives. Under integration, I have all the residual control rights over non-human assets and therefore possess strong investment incentives. Non-integration splits apart residual control rights, and therefore provides me with weaker investment incentives and you with stronger investment incentives. If I own an asset, you do not. Power is scarce and therefore should be allocated optimally.

Methodoligically, the PRT makes significant advances over the preceding theory. PRT's conceptual exercise is to hold technology, preferences, information, and the legal environment constant across prospective governance structures and ask, for a given transaction with given characteristics, whether the transaction is best carried out within a firm or between firms. That is, prior theories associated "make" with some vector $(\alpha_1, \alpha_2, \dots)$ of characteristics and "buy" with some other vector $(\beta_1, \beta_2, \dots)$ of characteristics. "Make" is preferred to "buy" if the vector $(\alpha_1, \alpha_2, \dots)$ is preferred to the vector $(\beta_1, \beta_2, \dots)$. In contrast, PRT focuses on a single aspect: α_1 versus β_1. Further differences may arise between "make" and "buy," but to the extent that they are also choice variables, they will arise optimally rather than passively. We will talk about why this is an important distinction to make when we talk about the influence-cost model in the next section.

Description There is a risk-neutral upstream manager U, a risk-neutral downstream manager D, and two assets A_1 and A_2. Managers U and D make investments e_U and e_D at private cost $c_U(e_U)$ and $c_D(e_D)$. These investments determine the value that each manager receives if trade occurs, $V_U(e_U, e_D)$ and $V_D(e_U, e_D)$. There is a state of the world, $s \in S = S_C \cup S_{NC}$, with $S_C \cap S_{NC} = \emptyset$ and $\Pr[s \in S_{NC}] = \mu$. In state s, the identity of the ideal good to be traded is s–if they trade good s, they receive $V_U(e_U, e_D)$ and $V_D(e_U, e_D)$. If the managers trade good $s' \neq s$, they both receive $-\infty$. The managers choose an asset allocation, denoted by g, from a set $G = \{UI, DI, NI, RNI\}$. Under $g = UI$, U owns both assets. Under $g = DI$, D owns both assets. Under $g = NI$, U owns asset A_1 and D owns asset A_2. Under $g = RNI$, D owns asset A_1, and U owns asset A_2. In addition to determining an asset allocation, manager U also offers an incomplete contract $w \in W = \{w : E_U \times E_D \times S_C \to \mathbb{R}\}$ to D. The contract specifies a transfer $w(e_U, e_D, s)$ to be paid from D to U if they trade good $s \in S_C$. If the players want to trade a good $s \in S_{NC}$, they do so in the following way. With probability $\frac{1}{2}$, U makes a take-it-or-leave-it offer $w_U(s)$ to D, specifying trade and a price. With probability $\frac{1}{2}$, D makes a take-it-or-leave-it offer $w_D(s)$ to U specifying trade and a price. If trade does not occur, then manager U receives payoff $v_U(e_U, e_D; g)$ and manager D receives payoff $v_D(e_U, e_D; g)$, which depends on the asset allocation.

Timing There are five periods:

1. U offers D an asset allocation $g \in G$ and a contract $w \in W$. Both g and w are commonly observed.

2. U and D simultaneously choose investment levels e_U and e_D at private cost $c(e_U)$ and $c(e_D)$. These investment levels are commonly observed by e_U and e_D.

3. The state of the world, $s \in S$ is realized.

4. If $s \in S_C$, D buys good s at price specified by w. If $s \in S_{NC}$, U and D engage in 50-50 take-it-or-leave-it bargaining.

5. Payoffs are realized.

Equilibrium A **subgame-perfect equilibrium** is an asset allocation g^*, a contract w^*, investment strategies $e_U^* : G \times W \to \mathbb{R}_+$ and $e_D^* : G \times W \to \mathbb{R}_+$, and a pair of offer rules $w_U^* : E_D \times E_U \times S_{NC} \to \mathbb{R}$ and $w_D^* : E_D \times E_U \times S_{NC} \to \mathbb{R}$ such that given $e_U^*(g^*, w^*)$ and $e_D^*(g^*, w^*)$, the managers optimally make offers $w_U^*(e_U^*, e_D^*)$ and $w_D^*(e_U^*, e_D^*)$ in states $s \in S_{NC}$; given g^* and w^*, managers optimally choose $e_U^*(g^*, w^*)$ and $e_D^*(g^*, w^*)$; and U optimally offers asset allocation g^* and contract w^*.

Assumptions As always, we will assume $c_U(e_U) = \frac{1}{2}e_U^2$ and $c_D(e_D) = \frac{1}{2}e_D^2$. We will also assume that $\mu = 1$, so that the probability that an ex-ante specifiable good is optimal to trade ex-post is zero. We will return to this issue later. Let

$$\begin{aligned}
V_U(e_U, e_D) &= f_{UU}e_U + f_{UD}e_D \\
V_D(e_U, e_D) &= f_{DU}e_U + f_{DD}e_D \\
v_U(e_U, e_D; g) &= h_{UU}^g e_U + h_{UD}^g e_D \\
v_D(e_U, e_D; g) &= h_{DU}^g e_U + h_{DD}^g e_D,
\end{aligned}$$

and define

$$\begin{aligned}
F_U &= f_{UU} + f_{DU} \\
F_D &= f_{UD} + f_{DD}.
\end{aligned}$$

Finally, outside options are more sensitive to one's own investments the more assets one owns:

$$\begin{aligned}
h_{UU}^{UI} &\geq h_{UU}^{NI} \geq h_{UU}^{DI}, h_{UU}^{UI} \geq h_{UU}^{RNI} \geq h_{UU}^{DI} \\
h_{DD}^{DI} &\geq h_{DD}^{NI} \geq h_{DD}^{UI}, h_{DD}^{DI} \geq h_{DD}^{RNI} \geq h_{DD}^{UI}.
\end{aligned}$$

The Program We solve backwards. For all $s \in S_{NC}$, with probability $\frac{1}{2}$, U will offer price $w_U(e_U, e_D)$. D will accept this offer as long as $V_D(e_U, e_D) - w_U(e_U, e_D) \geq v_D(e_U, e_D; g)$. U's offer will ensure that this holds with equality (or else U could increase w_U a bit and increase his profits while still having his offer accepted):

$$\begin{aligned}
\pi_U &= V_U(e_U, e_D) + w_U(e_U, e_D) = V_U(e_U, e_D) + V_D(e_U, e_D) - v_D(e_U, e_D; g) \\
\pi_D &= V_D(e_U, e_D) - w_U(e_U, e_D) = v_D(e_U, e_D; g).
\end{aligned}$$

Similarly, with probability $\frac{1}{2}$, D will offer price $w_D(e_U, e_D)$. U will accept this offer as long as $V_U(e_U, e_D) + w_D(e_U, e_D) \geq v_U(e_U, e_D; g)$. D's offer will ensure that this holds with equality (or else D could decrease w_D a bit and increase her profits while still having her offer accepted):

$$\begin{aligned}
\pi_U &= V_U(e_U, e_D) + w_D(e_U, e_D) = v_U(e_U, e_D; g) \\
\pi_D &= V_D(e_U, e_D) - w_D(e_U, e_D) = V_U(e_U, e_D) + V_D(e_U, e_D) - v_U(e_U, e_D; g).
\end{aligned}$$

In period 2, manager U will conjecture e_D and solve

$$\max_{\hat{e}_U} \frac{1}{2}\left(V_U(\hat{e}_U, e_D) + V_D(\hat{e}_U, e_D) - v_D(\hat{e}_U, e_D; g)\right) + \frac{1}{2} v_U(\hat{e}_U, e_D; g) - c(\hat{e}_U)$$

and manager D will conjecture e_U and solve

$$\max_{\hat{e}_D} \frac{1}{2} v_D(e_U, \hat{e}_D; g) + \frac{1}{2}\left(V_U(e_U, \hat{e}_D) + V_D(e_U, \hat{e}_D) - v_U(e_U, \hat{e}_D; g)\right) - c(\hat{e}_D).$$

Substituting in the functional forms we assumed above, these problems become:

$$\max_{\hat{e}_U} \frac{1}{2}(F_U \hat{e}_U + F_D e_D) + \frac{1}{2}\left(\left(h_{UU}^g - h_{DU}^g\right)\hat{e}_U + \left(h_{UD}^g - h_{DD}^g\right)e_D\right) - \frac{1}{2}\hat{e}_U^2$$

and

$$\max_{\hat{e}_D} \frac{1}{2}\left(F_U e_U + F_D \hat{e}_D\right) + \frac{1}{2}\left(\left(h_{DU}^g - h_{UU}^g\right) e_U + \left(h_{DD}^g - h_{UD}^g\right) \hat{e}_D\right) - \frac{1}{2}\hat{e}_D^2.$$

These are well-behaved objective functions, and in each one, there are no interactions between the managers' investments, so each manager has a dominant strategy, which we can solve for by taking first-order conditions:

$$\begin{aligned} e_U^{*g} &= \frac{1}{2}F_U + \frac{1}{2}\left(h_{UU}^g - h_{DU}^g\right) \\ e_D^{*g} &= \frac{1}{2}F_D + \frac{1}{2}\left(h_{DD}^g - h_{UD}^g\right) \end{aligned}$$

Each manager's incentives to invest are derived from two sources: (1) the marginal impact of investment on total surplus and (2) the marginal impact of investment on the "threat-point differential." The latter point is worth expanding on. If U increases his investment, his outside option goes up by h_{UU}^g, which increases the price that D will have to offer him when she makes her take-it-or-leave-it offer, which increases U's ex-post payoff if $h_{UU}^g > 0$. Further, D's outside option goes up by h_{DU}^g, which increases the price that U has to offer D when he makes his take-it-or-leave-it-offer, which decreases U's ex-post payoff if $h_{DU}^g > 0$.

Ex-ante, players' equilibrium payoffs are:

$$\begin{aligned} \Pi_U^{*g} &= \frac{1}{2}\left(F_U e_U^{*g} + F_D e_D^{*g}\right) + \frac{1}{2}\left(\left(h_{UU}^g - h_{DU}^g\right) e_U^{*g} + \left(h_{UD}^g - h_{DD}^g\right) e_D^{*g}\right) - \frac{1}{2}\left(e_U^{*g}\right)^2 \\ \Pi_D^{*g} &= \frac{1}{2}\left(F_U e_U^{*g} + F_D e_D^{*g}\right) + \frac{1}{2}\left(\left(h_{DU}^g - h_{UU}^g\right) e_U^{*g} + \left(h_{DD}^g - h_{UD}^g\right) e_D^{*g}\right) - \frac{1}{2}\left(e_D^{*g}\right)^2. \end{aligned}$$

If we let $\theta = \left(f_{UU}, f_{UD}, f_{DU}, f_{DD}, \{h_{UU}^g, h_{UD}^g, h_{DU}^g, h_{DD}^g\}_{g \in G}\right)$ denote the parameters of the model, total ex-ante equilibrium welfare for **governance structure** g is:

$$W^g(\theta) = \Pi_U^{*g} + \Pi_D^{*g} = F_U e_U^* + F_D e_D^* - \frac{1}{2}\left(e_U^{*g}\right)^2 - \frac{1}{2}\left(e_D^{*g}\right)^2.$$

The **Coasian Program** that describes the optimal governance structure is then:

$$W^*(\theta) = \max_{g \in G} W^g(\theta).$$

At this level of generality, the model is too rich to provide straightforward insights. In order to make progress, we will introduce the following definitions. If $f_{ij} = h_{ij}^g = 0$ for $i \neq j$, we say that investments are **self-investments**. If $f_{ii} = h_{ii}^g = 0$, we say that investments are **cross-investments**. When investments are self-investments, the following definitions are useful. Assets A_1 and A_2 are **independent** if $h_{UU}^{UI} = h_{UU}^{NI} = h_{UU}^{RNI}$ and $h_{DD}^{DI} = h_{DD}^{NI} = h_{DD}^{RNI}$ (i.e., if owning the second asset does not increase one's marginal incentives to invest beyond the incentives provided by owning a single asset). Assets A_1 and A_2 are **strictly complementary** if either $h_{UU}^{NI} = h_{UU}^{RNI} = h_{UU}^{DI}$ or $h_{DD}^{NI} = h_{DD}^{RNI} = h_{DD}^{UI}$ (i.e., if for one player, owning one asset provides the same incentives to invest as owning no assets). U's **human capital is essential** if $h_{DD}^{DI} = h_{DD}^{UI}$, and D's human capital is essential if $h_{UU}^{UI} = h_{UU}^{DI}$.

With these definitions in hand, we can get a sense for what features of the model drive the optimadescribe several aspects of the solution.

PROPOSITION (Hart 1995). If A_1 and A_2 are independent, then NI or RNI is optimal. If A_1 and A_2 are strictly complementary, then DI or UI is optimal. If U's human capital is essential, UI is optimal. If D's human capital is essential, DI is optimal. If both U's and D's human capital is essential, all governance structures are equally good.

These results are straightforward to prove. If A_1 and A_2 are independent, then there is no additional benefit of allocating a second asset to a single party. Dividing up the assets therefore strengthens one party's investment incentives without affecting the other's. If A_1 and A_2 are strictly complementary, then relative to integration, dividing up the assets necessarily weakens one party's investment incentives without increasing the other's, so one form of integration clearly dominates. If U's human capital is essential, then D's investment incentives are independent of which assets he owns, so UI is at least weakly optimal.

The more general results of this framework are that (a) allocating an asset to an individual strengthens that party's incentives to invest, since it increases his bargaining position when unprogrammed adaptation is required, (b) allocating an asset to one individual has an opportunity cost, since it means that it cannot be allocated to the other party. Since we have assumed that investment is always socially valuable, this implies that assets should always be allocated to exactly one party (if joint ownership means that both parties have a veto right). Further, allocating an asset to a particular party is more desirable the more important that party's investment is for joint welfare and the more sensitive his/her investment is to asset ownership. Finally, assets should be co-owned when there are complementarities between them.

While the actual results of the PRT model are sensible and intuitive, there are many limitations of the analysis. First, as Holmstrom points out in his 1999 JLEO article, "The problem is that the theory, as presented, really is a theory about asset ownership by individuals rather than by firms, at least if one interprets it literally. Assets are like bargaining chips in an entirely autocratic market... Individual ownership of assets does not offer a theory of organizational identities unless one associates individuals with firms." Holmstrom concludes that, "... the boundary question is in my view fundamentally about the distribution of activities: What do firms do rather than what do they own? Understanding asset configurations should not become an end in itself, but rather a means toward understanding activity configurations." That is, by taking payoff functions V_U and V_D as exogenous, the theory is abstracting from what Holmstrom views as the key issue of what a firm really is.

Second, after assets have been allocated and investments made, adaptation is made efficiently. The managers always reach an ex-post efficient arrangement in an efficient manner, and all inefficiencies arise ex-ante through inadequate incentives to make relationship-specific investments. Williamson (2000) argues that "The most consequential difference between the TCE and GHM setups is that the former holds that maladaptation in the contract execution interval is the principal source of inefficiency, whereas GHM vaporize ex post maladaptation by their assumptions of common knowledge and ex post bargaining." That is, Williamson believes that ex-post inefficiencies are the primary sources of inefficiencies that have to be managed by adjusting firm boundaries, while the PRT model focuses solely on ex-ante inefficiencies. The two approaches are obviously complementary, but there is an entire dimension of the problem that is being left untouched under this approach.

Finally, in the Coasian Program of the PRT model, the parties are unable to write formal contracts (in the above version of the model, this is true only when $\mu = 1$) and therefore the only instrument they have to motivate relationship-specific investments is the allocation of assets. The implicit assumption underlying the focus on asset ownership is that the characteristics defining what should be traded in which state of the world are difficult to write into a formal contract in a way that a third-party enforcer can unambiguously enforce. State-contingent trade is therefore unverifiable, so contracts written directly or indirectly on relationship-specific investments are infeasible. However, PRT assumes that relationship-specific investments, and therefore the value of different ex-post trades, are commonly observable to U and D. Further, U and D can correctly anticipate the payoff consequences of different asset allocations and different levels of investment. Under the assumptions that relationship-specific investments are commonly observable and that players can foresee the payoff consequences of their actions, Maskin and Tirole (1999) show that the players should always be able to construct a mechanism in which they truthfully reveal the payoffs they would receive to a third-party enforcer. If the parties are able to write a contract on these announcements, then they should indirectly be able to write a contract on ex-ante investments. This debate over the "foundations of incomplete contracting" mostly played out over the mid-to-late 1990s, but it has attracted some recent attention. We will discuss it in more detail later.

4.3 Influence Costs (Updated Apr 21, 2014)

- Unify Milgrom and Roberts (1988) on the costs of internal organization with Williamson (1975) on the costs of market exchange by asserting that the types of decisions that actors in separate firms argue about typically have analogues to the types of decisions that actors within the same firm argue about (e.g., prices versus transfer prices, trade credit versus capital allocation) and that there is no reason to think a priori that the ways in which they argue with each other differ across different governance structures. They may in fact argue in different ways, but this difference should be derived, not assumed.

- You can take away someone's right to make a decision. But you can't take away the fact that they care about that decision. And you can't take away their ability to try to influence whoever has the right to make that decision, at least not unless you are willing to incur additional costs

Description There are two risk-neutral managers, L and R, and two decisions, d_1 and d_2, that have to be made. The payoffs to the managers for a particular pair of decisions depend on an underlying state of the world, denoted by $s \in S$. The state of the world is unobserved; however, the two managers can potentially commonly observe an informative but manipulable signal, $\sigma \in \Sigma$. The two managers bargain over a **control structure** $g \in G = \{I_L, I_R, NI, RNI\}$, where under I_j, manager j controls both decisions; under NI, L controls d_1 and R controls d_2; and conversely under RNI. Additionally, the two managers bargain over an **organizational practice** $\theta \in \Theta = \{0, 1\}$. Under organizational practice $\theta = 0$, which I refer to as an **open-door policy**, the signal σ is commonly observed by the two managers, and under organizational practice $\theta = 1$, which I refer to as a **closed-door policy**, it is not. I refer to the bundle (g, θ) as a **governance structure**. I will assume that L makes an offer to R. This offer consists of a governance structure (g, θ) as well as a transfer $w \in \mathbb{R}$ to be paid to R. R can accept the offer or reject it in favor of an outside option that yields utility 0.

After the governance structure has been determined, each manager simultaneously chooses a level of **influence activities**, denoted by λ_i at private cost $k(\lambda) = \frac{1}{2}\lambda^2$. Influence activities are chosen prior to the observation of the public signal (if it is to be observed) and without any private knowledge of the state of the world, and they affect the conditional distribution of σ given s. After the signal has been observed (or not), the manager(s) with control of the decision rights immediately choose(s) a decision. THe managers cannot bargain over a signal-contingent decision rule ex-ante, and they cannot bargain ex-post over the decisions to be taken.

Timing The timing of the model is as follows:

1. L makes an offer of a governance structure $(g, \theta) \in G \times \Theta$ and a transfer $w \in \mathbb{R}$ to U. (g, θ) and w are publicly observed. R chooses whether to accept $(d = 1)$ or reject $(d = 0)$ this offer in favor of outside option yielding utility 0. $d \in D = \{0, 1\}$ is commonly observed.

2. L and R simultaneously choose influence activities $\lambda_L, \lambda_R \in \mathbb{R}$ at cost $k(\lambda)$; λ_i is privately observed by λ_i.

3. L and R publicly observe signal σ if and only if $\theta = 0$.

4. The manager with control of decision ℓ chooses $d_\ell \in \mathbb{R}$.

5. Payoffs are realized

Assumptions The signal is linear in the state of the world, influence activities, and a noise term. That is, $\sigma = s + \lambda_L + \lambda_R + \varepsilon$. All random variables are independent and normally distributed: $s \sim N(0, h^{-1})$ and $\varepsilon \sim N(0, h_\varepsilon^{-1})$. Each manager's payoffs gross of the costs of influence activities are

$$U_i(s,d) = \sum_{\ell=1}^{2} \left[-\frac{\alpha_i}{2}\left(d_\ell - s - \beta_i\right)^2 \right], \alpha_i > 0, \beta_i \in \mathbb{R}.$$

Both managers prefer that the decision is tailored to the state of the world, but conditional on the state of the world, manager i prefers that $d_1 = d_2 = s + \beta_i$, so there is disagreement between the two managers. Define $\Delta \equiv \beta_L - \beta_R > 0$ and assume that $\alpha_L \geq \alpha_R$.

In this model, $g = I_R$ will always be dominated by $g = I_L$, and $g = RNI$ will always yield exactly the same ex-ante equilibrium expected total payoffs as $g = NI$, so we will focus on only two prospective control structures. I will refer to divided control as **non-integration** (and denote it by $g = NI$) and unified control as **integration** (and denote it by $g = I$).

Equilibrium A **Perfect-Bayesian Equilibrium** of this game is a belief profile μ^*, an offer $(g^*, \theta^*), w^*$ of a governance structure and a transfer, a pair of influence-activity strategies $\lambda_L^* : G \times \Theta \times \mathbb{R} \times D \times \Delta(s) \to \mathbb{R}$ and $\lambda_R^* : G \times \Theta \times \mathbb{R} \times D \times \Delta(s) \to \mathbb{R}$, and a pair of decision rules $d_\ell^* : G \times \Theta \times \mathbb{R} \times D \times \mathbb{R} \times \Sigma \times \Delta(s) \to \mathbb{R}$ such that the influence-activity strategies and the decision rules are sequentially optimal for each player given his/her beliefs, and μ^* is derived from the equilibrium strategy using Bayes's rule whenever possible.

This model is a signal-jamming game, like the career concerns model earlier in the class. Further, the assumptions we have made will ensure that players want to choose relatively simple strategies. That is, they will choose public influence-activity strategies $\lambda_L^* : \Delta(s) \to \mathbb{R}$ and $\lambda_R^* : \Delta(s) \to \mathbb{R}$ and decision rules $d_\ell^* : G \times \Theta \times \Sigma \times \mathbb{R} \times \gtreqqless(s) \to \mathbb{R}$.

The Program Let us begin by solving for an equilibrium for an arbitrary governance structure (g, θ). Suppose manager i has control of decision ℓ under governance structure g. Let λ^* denote the equilibrium level of influence activities. Denote $\sigma(\theta) = \sigma$ if $\theta = 0$ and

$\sigma(\theta) = \emptyset$ if $\theta = 1$. Manager i will choose d_ℓ^* to minimize his expected loss given his beliefs:

$$\max_{d_\ell} E_s \left[-\frac{\alpha_i}{2} \left(d_\ell - s - \beta_i \right)^2 \middle| \sigma(\theta), \hat{\lambda}(i) \right],$$

where $\hat{\lambda}(i)$ denotes i's beliefs about the vector of influence activities. Since he faces a quadratic loss function, his decision will be equal to his conditional expectation of the state world, plus his bias term β_i:

$$d_\ell^{*g} \left(\sigma(\theta); \hat{\lambda}(i) \right) = E_s \left[s | \sigma(\theta), \hat{\lambda}(i) \right] + \beta_i.$$

The decision manager i chooses differs from the decision manager $j \neq i$ would choose for two reasons: first, $\beta_i \neq \beta_j$, so for a given set of beliefs, they prefer different decisions. Second, out of equilibrium, they may have different beliefs about λ. Manager i knows λ_i but only has a conjecture about λ_j.

Again, we can make use of the normal updating formula. The conditional expectation is a convex combination of the prior mean, which is 0, and a modified signal $\hat{s}(i) = \sigma(\theta) - \hat{\lambda}_L(i) - \hat{\lambda}_R(i)$. The weight that i's preferred decision rule attaches to the signal is given by the signal-to-noise ratio $\varphi = \frac{h_\varepsilon}{h+h_\varepsilon}$ times $(1-\theta)$, since if $\theta = 1$, the manager in fact does not observe a signal, and if $\theta = 0$, he does. The decision rule can be written as

$$d_\ell^{*g} \left(\sigma(\theta); \hat{\lambda}(i) \right) = \varphi \cdot \hat{s}(i).$$

Given decision rules $d_\ell^{*g}(\sigma(\theta); \lambda^*)$, we can now set up the program that the managers solve when deciding on the level of influence activities to engage in. Manager j chooses λ_j to solve

$$\max_{\lambda_j} E_{s,\varepsilon} \left[\sum_{\ell=1}^{2} -\frac{\alpha_j}{2} \left(d_\ell^{*g}(\sigma(\theta); \lambda^*) - s - \beta_j \right)^2 \right] - k(\lambda_j).$$

Taking first-order conditions, we get:

$$\begin{aligned}\left|k'\left(\lambda_j^*\right)\right| &= \left|E_{s,\varepsilon}\left[\sum_{\ell=1}^{2}-\alpha_j\underbrace{\left(d_\ell^{*g}\left(\sigma\left(\theta\right);\lambda^*\right)-s-\beta_j\right)}_{=0\text{ if } j \text{ controls } g;\ =\Delta \text{ otherwise}}\underbrace{\frac{\partial d_\ell^{*g}}{\partial\sigma}}_{\varphi(1-\theta)}\underbrace{\frac{\partial\sigma}{\partial\lambda_j^*}}_{=1}\right]\right| \\ &= \left(1-\theta\right)N_{\neg j}^{g}\alpha_j\Delta\varphi,\end{aligned}$$

where $N_{\neg j}^{g}$ is the number of decisions that manager j does not control under control structure g.

Finally, manager L will make an offer $(g,\theta), w$ to

$$\max_{(g,\theta),w} E_{s,\varepsilon}\left[\sum_{\ell=1}^{2}-\frac{\alpha_L}{2}\left(d_\ell^{*g}\left(\sigma\left(\theta\right);\lambda^*\right)-s-\beta_L\right)^2\right]-k\left(\lambda_L^*\right)-w$$

subject to manager R's individual-rationality constraint:

$$E_{s,\varepsilon}\left[\sum_{\ell=1}^{2}-\frac{\alpha_R}{2}\left(d_\ell^{*g}\left(\sigma\left(\theta\right);\lambda^*\right)-s-\beta_R\right)^2\right]-k\left(\lambda_R^*\right)+w\geq 0.$$

Again, w will be chosen so that the individual-rationality constraint holds with equality. Manager L's problem then becomes

$$\max_{(g,\theta)} E_{s,\varepsilon}\left[\sum_{i\in\{L,R\}}\sum_{\ell=1}^{2}-\frac{\alpha_i}{2}\left(d_\ell^{*g}\left(\sigma\left(\theta\right);\lambda^*\right)-s-\beta_i\right)^2\right]-\sum_{i\in\{L,R\}}k\left(\lambda_i^*\right).$$

Let χ denote a vector of parameters of the model, and let

$$W\left(g,\theta;\chi\right)=E_{s,\varepsilon}\left[\sum_{i\in\{L,R\}}\sum_{\ell=1}^{2}-\frac{\alpha_i}{2}\left(d_\ell^{*g}\left(\sigma\left(\theta\right);\lambda^*\right)-s-\beta_i\right)^2\right]-\sum_{i\in\{L,R\}}k\left(\lambda_i^*\right).$$

The **Coasian Program** is then:

$$\max_{(g,\theta)\in G\times\Theta} W\left(g,\theta;\chi\right).$$

4.3.1 Solution

Managers' payoff functions are quadratic, so ex-ante expected equilibrium total surplus, $W(g,\theta;\chi)$ can be decomposed using a mean-variance decomposition (recall that for two random variables X and Y, $Var(Y) = E[Var(Y|X)] + Var(E[Y|X])$):

$$W(g,\theta;\chi) = -\left(ADAP(\theta) + ALIGN(g) + INFL(g,\theta)\right).$$

$ADAP(\theta)$ represents the costs associated with basing decision making on a noisy signal. $ADAP(\theta)$ is higher for $\theta = 1$, because under $\theta = 1$, even the noisy signal is unavailable. $ALIGN(g)$ represents the costs associated with the fact that ex-post, decisions will always be made in a way that are not ideal for someone. Whether they are ideal for manager L or R depends on the control structure g. Finally, $INFL(g,\theta)$ are the influence costs, $k(\lambda_L^*) + k(\lambda_R^*)$. When $\theta = 1$, these costs will be 0, since there is no signal to manipulate. When $\theta = 0$, these costs will depend on the control structure.

There will be two trade-offs of interest.

Influence-cost/alignment-cost trade-off. First, let us ignore $ADAP(\theta)$ and look separately at $ALIGN(g)$ and $INFL(g,\theta)$. To do so, let us begin with $INFL(g,\theta)$. These can be written as:

$$INFL(g,\theta) = \frac{1}{2}\left((1-\theta)N_{\neg L}^{g}\alpha_L\Delta\varphi\right)^2 + \frac{1}{2}\left((1-\theta)N_{\neg R}^{g}\alpha_R\Delta\varphi\right)^2.$$

When $\theta = 1$, these are clearly 0. When $\theta = 0$, we can compute $INFL(I,0)$ and $INFL(NI,0)$ easily, since under $g = I$, $N_{\neg L}^{g} = 0$ and $N_{\neg R}^{g} = 2$, and under $g = NI$, they are both equal to 1.

$$\begin{aligned} INFL(I,0) &= \frac{1}{2}(2\alpha_R\Delta\varphi)^2 \\ INFL(NI,0) &= \frac{1}{2}(\alpha_L\Delta\varphi)^2 + \frac{1}{2}(\alpha_R\Delta\varphi)^2. \end{aligned}$$

Divided control minimizes influence costs, as long as $\alpha_R < \alpha_L < \sqrt{3}\alpha_R$:

$$INFL\,(I,0) - INFL\,(NI,0) = \frac{1}{2}\left(3\,(\alpha_R)^2 - (\alpha_L)^2\right)(\Delta\varphi)^2 > 0.$$

Next, let us look at $ALIGN\,(g)$. When $g = I$, manager L gets his/her ideal decisions on average, but manager R does not:

$$ALIGN\,(I) = \alpha_R \Delta^2.$$

When $g = NI$, each manager gets his/her ideal decision correct on average for one decision but not for the other decision:

$$ALIGN\,(NI) = \frac{\alpha_L + \alpha_R}{2}\Delta^2.$$

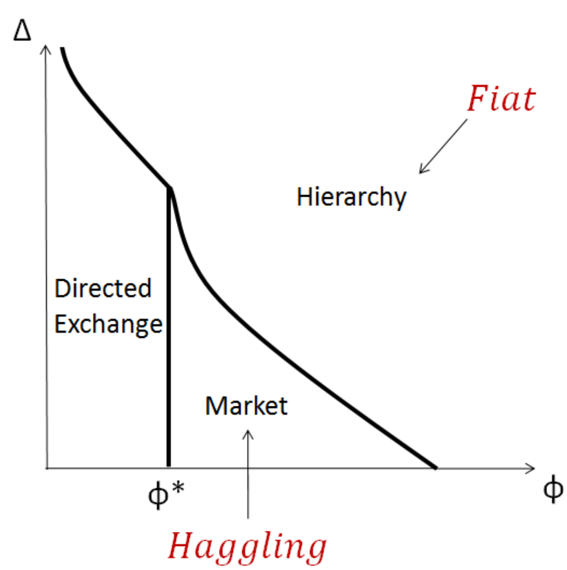
When $\alpha_L = \alpha_R$, so that $ALIGN\,(I) = ALIGN\,(NI)$, we have that $INFL\,(I,0) - INFL\,(NI,0) > 0$, so that influence costs are minimized under non-integration. When $\varphi = 0$, so that there are no influence costs, $ALIGN\,(I) < ALIGN\,(NI)$, since $\alpha_L > \alpha_R$, so that alignment costs are minimized under integration. Unified control reduces ex-post alignment costs and divided control reduces influence costs, and there is a trade-off between the two.

Influence-cost/adaptation-cost trade-off. Next, let us ignore $ALIGN\,(g)$ and look separately at $ADAP\,(\theta)$ and $INFL\,(g,\theta)$. It is not difficult to show that

$$ADAP\,(\theta) = \frac{\alpha_L + \alpha_R}{h + h_\varepsilon} + \theta\varphi\frac{\alpha_L + \alpha_R}{h}.$$

Adaptation costs are therefore higher when $\theta = 1$ (i.e., when the firm adopts a closed-door policy.) But when $\theta = 1$, influence costs are reduced to 0. Closed-door policies therefore eliminate influence costs but reduce the quality of decision making, so there is a trade-off here as well. Finally, it is worth noting that when $\theta = 1$, influence activities are eliminated, so the parties might as well unify control, since doing so reduces ex-post alignment costs. That is, closed-door policies and integration are complementary.

The next figure describes the optimal governance structure for different parts of the parameter space. On the horizontal axis is the signal-to-noise ratio, and on the vertical axis is the level of disagreement between the two parties. I refer to $(g, \theta) = (I, 0)$ as "directed exchange," $(g, \theta) = (I, 1)$ as "hierarchy," and $(g, \theta) = (NI, 0)$ as "market" to connect with the existing literature. The first observation is that $(g, \theta) = (NI, 0)$ is never optimal, for the reasons we just described. The second observation is that under "market" governance, the two managers will engage in positive levels of influence activities. I will interpret the associated costs as Williamson's "haggling costs." The third observation is that under "hierarchy" governance, the two managers will not engage in any influence activities, but the decision will also not be tailored to the signal, since they do not observe it. The costs associated with worsened decision making are referred to as the costs of "fiat."



4.4 Incentive Systems

Holmstrom and Milgrom (1994)

4.5 Evidence (Updated: April 21 2014)

4.5.1 Classic Evidence on VI (Updated: April 21 2014)

- Quick TCE recap:
  - Can B and S costlessly write complete contract ex-ante? Go with that, then.
  - Incomplete contracts but no ex-post lock-in? Wait, and then bargain or go with the alternative trader.
  - Incomplete contracts, ex-post lock-in, costless ex-post bargaining? Wait and then bargain.
  - Incomplete contracts, ex-post lock-in, costly ex-post bargaining? Perhaps integrate.
- Thus, need three things:
  - Incomplete contracts ex-ante
  - Lock-in ex-post (AQRs–something to fight *over*)
  - Costly ex-post bargaining (fighting $\Rightarrow$ lost surplus)
- Williamson (1971): "fiat is frequently a more efficient way to settle minor conflicts than is haggling."
  - Integration may reduce lost surplus from fighting
- Hypothesis: increase in interaction (of all three factors) increases relative benefits of integration

Operationalizing TCE

- Ideally, need transaction-specific measures of:

- Contractual incompleteness (how often will ex-post opportunism be a problem?)
- Ex-post lock-in (how much is there to fight over?)
- Costliness of ex-post bargaining
- Whether or not transaction is integrated

- Need consistent measures of these across transactions

Monteverde-Teece (Bell, 1982)

- One of the firmst attempts at taking TCE to data
- Industry contacts, survey data
- Automobile industry
  - Each car composed of thousands of parts
  - Significant variation in degree of integration across these parts
- Hypothesis:
  - Specialized, non-patentable know-how generates potential for opportunistic re-contracting
  - Ideal measure of non-patentable know-how in this setting: cost of developing a component
  - Actual measure of non-patentable know-how in this setting: estimate of engineering cost associated with a component.
- Data from GM and Ford in 1976, 133 components
  - An observation is at the (firm,component) level
  - "engineering cost" rated 1 (none) to 10 (a lot) by experts

- "specific" if procurer ID needed to find replacement part
- "integrated" if $> 80\%$ of component is produced in-house
- Control for component's "system"
- GM dummy

- Estimation and results
- Probit:

$$\Pr\left[VI_i = 1 | X_i\right] = F\left(X_i \beta\right)$$

  - Where does this probit come from?

- Results:
  - "Engineering cost" and "specific" significant
  - GM significant

- Concerns:
  - Measurement error?
  - What is "engineering cost" and how is it related to AQRs?
    * Lafontaine, Slade (2007) classify "engineering cost" as both specificity and complexity–which is it?
    * Could be that a "general-purpose" component could require lots of engineering effort
  - Are "engineering cost" and "specific" plausibly thought of as exogenous? Are in-house components likely to be un-"specific"?

  - Theory makes predictions about interactions, not individual terms.
    * What interactions should Monteverde-Teece include?
    * Contractual incompleteness in the background and constant across components?

- This was a first attempt at testing predictions of TCE
  - Largely a success: "engineering cost" & "specific" both significant
  - Influential: 1200 Google scholar cites
  - Provided a road map: pick a couple firms, use surveys to get at measures of specificity and uncertainty, estimate discrete-choice model
    * Masten (1984), Walker and Weber (1984), Anderson and Schmittlein (1984), Anderson (1985), Anderson and Coughlan (1987), Heidi and John (1988), Masten, Meehan, and Snyder (1991),...

Masten (JLE, 1984)

- Make-or-buy decision is about "differential efficiency of alternative organizational forms"
  - Framed the problem as one of "comparative-institutional analysis"
- Looks at make-or-buy decisions for a defense-related aerospace project
  - Specialized designs are common, investments are typically not redeployable
- Data on 1887 components, grouped into 34 (weighted) observations
- Measure of specificity–is component used (a) exclusively by this subcontractor, (b) by others in this industry, or (c) by other industries also?
- Co-location–is it important to be co-located?

- Complexity–internal ranking of complexity (scale: A, B, C)
- Costs of each governance structure:

$$
\begin{aligned}
C_{INT,i} &= \beta_0 + \beta_1 \cdot \text{complex}_i + \varepsilon_{INT,i} \\
C_{NI,i} &= \alpha_1 \cdot \text{complex}_i + \alpha_2 \cdot specific_i + \alpha_3 \text{coloc}_i + \varepsilon_{NI,i}
\end{aligned}
$$

  - β_0 – "administrative burden" of internal procurement
  - complex – complexity of part
  - speci – specificity of part
  - coloc – importance of colocation
  - $\varepsilon_{INT,i}, \varepsilon_{NI,i}$ – normally distributed error terms

- Costs are unobservable, so do probit:

$$
\begin{aligned}
\Pr\left[VI_i = 1 | X_i\right] &= \Pr\left[C_{INT,i} < C_{NI,i} | X_i\right] \\
&= \Pr\left[X_i\beta + \varepsilon_{INT,i} < X_i\alpha + \varepsilon_{NI,i} | X_i\right] \\
&= \Pr\left[\left.\frac{\varepsilon_{INT,i} - \varepsilon_{NI,i}}{\sigma} < X\frac{\alpha - \beta}{\sigma}\right| X_i\right] = F\left(X\gamma\right)
\end{aligned}
$$

  - Complexity, specificity significant
  - Co-location insignificant
  - Estimate of bureaucratic costs positive
  - Interaction important

Probability of Integration	Not Specific	Specific
Not Complex	$F\left(\hat{\beta}_0\right) < 0.01$	$F\left(\hat{\beta}_0 + \hat{\beta}_{Spec}\right) = 0.31$
Complex	$F\left(\hat{\beta}_0 + \hat{\beta}_{comp}\right) = 0.02$	$F\left(\hat{\beta}_0 + \hat{\beta}_{Comp} + \hat{\beta}_{Spec}\right) = 0.92$

- Concerns:
  - Interaction–why not just include the interaction term?
  - Identification
    * Orthogonal costs of bureaucracy?
    * Can only estimate $\gamma = (\alpha - \beta)/\sigma$
    * Is it plausible that the specificity of a component is orthogonal to the make-or-buy question? If I make something, it is less likely that others are going to use it.
  - Testing half a theory?
    * Empirical work focused on costs of market exchange (not of integration) as a function of AQRs
  - Are findings consistent with anything else?
    * "Knowledge-based theory of the firm" (e.g., Monteverde '95) imagines (no hold-up in markets but) hierarchy produces firm-specific language and routines (i.e., TCE getting it backwards: specificity irrelevant to market exchange but induced and managed within firms)

Masten, Meehan, and Snyder (JLEO, 1991)

- "The Costs of Organization"
- Structural "selection model" approach aimed at addressing the above concerns
  - What are the costs of choosing the wrong governance structure?
- Large Naval construction project
  - Information on managerial costs for integrated components

- Construction is very different than manufacturing, so we would expect different aspects to be important
  - Specialized assets not as important (pipe-welding equipment is relatively standardized and mobile)
  - No buffer inventories: timing and coordination critical (temporal specificity). "Where timely performance is critical, delay becomes a potentially effective strategy for exacting price concessions"
- Tasks that are more similar to existing capabilities cheaper to organize internally
  - This provides MMS with variation in the costs of internal organization
- Probit gives us estimates of $\gamma = (\alpha - \beta)/\sigma$
  - $\gamma > 0$ consistent with $\alpha > \beta > 0$ (TCE?) and $0 > \alpha > \beta$ (knowledge-based theory?)
- How to separately identify α and β?
- Approach here:
  - Have data on $C_{INT,i}$, use two-stage procedure:
    * Estimate selection equation (as in MT 82 and M84)
    * Correct for selection in cost regression
- Costs of production under integration:

$$C_{INT,i} = X\beta_X + Z\beta_Z + \varepsilon_{INT,i}$$

- Costs of production under non-integration

$$C_{NI,i} = X\alpha_X + Z\alpha_Z + \varepsilon_{NI,i}$$

- Probit for selection equation (analogous to the first stage of a two-stage regression):

$$\Pr\left[VI_i = 1 | X_i, Z_i\right] = F\left(X_i\gamma_X + Z_i\gamma_Z\right)$$

- Cost regression (second stage):

$$E\left[C_i | INT_i\right] = X_i\beta_X + Z_i\beta_Z + \rho\sigma_U\lambda\left(W_i\hat{\gamma}\right)$$

- Identifying restriction comes from $\alpha_Z = 0$
  - Need factor that affects costs under integration but not under non-integration
- First stage gives estimates of $\gamma = (\alpha - \beta)/\sigma$.
- Second stage gives estimates of β and σ. Can use these to back out α.
- Their findings are as follows
- First stage:
  - Temporal specificity, human capital, similarity of tasks significant
  - Asset specificity insignificant
  - Measure of complexity significant in "wrong" direction
- Second stage:
  - $\beta > 0$, consistent with TCE and not knowledge-based theory of the firm

- Costs associated with incorrect organization are substantial

- Identification off functional form (has to do with likelihood ratio in the tails)
- For structural estimation, should always think about identification in reduced-form model–where is variation coming from and how does it identify the coefficients of interest?
  - Want something that shifts propensity to integrate without directly affecting costs under integration
- Lots to like about this paper:
  - "Temporal specificity"
  - Distinguish between costs of integration and costs of non-integration
  - Recognizes selection problem and makes serious (1990's) attempt at identification
- More TCE evidence on VI:
  - Surveys: Shelanski and Klein (JLEO, 1995), Macher and Richman (Business and Politics, 2008), Lafontaine and Slade (JEL, 2007)
    * "TCE is an empirical success story..."
  - But see also: David and Han (SMJ, 2004), Carter and Hodgson (SMJ, 2006)
    * Have *all* aspects of TCE been confirmed? How often have main aspects been confirmed?
- Identification?
  - "So much smoke, there must be a fire somewhere."

4.5.2 Recent Evidence on Vertical Integration (Updated: April 21 2014)

- Takes identification more seriously
  - Instrumental variables?
  - Conditionally independent regressors?
- Examines theories other than TCE
  - PRT: Baker and Hubbard (AER), Woodruff (IJIO)
  - Incentive systems: Baker and Hubbard (QJE), Azoulay (AER)
  - Adaptation: Forbes and Lederman (AER)
- Examines effects in addition to determinants
  - Forbes and Lederman (RAND), Gil (JLEO)

Baker and Hubbard (AER 2003, QJE 2004)

- Explain governance in US trucking
- Key players
  - Shippers: Customers who want stuff shipped
  - Carriers: Companies who own trucks and carry stuff
  - Dispatchers: Find backhauls
  - Drivers

Baker and Hubbard (AER 2003)

- Make versus buy in trucking

- How does Walmart decide whether to ship via their own captive fleet or use Roadway Express?

- Three key elements:
  - Driver multi-tasking
  - Efficient dispatch
  - On-board computer technology
    * Trip recorder
    * Electronic Vehicle Monitoring System (EVMS)

- Multitasking: drivers do not just drive trucks
  - Service is often important: loading/unloading, waiting, going to different or additional locations

- When drivers are expected to perform service activities, it is harder to measure their performance
  - Agency costs are higher when drivers also perform non-driving services

- Dispatch: incomplete contracts about scheduling adaptations
  - For-hire fleet–if shipper has additional demands that come up, has to bargain with carrier
    * Carrier has incentive to improve bargaining position by making better investments in relevant (and irrelevant) backhauls–hold-up can be your friend (and your enemy)
  - Captive fleet–carrier has no outside option

        * Carrier has less incentive to invest in understanding general demand patterns when they will not improve bargaining position

- Service and dispatch interact:
    - Service = flexibility for the customer (willingness to wait/go somewhere else/make changes)
    - "This gives dispatchers a splitting headache"
- Model:
    1. Shipper and carrier bargain over asset ownership ($\delta = 0$ if captive fleet, $\delta = 1$ if carrier-owned) and service level (s, which is valued at m and results in agency costs $M(s, \sigma)$)
    2. Shipper searches for a "complementary haul," choosing a_1 (improves actual coordination) and a "substitute haul," choosing a_2 (improves outside option only)
    3. Shipper and carrier Nash bargain
- Given s, a_1, a_2, we have

$$
\begin{aligned}
V &= \underline{V} + (g_1 - \theta s) a_1 + ms - M(s, \sigma) \\
V_0^{\delta=1} &= (g_1 - \theta s) a_1 + g_2 a_2 + ms - M(s, \sigma) \\
V_0^{\delta=0} &= 0
\end{aligned}
$$

- Can derive $a_i(s, \delta)$. Total surplus:

$$W(s, \delta, m, \sigma, g_1, g_2)$$

- Supermodular in $(-s, \delta, -m, -\sigma, g_1, -g_2)$

- Empirical predictions: introduction of on-board computers depends on which type of on-board computers adopted
  - Trip recorder (TR): increase in σ (improve monitoring)
  - Electronic vehicle monitoring system (EVMS): increase σ and g_1 (improve monitoring and dispatch)

  1. $TR \uparrow \Rightarrow$ Captive Fleet $\uparrow$
  2. $EVMS \uparrow \Rightarrow$ Captive Fleet $\uparrow$ by less than if $TR \uparrow$
  3. $TR \uparrow \Rightarrow$ Captive Fleet $\uparrow$ more when service is important (say packaged goods to small outlets). No effect when service is unimportant ($m = 0$, say, bulk goods or raw materials)

- Results: Cohorts (state-product-trailer-distance combination) to get panel structure (can difference to control for time-invariant factors)
  - Three years: '87, '92, '97 of Truck Inventory and Use Surveys–information on on-board computers, ownership, type of hauls, distance

- Linear probability model (OLS)

  1. Cohorts with high adoption rates for TRs move towards captive fleet
  2. Cohorts with high adoption rates for EVMS move more towards carrier-owned than those with only TRs
  3. The effect of TRs on governance disappears for low-service hauls

- Unobserved third factors? Reverse causality?

- Instrumental variables–want factors that affect OBC adoption but do not directly affect changes in organizational form

  - Fraction of miles trucks are operated out of state (need OBC for state fuel tax purposes)
  - Regional dummies (more difficult for drivers to contact dispatchers in less densely populated areas)

- IV estimates are similar to OLS estimates, but noisy. The first stage is tricky, because instruments increase both TR and EVMS adoption.

- What theory are Baker and Hubbard testing?
  - PRT + job design
  - Asset ownership interacts with job design to determine optimal governance structure

- Effect of IT
  - Improved monitoring $\Rightarrow$ bigger firms
  - Improved coordination $\Rightarrow$ smaller firms
  - Identification: panel, instruments

Forbes and Lederman (AER, 2009)

- Tests an aspect of adaptation theory in US Airline industry
  - Great data (OAG–all fight schedules, RAA–ownership, BTS–flight-level performance, etc.)
  - Plausibly exogenous measure of need to make ex-post changes (unanticipated scheduling changes almost always driven by weather-related events–cannot contract on all contingencies ex-ante)
  - Variation in governance

- Majors–large network airlines (AA, United, etc.)
- Regionals–"subcontractors" on short- and medium-haul routes (American Eagle, Comair, etc.)
- Market: city-pair (e.g., ORD-BOS)
- Three-step decision fo major for each city pair
  1. Serve the market? (Ignored in this analysis)
  2. Fly own plane or subcontract to regional? (56% major-operated)
  3. Owned or independent regional? (26% owned regional, 19% independent regional)
- Decision to fly own plane or subcontract is usually determined by size of desired plane for the route.
- Costs of owning regional?
  - Regionals have lower labor costs (often first step toward getting a job at a major)
  - Unionization (Railway Labor Act): union agreements would apply to all employees
  - But: owned regionals are operated as subsidiaries, which seems to get around this
- Can easily contract on ex-ante schedules
- Cannot contract on ex-post changes
  - Possible schedule disruptions can be very complex
  - Disagreement on how to adapt to disruptions (key assumption–major has more to lose from maladaptation)
- Contingent control over schedules? Observed in practice

- But how to implement desired changes? Must still be carried out by regionals

- Key hypotheses:
  - Adverse weather more likely $\Rightarrow$ ex-post adaptations more likely $\Rightarrow$ owned regional
  - More interconnected network $\Rightarrow$ ex-post maladaptation more costly $\Rightarrow$ owned regional

- Observation is at the (major airline, city-pair) level

- Key variables:
  - Regional
  - Owned regional (governance)
  - Precipitation/snowfall (incompleteness)
  - Hub/number of flights out of each airport (scope for costly fighting)
  - Major dummies (identified off variation in ownership for particular major across city-pairs)

- Nested logit
  - "Top nest" – fly own plane or subcontract?
  - "Bottom nest" – owned or independent regional?

- Nothing interest in "top nest"

- "Bottom nest" (make-or-buy)
  - Precipitation/snowfall$\uparrow$ $\Rightarrow$ $\Pr[owned]\uparrow$
  - Hub/#flights $\uparrow$ $\Rightarrow$ $\Pr[owned]\uparrow$

- Very nice contribution:

  - Simple
  - Good data (many firms, governance structure measures, transaction-level characteristics)
  - Convincing proxy for need for ex-post adaptation (and weather is plausibly exogenous!)

- But:
  - Where is the interaction term?
  - Are we convinced by the story about costs of owning regional? Test using variation in strength of unions? (e.g., HQs in states with "right-to-work" laws?)
  - Hubbing and interconnectedness predetermined? (Hub locations are thought to be artifact of deregulation, but interconnectedness?)

Other recent evidence

- Forbes and Lederman (RAND, 2010)–how does owning a regional affect the performance of a major? Effects, rather than determinants, of integration.
- Atalay, Hortacsu, Syverson (AER, 2014)–within integrated firms, 2% of upstream plants are exclusively dedicated to their firms, but $\frac{1}{3}$ **make no internal shipments**, and $90th$ percentile of internal shippers send 42% outside. Is this consistent with TCE? Other theories?
- Novak and Stern (2009)–complementarities among transactions within a given firm
- Bresnahan and Levin (2011)–role of market structure and impact on integration

4.6 Foundations of Incomplete Contracts

Maskin and Tirole (1999)

Aghion, Fudenberg, Holden, Kunimoto, and Tercieux (2012)

Fehr, Powell, and Wilkening (2013)

5 Careers in Organizations

5.1 Internal Labor Markets

Baker, Gibbs, and Holmstrom (1994)

Gibbons and Waldman (1999)

5.2 Mandatory Retirement

Lazear (1979)

5.3 Efficiency Wages

Ke, Li, and Powell (2014)

6 Relational Incentive Contracts, Revisited

6.1 Imperfect Public Monitoring

Levin (2003)

6.2 Limited Transfers

Fong and Li (2013)

Li and Matouschek (2013)

6.3 Subjective Performance Measures

Fuchs (2007)

6.4 Persistent Private Information

Halac (2012)

6.5 Multiple Agents

Board (2011)

6.6 Imperfect Private Monitoring

Andrews and Barron (2014)

Barron and Powell (2014)

7 Persistent Performance Differences

7.1 Overview and Misallocation

Syverson (2011)

Hsieh and Klenow (2009)

7.2 Evidence and Management Practices

Bloom and Van Reenen (2007)

Bloom, Eifert, MacKenzie, Mahajan, Roberts (2013)

7.3 Theories

Chassang (2010)

Li, Matouschek, Powell (2014)

8 Organizations in Market Equilibrium

8.1 Information

Gibbons, Holden, and Powell (2012)

8.2 Price Levels

Legros and Newman (2013)

8.3 Competitive Rents

Powell (2013)