Contracting inside an organization: an experimental study*

Paul J. Healy, John O. Ledyard, Charles Noussair,
Harley Thronson, Peter Ulrich, and Giulio Varsi†

December 9, 2005

Abstract

In this paper we propose and test a contracting mechanism, Multi-Contract Cost Sharing (MCCS), for use in the management of a sequence of projects. The mechanism is intended for situations where (1) the contractor knows more about the true costs of various projects than does the contracting agency (adverse selection,) and (2) unobservable effort on the part of the contractor may lead to cost reductions (moral hazard.) The proposed process is evaluated in an experimental environment that includes the essential economic features of the NASA process for the acquisition of Space Science Strategy missions. The environment is complex and the optimal mechanism is unknown. The design of the MCCS mechanism is based on the optimal contract for a simpler related environment. We compare the performance of the proposed process to theoretical benchmarks and to an implementation of the current NASA 'cost cap' procurement process. The data indicate that the proposed MCCS process generates significantly higher value per dollar spent than using cost caps, because it allocates resources more efficiently among projects and provides greater incentives to engage in cost-reducing innovations.

1 Introduction

Many projects that provide a benefit to an entire organization are assigned to a specialized division for management while being funded through budgets at the headquarters level. Examples include the research division of a corporation, a team from a construction firm assigned to a building project, or a group of engineers and scientists assigned to develop a space mission. Often the division has better information about the eventual cost of the project than does headquarters. If the division's actions are not fully observable and if multiple divisions compete for the assignment of a project, the principal faces both moral hazard and adverse selection problems. Without a proper contract to mitigate these problems, inefficiencies will arise in the assignment of the project and in the effort level of the division.

*We thank Charles Holt, Susan Laury, Molly Macauley, Thomas Palfrey, Charles Plott, David Porter, Andrew Schotter, and an anonymous referee for comments. We thank the NASA personnel who participated in an experimental session conducted at NASA headquarters and who participated in interviews with us. We thank Steve Tucker for research assistance and James Braun for helpful suggestions. At the time the research was conducted, Noussair was a faculty member and Healy was an undergraduate student at Purdue University. We thank the NASA Office of Space Science and the Krannert School of Management at Purdue University for financial support.

†Healy: Tepper School of Business, Carnegie Mellon University, Pittsburgh, PA 15213, phealy@cmu.edu. Ledyard: Division of the Humanities and Social Sciences, California Institute of Technology, Pasadena, CA 91125, jledyard@hss.caltech.edu. Noussair: Department of Economics, Emory University, Atlanta, GA 30322-2240, cnoussa@emory.edu. Thronson, Ulrich and Varsi: Office of Space Science, NASA headquarters, Washington DC, hthronso@hq.nasa.gov, pjulrich@radix.net and gvarsi@hq.nasa.gov.

If the division (or more generally, the agent) and headquarters (the principal) are profit maximizers, then the structure of the optimal relationship or contract is well known.[1] We summarize this theory in Section 2. Under appropriate assumptions on the distribution and timing of information, the optimal contract can be implemented by the headquarters by offering a menu of linear contracts to the division such that for each realized private cost value, there is exactly one optimal contract for the division. Thus, the division's choice of the contract truthfully reveals its private information. The contract then specifies the percentage of costs to be shared between the division and the headquarters. When used in government procurement with private contractors, cost-sharing contracts have generally met with success.[2]

In most applications, the structure of the environment differs from that assumed by the theory. For example, the headquarters may be physically or legally constrained from incentivizing the division with monetary transfers.[3] When the division and the headquarters are units within the same organization, their preferences may be more aligned than is typically assumed in the principal-agent literature. The construction of a project may be time consuming and the division's private information is often resolved over time (and only after the project has been assigned) rather than being known at the outset. In this paper we focus on a particular setting, NASA's mission acquisition process, which features these complications, as well as several others. The NASA environment is detailed in Section 2.

Our goal is to identify and exploit key aspects of the solution to the canonical model in order to improve outcomes in the applied environment. We propose a new contracting process, called Multi-Contract Cost Sharing (MCCS). Under this system, headquarters offers a small number of linear cost-sharing contracts to the division. Since the the fully optimal solution features an infinite menu of linear cost-sharing contracts, the MCCS is a simplified approximation to the full solution. We then compare the MCCS process to the "Cost Cap" system currently in place at NASA. Under this system, the headquarters assigns a cost cap to each project and fully refunds any reported costs up to the cap. The moral hazard problems created by the Cost Cap rule are obvious, and we conjecture that the MCCS process will significantly improve welfare.

Because the MCCS process is only an approximation of the optimal solution and because the environment of interest differs from the canonical model, it is not immediately clear that the MCCS process would improve welfare. Although the Cost Cap system may appear inferior, it is has evolved in the NASA environment and was put in place by actors with very large stakes in the project outcomes. This implies that the existing process has properties that at least some parties view as desirable. Furthermore, decision makers may exhibit biases, errors, and learning effects not captured by the standard model, making the simple Cost Cap system the better choice. For these reasons, we turn to experimental methods to compare the outcomes under the MCCS process to the outcomes under the Cost Cap system. By simulating these processes in the lab, we can directly test our hypotheses and reach conclusions where theory remains inconclusive.

This research is of independent interest to experimental economists because it provides one possible roadmap for using experiments in the design of real-world institutions. In particular, when existing theory cannot provide a complete solution to a given problem, simplified approximations to the solutions of similar models may capture the desired incentive effects. It is then the goal of the experimentalist to test whether or not these effects survive the differences in environment and the simplifications of the solution. In that sense, experiments on institutional design can be thought of as robustness checks of existing theory. Theoretical solutions that are shown to be robust to the idiosyncrasies of real-world environments then become useful tools for practical implementation problems. Other experiments that have used this technique are briefly surveyed in Section 7.

In order to capture the many idiosyncracies of the NASA environment, our experiment is necessarily very complex. The particular stages of the experiment, the timing of information, the choice of decision variables, and the induced preferences are all the result of many discussions with NASA agents. A fairly large portion of this paper is dedicated to describing the actual environment (Section 3), our proposed mechanism (Section 4), and our laboratory simulation (Section 5). For the reader uninterested in such details, the key finding is this: The proposed MCCS mechanism outperforms the existing Cost Cap system in essentially every dimension of interest. The success of the proposed system apparently stems from the features it shares with the theoretically optimal contract. In particular, approximate incentive compatibility prevents low-cost agents from 'gold plating' cheap missions while providing adequate funding to high-cost agents, cost sharing provides insurance to the agent that helps incentivize risky innovation attempts, and intertemporal budget flexibility relaxes budget constraints and encourages agents to spend in costly periods and save in cheaper periods. The experimental results are provided in Section 6.

2 The Theory

In this section we review the standard model of contracting in the presence of adverse selection (private information,) and moral hazard (unobservable effort.) Refer to Laffont and Tirole [21, Chapters 1,4] for a more complete treatment and for relevant extensions.

In our setting, an agent who builds a project is faced with an exogenous 'luck' cost $L \in \left[\underline{L}, \overline{L}\right]$ that is private information. She must then choose a level of effort $e \geq 0$, yielding a final cost of $C(e; L) = L - e$. The disutility of her effort is $\psi(e)$, an increasing, convex function with $\psi(0) = 0$. A moral hazard problem exists because the principal observes $C = C(e; L)$, but not L or e. He then reimburses the agent by $t(C)$.

The agent's profit is then $U(e;L) = t(C(e;L)) - C(e;L) - \psi(e)$.

The principal has a fixed value of S for the completed project and is interested in maximizing overall welfare. Therefore, the principal's *ex-post* payoff becomes

$$V(e;L) = S - (1+\lambda)\, t(C(e;L)) + U(e;L), \tag{1}$$

where $\lambda > 0$ is the opportunity cost of funds spent.

Laffont & Tirole [21, Chapter 1] show that the optimal, incentive compatible, individually rational transfer $t^*(C)$ is a convex function of C that depends on the functional form of ψ. For each level of luck L, the agent will prefer the level of effort $e^*(L)$ (and, consequently, the final cost $C^*(L)$) that is socially optimal. By its equilibrium choice of $C^*(L)$, the agent fully reveals its private information L. Note that $C^*(L)$ being the agent's profit maximizing choice implies that her indifference curve in (C,t) space is tangent to t^* at $C^*(L)$.

Instead of using this convex transfer scheme directly, the principal can implement the same outcome using a family of linear transfer functions, each of which is tangent to $t^*(C)$ at a unique point. Specifically, the principal asks the agent for a cost estimate C^E, observes the final cost C, and pays an incentive transfer

$$T^*\left(C^E, C\right) = t^*\left(C^E\right) + \beta\left(C^E\right)\left(C - C^E\right), \tag{2}$$

where $\beta\left(C^E\right) = t^{*\prime}\left(C^E\right)$. This menu of tangents forms the lower envelope of the convex transfer function $t^*(C)$. Since the agent's indifference curve is tangent to t^* at $C^*(L)$, it is also tangent to $T^*(C^*(L), C)$ at $C = C^*(L)$. Furthermore, any other cost estimate announcement $C^E \neq C^*(L)$ necessarily gives the agent a lower payoff, regardless of the final cost C. Finally, $T^*(C^*, C^*) = t^*(C^*)$, so the same transfer results in either scheme. Thus, the menu of linear contracts is incentive compatible and results in the same outcome as t^*. We have that the agent will reveal a cost estimate of $C^E = C^*(L)$, which perfectly reveals her information, and will choose the effort that results in a final cost of $C = C^*(L)$.[4]

To see this graphically, consider Figure 1. An agent with a low luck parameter has a socially optimal cost C^L. Because ψ is concave, the agent's indifference curve in (C,t)-space is convex. The optimal convex contract t^* is tangent to the indifference curve U_L at C^L, so the agent prefers her final cost to equal C^L. The linear contract $T^*\left(C^L, C\right)$ is also tangent to U_L at C^L, so if the agent announces C^L, she will prefer her final cost to equal C^L and she will receive the same transfer, $t^*\left(C^L\right)$. Suppose the low-cost agent chooses a different contract, say by announcing C^H. Since every point on the contract $T^*\left(C^H, C\right)$ is below U_L, the agent is made strictly worse off. Thus, the agent prefers to announce C^L. In this way, every agent prefers to announce the socially optimal cost level (given his private information) and then exert effort such that the final cost matches this level.

The function $\beta\left(C^E\right)$ in equation (2), which is increasing in C^E, determines the percentage of cost overruns the agent must bear (or what fraction of cost savings the agent may keep.) For lowest value of C^E (the most efficient agents,) $\beta = 0$ and the contract is simply a fixed-price contract. In this case, the agent faces the highest-powered incentive scheme because they bear the entire cost of an overrun but may keep the entire benefit of any cost savings. The reward for bearing this risk is a higher fixed payment (represented by the intercept of T^*.) Less efficient (high-cost) types self-select a higher β and therefore face weaker cost-saving incentives, as in a 'cost-plus' contract.

Laffont & Tirole [21, Chapter 7] analyze the same model with $n > 1$ agents. The agents compete in an auction in which each submits a bid representing their cost estimate. The optimal incentive compatible auction awards the project to the lowest bidder and uses the optimal linear contract above, where the agent's bid is used as the cost estimate C^E. As the number of agents increases, the information rents accruing to the winning agent are reduced and the outcome approaches the first-best outcome.

3 Application: NASA Mission Acquisition

In this section, we describe in detail the system whereby the National Aeronautics and Space Administration (NASA) procures upcoming missions. Many individuals within NASA view the existing process as unsatisfactory, in part because adverse selection and moral hazard problems inherent in this process have contributed to several recent mission failures. There is interest in developing a new system for mission acquisition. However the various complications of reality make the applicability of the theory from the previous section questionable. We now describe the current process in detail. With an understanding of the idiosyncracies of the environment, we then propose an alternative process inspired by aspects of the theory described in Section 2.[5]

3.1 The NASA Mission Acquisition Environment

The organization of the main actors in NASA's mission acquisition process is given in figure 2. The three key players are the Office of Management and Budget (OMB), who oversees NASA's budget and approves mission requests by NASA, the Space Science Enterprise office at NASA headquarters (Associate Administrator and five Science Theme Directors), and the Implementing Centers (which can be thought of as divisions of the main NASA organization). In addition, the space science community at large and their committees and councils are involved as first line customers, the American public as the ultimate customer, and various levels of management in the organizations of the key three players participate in the decision processes.

We focus here on the relationship between NASA headquarters, who has the role of procuring new missions, and the Implementing Centers, who are the contractors, but who are also part of the overall NASA agency. The Implementing Centers manage the construction of missions, often subcontracting with the private sector. The four largest Implementing Centers for the Space Science Enterprise are the Jet Propulsion Laboratory, Goddard Space Flight Center, Marshall Space Flight Center, and Ames Research Center. These organizations rely on formal agreements and contracts to develop and manage NASA missions as their principal source of revenue, and although they have an interest in the overall success of NASA, they also have an interest in larger shares of the Space Science budget being allocated to their own part of the organization rather than to the other centers.

3.2 The Current Process of Mission Acquisition

Every three years, the Space Science Enterprise at NASA Headquarters develops a menu of missions on which NASA will focus.[6] One or more Implementing Center ('IC') reviews each mission and provides a cost estimate. The Associate Administrator within NASA then assigns each mission to an IC. Cost estimates are further refined and tested for viability within the overall budget. As a result, the IC is given a cost cap that indicates the maximum allowable expenditure for the mission. Expenditures are fully reimbursed up to the cost cap, but ICs are not permitted to spend above the cap.[7] Panel (a) of figure 3 illustrates the reimbursement structure of a Cost Cap contract.[8]

As a mission enters the planning and construction phase, it often becomes clear that the initial estimates were optimistic and the cost cap constraint will bind. In this case, the IC may either (1) descope the mission by reducing the mission's goals, (2) remove tests, analyses, and redundancies, increasing the risk of mission failure, (3) request additional funding from headquarters, or (4) cancel the mission. Descoping, although undesirable, may be admissable as long as the resulting science content of the mission remains above some lower bound. When this 'science floor' is reached, costs can be further reduced only through increased risk.

When the IC requests an increase in funding, NASA Headquarters may (a) cancel the mission and reallocate the funds, (b) move funds from another mission to this mission, or (c) request additional funding for this mission from OMB and Congress. Any reshuffling of the budget requires authority from Congress and is a non-trivial matter. Requesting an increased budget is considered undesirable because it damages NASA's reputation within Congress, and Congress may respond by denying the request and cancelling the mission.

Clearly the choice of the cost cap is a delicate matter. For the Headquarters, high caps are wasteful, but low caps are risky. Future budgets allocated from Congress are sensitive to these choices. From the IC's point of view, reporting high cost estimates may lead Headquarters to assign the project to another center. However, reporting low cost estimates can result in a higher probability of cost overruns and failures. The failure of two missions in late 1999 (the Mars Climate Orbiter and Mars Polar Lander) was blamed in part on aggressive bidding by sub-contractors and cost pressure created by a tight cost cap.

4 The Proposed Process – MCCS

To develop an improved process for mission acquisition, we look to the theory of section 2. The optimal solution in that environment is an infinite menu of linear contracts. Agents with low cost estimates self-select a flatter reimbursement schedule that gives them no incentive to overspend (mimicing the high cost types,) but allocates the majority of any cost savings to the agent. High cost agents self-select steeper contracts that guarantee enough funding to cover their higher cost, but they retain a relatively small percentage of any realized cost savings.

Our proposed system differs in two fundamental ways from the system derived from theory. First, only three linear contracts are offered: high, middle and low. Agents with relatively high cost estimates will prefer the 'high' contract because, like the optimal solution above, it guarantees enough funding for the project to be built. Low cost agents will prefer the 'low' contract because it allows them to retain a larger share of the cost savings. Reducing the number of available contracts to three simplifies the decision problem for the agents.[9] Note that although these three contracts may not be fully incentive compatible, the fact that the utility function of the agent is not really known by the principal implies that even an infinite menu of contracts would not necessarily be incentive compatible.

The second difference is that agents negotiate the three contracts with the principal early in the process, then choose the actual contract late in the process after some of the cost uncertainty has been resolved. This allows the parties to establish a set of contracts early that are sufficiently flexible so that renegotiations are not necessary. As cost uncertainty is resolved, the agent knows that, regardless of actual costs, she will receive adequate funding to complete the project. The principal knows that the resulting contract will be (roughly) incentive compatible, so that low types do not overspend in an effort to maximize their reimbursement.

Specifically, the MCCS process consists of the following steps.

1. The principal negotiates a cost baseline C^B for each agent for each project.

2. Three contracts are specified according to the formula

$$T\left(C, C^k\right) = \alpha^k + \beta^k \left(C - C^k\right), \tag{3}$$

where C represents the final cost, $k \in \{L, B, H\}$ indexes the three contracts, and α^k, β^k, and C^k are all known functions of the negotiated baseline C^B.

The three contracts satisfy the following incentive compatible constraints: there exist C^* and C^{**} such that $C^L < C^* < C^B < C^{**} < C^H$ and

$$T\left(C^F, C^L\right) > T\left(C^F, C^k\right) \ \forall k \neq L \text{ if } C^F < C^* \tag{4}$$

$$T\left(C^F, C^B\right) > T\left(C^F, C^k\right) \ \forall k \neq B \text{ if } C^* < C^F < C^{**} \tag{5}$$

$$T\left(C^F, C^H\right) > T\left(C^F, C^k\right) \ \forall k \neq H \text{ if } C^F > C^{**} \tag{6}$$

Equations (4)-(6) require that agents with final costs below C^* prefer the low contract, agents with intermediate costs prefer the baseline contract, and agents with final costs above C^{**} prefer the high contract. Thus, $\beta_H > \beta_B > \beta_L$.

3. At a later stage at which agents have more precise cost estimates, one of these three contracts is selected by the agent.

4. The final cost C is observed and $T\left(C, C^k\right)$ is paid to the agent, where k indicates the selected contract. The difference $T\left(C, C^k\right) - C$ is either allocated to or drawn from a bank balance of funds held by the agent. These funds can be used for future projects.[10]

A hypothetical example of such a system of contracts is shown in panel (b) of figure 3. The agent's cost of completing the project is shown on the horizontal axis and the reimbursement received from the principal is given on the vertical axis. Notice that, as required, each of the three contracts is optimal for the agent for a range of cost realizations.

5 The Experiment

In our experiment, we compare the existing Cost Cap system to the proposed MCCS system by creating an environment in the laboratory that simulates the salient features of the NASA mission acquisition process. Subjects assume either the role of Headquarters (HQ) or an Implementing Center (IC) and proceed to negotiate contracts and 'build' missions in this simulated environment.

In comparing the two contracting systems, we ask three specific questions. First, how does the MCCS system perform (relative to the Cost Cap system) with and without adverse selection? Second, how does the MCCS system perform when cost shocks are particularly extreme? Third, does experience with the MCCS system improve performance? These questions suggest a $2 \times 2 \times 2 \times 2$ factorial design in which we vary the number of ICs (one or two), the variance of the 'luck' shocks (high or low), the experience of the subjects (inexperienced or experienced), and the contracting system used (Cost Cap or MCCS.) We then compare performance across these various treatments.

5.1 The General Environment

In each session in our design, one HQ interacts with the same ICs for five periods. In each period, an identical menu of two missions $j \in \{A, B\}$ is available, each with three different 'levels' $l \in \{1, 2, 3\}$. Missions are labelled A1–A3 and B1–B3. No more than one A mission and one B mission can be built in a given period, but the same IC could be assigned both missions. The $j1$ mission represents the high-cost, high-risk, high-reward design of the j mission, while $j3$ represents the low-cost, low-risk, low-reward design. When there is no confusion, the index l is dropped for convenience.

After negotiation, HQ assigns missions to ICs, who then build their assigned mission(s). This is done by selecting the 'science content' $S_{ij} \geq 0$ and 'reliability' $R_{ij} \in [0, 1]$ for each mission j assigned to IC i. Reliability is the probability of success for the mission. If the mission succeeds, the chosen science content S_{ij} is realized. Both the HQ and the IC building the mission receive benefits of S_{ij} for successful missions. If the mission fails, a fixed failure cost F_j is paid. Thus, the expected value of mission j for the HQ and IC i is $R_{ij}S_{ij} - (1 - R_{ij}) F_j$. Both R_{ij} and and S_{ij} are costly, so the IC faces a trade-off between science content and mission reliability. All other ICs receive zero payout from this mission.

Each level of each mission has an exogenously determined science 'floor' $\underline{S}_j$. An IC is not permitted to launch a mission with $S_{ij} < \underline{S}_j$. If an IC has inadequate funding (perhaps due to unfortunate luck shocks,) she may be unable to choose $S_{ij} \geq \underline{S}_j$ while keeping R_{ij} at an acceptable level. In this case, the IC may choose to cancel the mission, yielding zero payoff to herself and to the HQ. The funds allocated to a canceled mission are not recovered and can not be used for other missions.

Science content is measured in 'points', and subjects are ultimately paid based on the number of points they earn from successful missions. Budgets within the simulation are measured in 'francs', an experimental currency that has no intrinsic value to the subjects. The cost of building S_{ij} and R_{ij} is measured in francs, and the total budget of francs available dictates the trade-off that must be made between S_{ij} and R_{ij}. Francs are never converted into actual payouts for the subjects, so there is no incentive to save francs except to use them in building other missions.

Before building a mission, ICs have the ability to lower construction costs by attempting some innovation. Specifically, if an IC chooses to spend e_{ij} francs on innovation, then with some probability $P(e_{ij})$, the building cost of the mission will decrease by a fixed amount.

Finally, the cost of construction is affected by a random variable L_{ij} representing the luck encountered by the IC during planning and construction. The IC knows L_{ij} when construction begins, but the HQ knows only the distribution of possible values of L_{ij}. The expected value of L_{ij} is zero, and larger L_{ij} means a higher final cost. In practice, three independent mean-zero luck shocks will be drawn at different points in time. The total luck L_{ij} will then equal the sum of these three mean-zero shocks, L^1_{ij}, L^2_{ij}, and L^3_{ij}.

The total cost to i of building a mission j is given by

$$C_j(S_{ij}, R_{ij}) + e_{ij} + L_{ij} = a_j S_{ij}^2 + b_j \ln\left((1 - R_{ij})^{-1}\right) + e_{ij} + L_{ij}.$$

The coefficient a_j determines the cost of science and the coefficient b_j determines the cost of reliability. Successful innovation results in a lowering of a_j by $1/3$, thus making science relatively more affordable. Expending e_{ij} yields such an innovation with probability

$$P(e_{ij}) = 1 - z_j^{-e_{ij}},$$

where z_j is a mission-specific parameter.

These particular cost and innovation functions are the result of discussions with employees of NASA. They capture the salient trade-offs between science, reliability, and innovation spending. The cost functions have the property that each design is the lowest cost method for attaining some level of S_{ij}; level 3 for relatively low science output, level 2 for intermediate output, and level 1 for high output.

At the beginning of each period, an estimated cost C_j^E is calculated by

$$C_j^E = C_j\left(\underline{S}_j, 0.95\right).$$

This represents the cost of building $S_{ij} = \underline{S}_j$ with 95% reliability and no innovation spending or luck shocks. This number is known to all subjects. It is also common knowledge that HQ is endowed with a budget $B = 1500$ francs each period.

Agents within NASA have a known preference for building missions with high science content. To capture this incentive, we award a bonus of $y_{j1} = 500$ points to the IC for a successful 'level 1' (high-cost/high-value) mission. We set $y_{jl} = 0$ for $l > 1$. The actual coefficients used in the experiment for each mission and each design are given in table 1. Recall that expected payoffs for a mission j are given by

$$E[\pi_{HQ}] = R_{ij} S_{ij} - (1 - R_{ij}) F_j, \text{ and} \tag{7}$$

$$E[\pi_{IC}] = R_{ij} (S_{ij} + y_j) - (1 - R_{ij}) F_j. \tag{8}$$

5.2 Timing: Cost Cap Process

To provide a standard of comparison for our proposal, we implement a stylized version of the current contracting process used at NASA for mission acquisition in the context of our experimental environment. The sequence of events in a period of the Cost Cap process is described below and summarized in panel (a) of Figure 4.

Initial Information All subjects see the budget for the period (1500 francs) and the initial cost estimates C_j^E. Each IC i also sees its first luck shock L_{ij}^1, which is private information.

1st Negotiation Each IC i requests an initial cost cap κ_{ijl}^1 for each mission jl. A request of zero indicates the IC does not want to build mission jl. The HQ responds with a counteroffer for each κ_{ijl}^1 such that $\sum_{i,j,l} \kappa_{ijl}^1 \leq B$ and, for each j, $\kappa_{ijl}^1 > 0$ for at most one (i, l) pair.[11] This process repeats three times.[12] After the third round, the HQ's counteroffer κ_{ijl}^1 becomes the assigned cost cap.[13]

1st Innovation Each IC i chooses $e_{ij}^1 \geq 0$ for each j such that $\kappa_{ij}^1 \geq 0$. With probability $P(e_{ij}^1)$, a_j is reduced by 1/3.

1st Construction Each IC i observes L_{ij}^2 and then chooses an initial level of science S_{ij}^1 and reliability R_{ij}^1 for each j such that $\kappa_{ij}^1 > 0$. At the end of this phase, the IC has spent $C_j(S_{ij}^1, R_{ij}^1) + L_{ij}^1 + L_{ij}^2 + e_{ij}^1$ francs.

2nd Negotiation Exactly like 1st Negotiation, but with only one offer and counteroffer of the final cost cap, κ_{ijl}^2. If the level of a particular mission assignment changes ($\kappa_{ijl}^1 > 0$, $\kappa_{ijl}^2 = 0$ and $\kappa_{ijl'}^2 > 0$), then 50% of the initial science and reliability transfers to the new level ($S_{ijl'}^1 = S_{ijl}^1/2$ and $R_{ijl'}^1 = R_{ijl}^1/2$.)[14]

2nd Innovation Each IC i chooses $e_{ij}^2 \geq 0$ for each j such that $\kappa_{ij}^2 \geq 0$. With probability $P(e_{ij}^2)$, the current value of a_j is reduced to 2/3 its current level (this will be 4/9 of the original level if the first innovation was also successful.)

2nd Construction Each IC observes L_{ij}^3 and then chooses S_{ij}^2 and R_{ij}^2. Total science and reliability are $S_{ij} = S_{ij}^1 + S_{ij}^2$ and $R_{ij} = R_{ij}^1 + R_{ij}^2$, respecitvely, and total science must satisfy $S_{ij} \geq \underline{S}_j$. The total cost is $C_j(S_{ij}, R_{ij}) + \sum_{t=1}^{2} e_{ij}^t + \sum_{t=1}^{3} L_{ij}^t$. The IC may choose to cancel the mission at this point instead of selecting S_{ij}^2 and R_{ij}^2.

Launch Missions that were not canceled are launched. With probability R_{ijl} the mission is successful and pays $S_{ijl} + y_{jl}$ points to IC i and S_{ijl} points to the HQ. With probability $(1 - R_{ijl})$ the mission fails and costs IC i and the HQ F_{jl} points. In either case, IC $i' \neq i$ earns no points for the mission. Canceled missions pay zero points. Outcomes of all missions are observed by all subjects.

There is no cost reimbursement here. All expenditures in Stages 4 and 7 were required to be less than or equal to the cost caps negotiated in Stages 2 and 5. Any excess francs at this point are removed from the experiment; neither the IC nor HQ may keep them for future periods. Success and failure do not affect the budget HQ has in each period, though they may affect HQ's budget allocation policy toward the ICs in later periods.

It is important to note that the HQ is effectively a dictator in the negotiation of the cost caps, so it may seem that the later stages of negotiation are irrelevant. However, the entire process is included in the simulation because it not only models the real-world interaction, but it is conjectured that important private information may be revealed by the ICs through their requests.

5.3 Timing: MCCS

The simulation of the MCCS system in the experimental environment described above yields the sequence of events during a single period illustrated in panel (a) of Figure 4. Activity in each period can be divided into a series of stages.

Initial Information Same as the Cost Cap process above, except the annual budget of 1500 is added to the HQ's bank account. The annual budget is commonly known. HQ's bank account balance is not shown to the ICs, but could be calculated from the results of previous periods.

Negotiation Same as 1st Negotiation under the Cost Cap process above, except that offers and counteroffers are in terms of the baseline cost C_{ijl}^B of each mission jl rather than a cost cap. The HQ's third counteroffer is the effective cost baseline from which the menu of three cost contracts is calculated by equation (3).[15] The values of C_{jl}^k and α_{jl}^k are fixed multiples of C_{ijl}^B, and β_{jl}^k is a fixed constant independent of C_{ijl}^B. The exact formulas used are given in Table 2. During negotiation, the HQ is constrained to have, for each j, $C_{ijl}^B > 0$ for at most one (i, l) pair, but the sum of the cost baselines is *not* constrained to be below the annual budget.

1st Innovation Same as the Cost Cap process above.

Contract Choice Each IC i observes L_{ij}^2 and then picks a contract $k_{ij} \in \{L, B, H\}$ for each mission j for which $C_{ij}^B > 0$.

2nd Innovation Same as the Cost Cap process above.

Construction Each IC i chooses science S_{ij} and reliability R_{ij} for each mission j for which $C_{ij}^B > 0$, or else chooses to cancel the mission. The final cost is given by $C_{ij}^F = C_j(S_{ij}, R_{ij}) + \sum_{t=1}^2 e_{ij}^t + \sum_{t=1}^3 L_{ij}^t$.

Reimbursement HQ reimburses IC i by $T_i = \sum_{j:C_{ij}^B>0} T(C_{ij}^F, C_{ij}^k)$ according to equation (3). The HQ's bank account is reduced by $\sum_i T_i$ and IC i's bank account is reduced (or increased) by $\sum_j C_{ij}^F - T_i$. Positive bank balances are carried forward into future periods.[16]

Launch Same as the Cost Cap process above.

5.4 Comparing the Experiment to the Model

There are significant differences between the experimental environment and the assumptions of the theoretical model described in Section 2. These differences arise naturally because our current application (NASA mission acquisition) does not fit well the assumptions of the theory. We have therefore deviated from the optimal contracting theory in two directions; our proposed mechanism is only a weak approximation of the optimal mechanism, and our environment differs from that assumed by the theory. The following is a short list of some key differences between the two environments. Each of these may impact our results and our ability to apply the theory to our particular application.

1. In the experiment, IC i learns the components of the random variable L_{ij} as the game proceeds, rather than at the beginning. There are three independent shocks that comprise L_{ij}, at least one of which occurs after negotiations are complete. In the theory, the single unknown variable is resolved (to the IC) before negotiations occur. In the model, the adverse selection (drawing of L_{ij}) occurs before the moral hazard (choice of e_{ij}). In the experiment, adverse selection problems and moral hazard opportunities occur in overlapping stages during the process.

2. Missions are not of a fixed size. The IC's choice of science scales the benefit of the mission continuously. There are global budget constraints restricting the total expenditure on missions. In fact, missions

are two-dimensional commodities, so preferences must be defined over the space of feasible science-reliability pairs.

3. In the experiment, greater effort yields a greater probability of lowering the cost, but it is not deterministic.

4. In the model, T_i represents wealth transferred to the agent from the principal. In the experiment, T_i can only be used to fund future missions. It is an input into future production that will benefit the IC, so any utility value assigned to the transfer must be a measure of the expected value of its future production output.

5. In the current application, preferences are nearly aligned. The agent and the principal both receive utility from the completion of the mission. For example, engineers and scientists working on a mission often use the scientific output from the mission in their own research, and thus are also consumers of the science the mission generates.

6. In the experiment, multiple missions may be contracted simultaneously in a period. Two missions may be assigned to the same IC. There are multiple designs for each mission. These complications introduce additional competition and coordination issues missing from the model.

5.5 Experimental design, treatments, and sessions

We systematically vary the mechanism, number of ICs, luck shock variance, and subject experience in our 2^4 factorial design, yielding 16 possible treatments. We vary the mechanism between the Cost Cap and MCCS processes to compare our proposed system to the established system. We vary the number of ICs from one to two to test how the processes handle the case of almost perfectly aligned preferences (1 IC) and the case of adverse selection and competition among ICs (2 ICs). Changing the luck shock variance tests the robustness of the data to changes in the underlying properties of the project; perhaps one mechanism is better adapted than the other for situations with inherently higher risk. In the low variance treatments, each luck shock is drawn from the uniform distribution over $[-200, 200]$ for level 1 missions and $[-50, 50]$ for level 2 and 3 missions, so that each luck shock may be as large as 17%, 7%, and 15% of the initial cost estimate, respectively. In the high variance treatments, level 1 luck shocks are uniform over $[-500, 500]$, level 2 luck shocks are uniform over $[-300, 300]$, and level 3 luck shocks are uniform over $[-100, 100]$ (as large as 41%, 43%, and 30% of initial cost estimates.) Finally, some sessions were run using subjects who had already participated in one treatment of this study.

There were 48 independent sessions where the same pair or group of subjects interacted. Each session consisted of five periods. Eighty five subjects were recruited from the graduate student population at Purdue University. The experiment was entirely computerized. The program, which was developed specifically for this experiment, was written in Perl and subjects interacted via a web site accessed from the laboratory. The instructions for the experiment are available from the authors. Multiple groups participated simultaneously and independently in the same laboratory and subjects were not informed which subjects were in their group. Each subject remained grouped with the same person(s) for the entire session. There were three practice periods for inexperienced groups (in which earnings were not paid) and one practice period for experienced groups. Inexperienced sessions averaged approximately three hours in length, and experienced sessions averaged approximately 90 minutes. Table 3 describes the amount of data gathered under each of the 16 possible treatments.

6 Results

6.1 Performance of the Multi-Contract Cost Sharing Process

Table 4 shows the per-period average expected payoff of the HQ and the ICs under each treatment. These are the expected pre-launch payoffs given by equations (7) and (8), respectively. Expected payoffs are compared to two relevant benchmarks: the No-Carryover Benchmark (NCB) and the Carryover Benchmark (CB).

The NCB is calculated by finding the allocation of missions to ICs and subsequent choices of science and reliability that maximize the *sum* of the HQ and IC's expected payoff. As in the experiment, this optimization problem is constrained to assign at most one level of each mission, to choose $S_{ij} \geq \underline{S}_j$ for each mission j, and to remain within a one-period budget constraint. The one-period horizon on the budget mimics the Cost Cap environment in which unused funds are not available in future periods. The NCB program also assumes zero luck and zero innovation effort for every IC and every mission. Note that realized expected payoffs in the Cost Cap experiment can be higher than the NCB because missions may be assigned to ICs with negative (good) luck shocks and ICs may successfully innovate, reducing overall costs. The expected payoffs of the NCB solution are 672.2 points for the HQ and 1160.2 *total* points for the ICs.[17]

The CB is identical to the NCB taken over a five-period budget horizon. It is the optimal allocation of missions and choices of science and reliability for a five period time span assuming that money can move freely between periods. This is equivalent to solving a single period problem with five times the budget and five copies of the menu of missions. As with the NCB, luck and effort are fixed at zero, so actual subjects in the MCCS treatments can outperform the CB by choosing missions with negative luck shocks and successfully innovating those missions. The expected payoffs of the CB solution are 188 points higher for the HQ and 95 points higher for the ICs than the NCB. These differences provide a rough estimate of the value of intertemporal budget flexibility.

We argue in Observation 1 that the MCCS process earns higher payoffs for experienced subjects than the Cost Cap process *and* the two benchmarks. Although inexperienced subjects in the MCCS process often fail to exceed the CB, expected payoffs are still greater than under the Cost Cap system with *experienced* subjects. These results indicate that the MCCS process unambiguously outperforms the Cost Cap process.

The analysis of the following observations is done simply by examining the main effects of the 2^4 factorial design. For example, payoffs under the Cost Cap process can be compared to payoffs under the MCCS process under 8 separate treatments: [1IC vs. 2ICs] $\times$ [Experienced vs. Inexperienced Subjects] $\times$ [High vs. Low Variance]. If these treatments are independent and the MCCS and Cost Cap processes generate the same payoffs on average, the probability of the MCCS outperforming the Cost Cap process across all eight treatments is approximately 0.39%.[18] If we then observe that the MCCS payoffs are greater than the Cost Cap payoffs across all treatments (or even across seven of eight treatments,) we can reject the hypothesis of equal average payoffs at any standard significance level. This is indeed the case.

Observation 1 *The payoff to both HQ and the IC's is (a) greater under MCCS than under Cost Cap and (b) greater under MCCS with experienced subjects than both benchmark payoff levels.*

Support. Table 4 show that MCCS yields higher average values of both $E[\pi_{HQ}]$ and $E[\pi_{IC}]$ than Cost Cap under all 8 comparible treatments. With experienced subjects, the MCCS process outperforms the No-Carryover and the Carryover benchmarks for 4 of 4 treatments for the HQ and 4 of 4 treatments for the ICs, each of which is significant at the 10% level. This also implies that the sum of HQ and IC payoffs is significantly larger for experienced subjects under MCCS than Cost Caps. ■

Thus both HQ and the IC receive gains from a change from Cost Cap to MCCS. The added benefit of the MCCS process over the Cost Cap process is larger for the ICs. This is at least partially due to the fact that the cost sharing allows enough budget flexibility for the ICs to produce level 1 missions more frequently and consequently earn more 'prestige' bonuses.

It is also clear that benefit of subject experience is smaller under the Cost Cap process than under the MCCS process, particularly for the ICs.[19] We offer two explanations for this phenomenon. First, the optimization problem under the Cost Cap system is simpler because budget inflexibility causes each period's decision problem to be independent of the next. In the MCCS system, the ability to carry unspent budgets into future years introduces a dynamic component to the decision problem, making the optimum more difficult to identify or approximate. Second, myopic decision-making would reduce efficiency in the MCCS process but would not affect earnings in the Cost Cap process; an effect that would disappear with experience if subjects learn to maximize overall earnings rather than current period earnings.

One interesting observation is that experience apparently *reduces* HQ and IC payoffs in the Cost Cap system with 2 ICs and low variance luck shocks. Given that only one group participated in the 2IC, low variance, experienced subjects treatment, we are cautious about the value of such comparisons. We do speculate that in the presence of 2 ICs, the HQ may exhibit a preference for equity by always assigning one mission to each IC (rather than giving both missions to one IC and forcing the other to earn zero for the period.) This added constraint should lower payoffs and increase cancelations, which we observe (Table 5.) However, we do not observe such phenomena in other treatments and cannot conclude that the effect is robust.

A natural question is whether the superior performance of the MCCS process is due entirely to the budget flexibilty. If we permitted agents in the Cost Cap system to carry savings forward for use in future years, would it perform as well as the MCCS system? Or is there some benefit due to the risk-sharing and (approximate) incentive compatibility aspects of the MCCS mechanism? To address this issue, we compare the sum of expected payoffs ($E\left[\pi_{HQ}\right] + E\left[\pi_{IC}\right]$) between the MCCS process and the Cost Cap process in each of the four treatments. This difference represents the efficiency benefit of the MCCS over the Cost Cap system. In all four treatments, this difference is larger than 283 points, which is the difference between the sum of payoffs under the CB and the sum of payoffs under the NCB. This suggests that budget flexibility is not the only source of benefit of the MCCS system over the Cost Cap system.[20]

One final observation gleaned from Table 4 is that, compared to the NCB, the HQ earns more and the ICs earn less under the Cost Cap than predicted (with the exception of the anomalous 2IC, low variance, experienced subjects treatment discussed above.) In the experiment, HQ acts as a dictator in the negotiation phase. Even though preferences are nearly aligned, the HQ may choose suboptimal allocations by assigning fewer level 1 missions for which the IC receives a bonus. This rational bias could reduce the average IC payoff by as much as 500 points without significantly reducing the average payoff to the HQ.

Since the MCCS mechanism realizes higher value than optimal decision making when luck and effort are set to zero, it must be the case that the MCCS process mitigates the adverse selection problem by helping the HQ assign missions to ICs with lower (better) luck shocks, or it mitigates the moral hazard problem by sharing the risk of innovation between the IC and the HQ, or both. These effects can be examined by looking at measures of mission performance. Observation 2 below summarizes differences between the two systems in terms of final costs, the frequency of of non-delivery, the probability of failure, and the number of innovations realized per period.

Observation 2 *On average, there is more innovation, lower final cost and less frequent non-delivery of projects under MCCS than under Cost Cap.*

Support. The data in table 5 show that there is more innovation per mission under MCCS than under Cost Cap in seven of eight treatments (and an equal amount in the eighth treatment). The cancellation of missions is more frequent under Cost Cap than under MCCS for each of the eight relevant comparisons. Average final cost relative to original estimated costs are lower under MCCS than under Cost Cap for seven of eight possible comparisons, even when the missions that are not delivered (which occur more frequently under Cost Cap and which tend to have unfavorable cost shocks) are censored from the data. ■

It appears that the MCCS system yields more innovation than Cost Cap because of the intertemporal budget flexibility and the risk-sharing of innovation between the IC and the HQ. Under the high contract, the cost sharing parameter is larger, so the HQ provides more insurance to the IC. Risk sharing is reduced under the low contract, but the savings from innovation are still transferable to future years. In either case, the risk to the IC of attempting an innovation is lower relative to the Cost Cap design.[21]

The most economically significant result is the reduction in cancelled missions when switching to the MCCS system. Under the Cost Cap system, more than one out of every five missions was cancelled, while fewer than one in thirty was cancelled with the MCCS design. In practice, a cancelled mission is delayed to a future budget cycle, creating a backlog of incomplete, costly projects. Removing these delays frees up present and future resources for other missions.

Under Cost Cap, the incidence of non-delivery does not appear to decrease with experience, so there is no evidence that it is a transitory phenomenon that would disappear over time. Non-delivery is more common under HiVar and when there are two IC's. With high variance, large positive (bad) luck shocks can run a mission over its budget. Without cost sharing and intertemporal budget flexibility, the mission must be cancelled. With two ICs, strategic behavior leads to 'lowball' cost estimates and cost caps, ultimately increasing the number of delayed missions. Under an incentive compatible mechanism, such lowball bids are not profitable for the ICs and fewer delays will result. The approximate incentive compatibility of the MCCS system provides these incentives.

Final costs are roughly 15% lower for completed missions on average under MCCS than under Cost Cap. Under both of the mechanisms, conditional on completion, average costs are lower under HiVar than under LoVar. This reflects the subjects' ability to select missions with low initial cost shocks. The cost per mission shows little tendency to decrease over time under either process, but the composition of missions changes over time in MCCS to shift to higher value missions, leading to increases in payoffs.

6.2 Where does MCCS under-perform?

We have argued that the MCCS process is a significant improvement over Cost Caps in those dimensions that we consider important for success. The experimental data, however, allow us to identify three systematic biases in decision-making by participants in the MCCS system. The modification of MCCS in ways that would reduce these biases could further improve the performance of the system.

Bias 1 *MCCS exhibits a tendency toward overinvestment in effort relative to optimal behavior, particularly with 1 IC.*

Support. Optimal investment in innovation is defined as the level of investment that minimizes the expected cost of completing the mission at the target value S_j. Under MCCS and 1IC, 81.8% of design 1 missions were characterized by greater than optimal effort; while 12.8% had lower than and 5.4% had exactly the optimal level. Under 2IC, 54.2% of design 1 missions included overinvestment, 41.7% had underinvestment and the rest were optimal. ■

Bias 2 *In the 2IC treatment of MCCS, there is some tendency to award contracts to the IC that does not have the lowest estimated cost at the time of the award.*

Support. Under MCCS, only 65.5% of missions were awarded to the IC with the lowest initial cost estimate.[22] Most inefficient allocations appear to be due to a tendency for HQ to try to distribute the missions evenly between the ICs, even though this means often not awarding missions to the lowest cost suppliers. Although an even distribution of contracts may be a goal of NASA in practice, there were no direct incentives in the experiment for this to occur. ■

Bias 3 *At the levels of science and reliability chosen by the IC's, the marginal return on expenditure tends to be lower on science than on reliability. Expected payoffs would be increased by shifting money from science to reliability.*

Support. The marginal return on R_{ij} was greater than on S_j in all but 6 of 37 completed missions under MCCS in which $S_{ij} > \underline{S}_j$ (we exclude cases where $S_{ij} = \underline{S}_j$ because spending on science can not be reduced.) Of the 6 exceptions, 5 were under 1IC. In contrast, under Cost Cap, in 50 of 103 missions in which $S_{ij} > \underline{S}_j$, the marginal return to reliability exceeded that of science. ■

7 Discussion

We have used the theory of optimal contracts to design a budget allocation mechanism that outperforms the existing process despite the fact that the proposed mechanism only approximates the features of the optimal contract *and* the environment of interest does not fit the assumptions of the theory. The success of the mechanism comes not from the exact specifications of the theory, but from the incentive properties of the theoretical solution that can be applied more generally. Without knowing the exact functional form of the agent's utility, a principal can still use a menu of contracts to estimate an agent's type and induce the agent to choose approximately optimal actions given her type. If the principal is better informed about the agent's preferences *ex ante*, the estimate of the type ought to be more accurate and the induced behavior closer to optimal.

Our mechanism also sacrificies some efficiency in exchange for transparency. When the space of possible types is large, incentive compatibility typically requires an equally large menu of available contracts. By grouping agents of 'similar' types into one contract that is most preferred for the group, full incentive compatibility is lost, but the menu of contracts (and therefore the choice set of the agent) is significantly reduced. In a world with boundedly rational agents, this reduction of the mechanism may in fact be efficiency *improving* if simplifying the agent's decision makes her more likely to select the contract that maximizes her payoff. The number of contracts that the principal can feasibly offer may itself be constrained, either by the boundedness of the agent's rationality or by some exogenous legal or technical constraints.[23]

It is not transparency alone that is responsible for the mechanism's success. In our application, the proposed 3-contract mechanism is almost certainly less transparent than the existing single-contract Cost Cap mechanism, despite its simplification relative to the fully optimal solution. The Cost Cap system is completely specified by the negotiated cost cap, while our 3-contract mechanism requires negotiating over a single parameter that determines the menu of available contracts, and then requires the agent to select from the resulting menu. It is clear that the incentive aspects of the mechanism play a role in its success relative to the established process.

In our case, the properties of our proposed mechanism that appear to enhance performance are the following. (1) Risk sharing between the principal and agent mitigates the moral hazard problem by encouraging risky innovation attempts. (2) Intertemporal budget flexibility provides insurance against bad luck in certain periods by drawing resources from earlier or later periods. This increases innovation attempts and reduces costly project delays and cancellations. (3) Cost sharing between the principal and agent allows additional flexibility for covering costly bad luck and failed innovation. (4) Approximate incentive compatibility induces low-cost agents to select the contract with less cost sharing. This prevents low-cost agents from unnecessarily 'gold plating' successful projects. (5) The contract is chosen late in the process when cost estimates are more precise. A contract that is completely determined early in the process cannot incorporate relevant cost information that is revealed to the agent during planning and construction. Allowing the agent to select the contract from the given menu near the end of the process allows this information to be incorporated.

The advantage of the experimental methodology in this context is clear. Our mechanism is an approximation of an optimal solution for a substantially different environment. The theory is silent about its performance in our more complex setting. Using human subjects in a laboratory simulation allows us to testbed the mechanism and compare its performance with the existing process. Of course, the degree to which the experimental results presented here parallel those that would be observed in field applications cannot be known with certainty until the mechanism is implemented.

There are several questions that arise naturally when extrapolating laboratory results to field applications. The field application certainly involves higher stakes, particularly in the case of building large space missions.[24] Experiments in which the scale of payoffs is varied show that the scale of payoffs does not influence results qualitatively (see for example Smith and Walker [31]). Experiments in developing countries, (for example Kachelmeier and Shehata [18] and Cooper *et al.* [8]) in which large sums of cash are paid relative to subjects' overall incomes, also show only minor differences in behavior from experiments using university students in developed countries. There is evidence that players are more risk averse at higher stakes (Holt and Laury [16]), but we conjecture that risk aversion would only increase the performance of our proposed mechanism relative to the Cost Cap system because it provides insurance to the agent.

Recent experiments comparing behavior in the laboratory and on television game shows are also useful in addressing the generalizability of laboratory results because the precise rules of the game show can be recreated in the laboratory. Several studies show similar behavior in the two settings (see for example Cason and Tenorio [6] or Healy and Noussair [15]). Existing studies also suggest that experimental results are not substantially different if, instead of university students, professionals with relevant real-world experience are used as subjects. Examples include King *et al.* [19] who studied the behavior of experimental asset markets with stock traders as subjects and Bohm and Carlen [4] who studied negotiation in the laboratory using diplomats as subjects.[25]

The methodology of testbedding mechanisms in the laboratory is hardly new, dating at least as far back as the study of Grether *et al.* [13] which experimentally verified inefficiencies in the existing airport landing slot allocation mechanism. Other applications that have been studied include inland water transport rate filing policies (Hong and Plott [17]), the FTC's antitrust litigation model (Davis and Wilson [10]), railroad deregulation (Brewer and Plott [5]), spectrum auctions (Binmore and Klemperer's [3]), pollution permit trading markets (surveyed by Muller and Mestleman [25]), and studies of matching market mechanisms (see Roth and Peranson [29], Roth *et al.* [30], Haruvy *et al.* [14], and Chen and Sönmez [7], for example.) Other experiments designed to test mechanisms relevant for space mission applications include Banks *et al.* [2], Noussair and Porter [26], Ledyard *et al.* [23], and Ledyard *et al.* [22].

Notes

[1]This theory is due to Laffont and Tirole, [20] and [21].

[2]There have been several instances of the use of cost-sharing contracts in US government procurement. All have used different cost sharing rules than the MCCS process. The Air Force Peace Shield program contract had an incentive structure where there was an agreed upon baseline cost, the Air Force paid 75% of any overrun, and the contractor kept 75% of any underrun. In addition, the contractor received a $50 million bonus if the project was completed early, and incurred a $50m penalty for late delivery. There was a payment ceiling of 125% of the baseline cost. Another example is Lockheed Martin's contract for the F-117A in which the Defense Department and Lockheed Martin shared the cost of any overrun and savings from any underrun at a rate of 50%. 50-50 cost sharing also applied to the US Army's procurement of the Multiple Launch Rocket System (MLRS). In 1987, the GAO conducted a review of 60 DOD incentive contracts and found that the final costs of the majority of contracts were within 5% of target costs. 47% were below and 53% were above target. 21% exceeded the original ceiling price (US Army, [32]).

Cox et al. [9] compared cost sharing and fixed-price contracting in the laboratory. Their results indicate that cost sharing allows projects to be completed at lower cost than fixed price contracting. However, cost sharing is less efficient than fixed price contracting in the sense that the contract is less likely to be awarded to the lowest-cost contractor.

[3]This is especially true of many government agencies such as NASA who are constrained by Congress from transferring money between many budget categories.

[4]The reason the optimal contact differs from the first best solution is that the principal's transfer necessarily provides the agents with some information rents.

[5]There is a reasonable likelihood that our proposal will be field tested and benchmarked against the current process.

[6]This menu is derived from reports published by the Space Studies Board of the National Academy of Sciences, which help to define an overall strategy for space science within NASA.

[7]The capping of costs is very unusual for contracts in which new technologies are involved and the costs to the contractors are highly uncertain. In most such cases in the private sector, the contractor will typically only agree to a cost-plus contract structure, in which he is paid the full amount of his cost plus a fixed amount. At NASA, before 1993, cost-plus contracting was typical, and the current system of cost caps was put in place in response to a series of severe cost overruns under the cost-plus system.

[8]Note that the cost cap process differs from a fixed price contract. Under a fixed price contract, if the final cost to the contractor is less than the fixed price, the contractor keeps the difference. Under a cost cap system, the contractor must rebate the difference to the contracting agency. The cost cap system does not contain any positive incentive to reduce costs below the cap. There is an incentive to hold costs equal to the cap, because of the threat of cancellation of the mission, which is costly both because the Center values the mission, and because its reputation would then be damaged. It would then be less likely to receive contracts for future missions.

[9]The choice of three contracts is somewhat arbitrary. The goal is to pick a finite number of contracts large enough to separate centers into groups according to their final cost estimate, but not so large that the menu of contracts is overwhelming and the mechanism is not 'transparent' to all parties involved.

[10]In the case of a non-profit enterprise or an organization in which agents cannot maintain actual balances of money, balances of 'credits' could be used, where the agent effectively borrows against future budget allocations.

[11]In other words, the HQ cannot spend more than his budget, and only one level of each mission j can be built.

[12]All negotiations were carried out through the experimental software so communication was limited.

[13]Although the HQ is a dictator in this 'negotiation' process, it is believed that the offer/counter-offer process helps coordinate mission assignment decisions. Since preferences are so nearly aligned, the offers and counter-offers may be useful in voluntarily reducing some of the information asymmetry.

[14]Again, this specification came from conversations with NASA engineers who indicated that switching a mission design mid-stream did not result in a complete loss of spent efforts.

[15]During the negotiation process, subjects have access to a computer screen that allows them to compute their reimbursement for any baseline cost, any of the three contracts, and any final cost.

[16]At the end of a period, the account of an IC or HQ could have a negative balance up to −1000 francs, but at the end of period 5, the balance was required to be positive. If either of these rules were violated, the subject was required to pay a fine of 3000 francs, which was prohibitive given the earnings in the experiment. If the balance fell below −1000 before period 5, the subject was declared bankrupt. This did not happen in our data.

[17]Since the NCB program assumes no luck or innovation, ICs are identical and the solution is the same whether there are one or two ICs.

[18]If treatments are not independent, this number will be larger. If correlation between treatments is sufficiently low, a positive difference for all eight treatments would still be significant at standard levels.

[19]Under Cost Caps, only 4 of 8 treatment cells (1 of 4 for the HQ and 3 of 4 for the ICs) show higher payoffs for the experienced groups. For MCCS, experienced payoffs are higher for all 8 treatment cells.

[20]Another compelling argument comes from data collected from a version of MCCS with only one contract, denoted Single Contract Cost Sharing (SCCS). In a working paper version of this manuscript we show that the payoffs of the SCCS system are intermediate between MCCS and Cost Caps. This indicates that the one source of benefit of the MCCS stems directly from the approximate incentive compatibility of the menu of missions, and not simply budget flexibility or risk sharing.

[21]In the unreported sessions with a single contract, innovation frequency was between that of the Cost Cap design and the MCCS design.

[22]With a single contract system, only 58.8% of contracts were awarded to the IC with the lowest cost estimate.

[23]This is reminiscent of the study of communication and complexity in mechanism design. For example, Green and Laffont [12] examine the efficiency loss in implementation when the dimension of the message space is necessarily limited. See also Deneckere and Severniov [11].

[24]One should be careful about the magnitude of the stakes, however. Although a mission may *cost* NASA over a billion dollars to build, the *rewards* realized by the individual agents if the mission succeeds are likely far smaller.

[25]In addition to the data reported in this paper, we conducted one session with senior managers from NASA Headquarters and from the two largest Implementing Centers. Six periods of data were generated, of which three were under the Cost Cap process and three were under the MCCS process. In this session, HQ payoff averaged 753 under Cost Cap and 900 under MCCS, and IC payoffs averaged 1086 under Cost Cap and 1400 under MCCS.

References

[1] Bajari, P., and S. Tadelis (1999), “Incentives versus Transaction Costs: A Theory of Procurement Contracts,” The RAND Journal of Economics, 32(3), pp. 387–407.

[2] Banks, J., J. Ledyard, and D. Porter (1989), “Allocating Uncertain and Unresponsive Resources: An Experimental Approach”, The RAND Journal of Economics, 20(1), pp. 1–25.

[3] Binmore, K., and P. Klemperer (2002), “The Biggest Auction Ever: The Sale of the British 3G Telecom Licenses”, The Economic Journal, 112, pp. C74–C96.

[4] Bohm, P., and B. Carlen (1999), “Emission Quota Trade Among the Few: Laboratory Evidence of Joint Implementation Among Committed Countries”, Resource and Energy Economics, 21(1), pp. 43–66.

[5] Brewer, P. J., and C. R. Plott (1996), “A Binary Conflict Ascending Price (BICAP) Mechanism for the Decentralized Allocation of the Right to Use Railroad Tracks”, International Journal of Industrial Organization, 14(6), pp. 857–886.

[6] Cason T., and R. Tenorio (2002), “To Spin or Not To Spin? Natural and Laboratory Experiments from ‘The Price is Right”’, The Economic Journal 112, pp. 170–195.

[7] Chen, Y. and T. Sönmez (2005), “School Choice: An Experimental Study”, Journal of Economic Theory, forthcoming.

[8] Cooper, D., J. Kagel, W. Lo, and Q. Gu (1999), “Gaming Against Managers in Incentive Systems: Experimental Results with Chinese Students and Chinese Managers”, American Economic Review, 89(4), pp. 781–804.

[9] Cox, J. C., R. M. Isaac, P. Cech, and D. Conn (1996), “Moral Hazard and Adverse Selection in Procurement Contracting”, Games and Economic Behavior, 17, 147–176.

[10] Davis, D. and B. J. Wilson (2005), “Differentiated product competition and the Antitrust Logit Model: an experimental analysis”, Journal of Economic Behavior & Organization, 57(1), pp. 89–113.

[11] Deneckere, R., and S. Severinov (2003), “Mechanism Design and Communication Costs”, Fuqua School of Business, Duke University manuscript.

[12] Green, J. R., and J.-J. Laffont (1987), “Limited Communication and Incentive Compatibility”, in Information, Incentives, & Economic Mechanisms, ed. by Groves, T., R. Radner, and S. Reiter, University of Minneapolis Press, Minneapolis.

[13] Grether, D., R. M. Isaac, and C. R. Plott (1981), "The Allocation of Landing Rights by Unanimity Among Competitors", American Economic Review, 71(2), pp. 166–171.

[14] Haruvy, E., A. E. Roth, and M. Ünver (2005), "The Dynamics of Law Clerk Matching: An Experimental and Computational Investigation of Proposals for Reform of the Market", Journal of Economic Dynamics and Control, forthcoming.

[15] Healy, P. and C. Noussair (2003) "Bidding Behavior in the Price is Right Game: An Experimental Study", Journal of Economic Behavior and Organization, forthcoming.

[16] Holt, C. and S. Laury (2002) "Risk Aversion and Incentive Effects", American Economic Review, forthcoming.

[17] J. Hong and C. R. Plott (1982) "'Rate filing policies for inland water transportation: an experimental approach", Bell Journal of Economics, 13, pp. 1–19.

[18] Kachelmeier S. and M. Shehata (1992) "Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence from the People's Republic of China," American Economic Review, 82(5), pp. 1120–1141

[19] King R., V. Smith, A. Williams, and M. Van Boening (1993), "The Robustness of Bubbles and Crashes in Experimental Stock Markets," in Nonlinear Dynamics and Evolutionary Economics, I Prigogine, R. Day and P. Chen eds., Oxford University Press.

[20] Laffont, J, and J. Tirole (1986), "Using Cost Observation to Regulate Firms", Journal of Political Economy, 94, pp 614–641.

[21] Laffont, J, and J. Tirole (1993) A Theory of Incentives in Procurement and Regulation, The MIT Press, Cambridge, MA, USA.

[22] Ledyard, J., D. Porter and R. Wessen (2000) "A Market-Based Mechanism for Allocating Space Shuttle Secondary Payload Priority", Experimental Economics, Vol. 2, No. 3, pp. 173–195.

[23] Ledyard, J., C. Noussair, and D. Porter (1996), "The Allocation of a Shared Resource Within an Organization", Economic Design, Vol. 2, No. 2, pp. 163–192.

[24] McAfee R. P. and J. McMillan (1986) "Bidding for Contracts: A Principal Agent Analysis, RAND Journal of Economics, 17, 326–338.

[25] Muller, R. A., and S. Mestleman (1998), "What have we learned from emissions trading experiments?", Managerial and Decision Economics, 19, pp. 225–238.

[26] Noussair, C., and D. Porter (1992), "Allocating Priority with Auctions: An Experimental Analysis", Journal of Economic Behavior and Organization, pp. 169–195.

[27] Polk, C. (1994) "Subcomponent Innovation and Moral Hazard: Where Technological Progress Meets the Division of Labor", California Institute of Technology Social Science working paper #895.

[28] Reichelstein, S. (1992) "Constructing Incentive Schemes for Government Contracts: An Application of Agency Theory," The Accounting Review, pp. 712–731.

[29] Roth, A. E. and E. Peranson (2002), "The Redesign of the Matching Market for American Physicians: Some Engineering Aspects of Economic Design", Game Theory in the Tradition of Bob Wilson, Chapter 16.

[30] Roth, A. E., T. Sönmez, and M. U. Ünver (2004), "Kidney Exchange", Quarterly Journal of Economics, 119, pp. 457–488.

[31] Smith, V., and J. Walker (1993), "Rewards, Experience and Decision Costs in First-Price Auctions", Economic Inquiry, 31, 237–244.

[32] U.S. Army (2001) "Innovation in Contractual Incentives: Phase 1 Report."

Mission	a_{jl}	b_{jl}	z_{jl}	$\underline{S}_{jl}$	F_{jl}	C^E_{jl}	y_{jl}
A1	.003200	60	1.010	500	1500	980	500
A2	.004375	20	1.005	400	1200	760	0
A3	.004750	30	1.005	200	600	280	0
B1	.001400	60	1.010	1000	3000	1580	500
B2	.002400	20	1.005	500	1500	660	0
B3	.007500	30	1.005	200	600	390	0

Table 1: Cost parameters for each design of each mission.

		Low Variance			High Variance		
Mission	Contract	C^k/C^B	α^k/C^B	β^k	C^k/C^B	α^k/C^B	β^k
$A1$	$k = H$	1.204	1.204	0.916	1.306	1.306	0.928
	B	1.000	1.0375	0.633	1.000	1.0375	0.755
	L	0.796	0.946	0.349	0.694	0.844	0.433
$A2$	H	1.066	1.066	0.872	1.263	1.263	0.958
	B	1.000	1.015	0.544	1.000	1.0375	0.715
	L	0.934	0.994	0.216	0.737	0.887	0.473
$A3$	H	1.172	1.172	0.883	1.345	1.345	0.911
	B	1.000	1.0375	0.565	1.000	1.0375	0.783
	L	0.828	0.978	0.248	0.655	0.805	0.474
$B1$	H	1.127	1.127	0.853	1.316	1.316	0.902
	B	1.000	1.025	0.605	1.000	1.0375	0.763
	L	0.873	0.973	0.388	0.684	0.834	0.445
$B2$	H	1.076	1.076	0.885	1.455	1.455	0.938
	B	1.000	1.0125	0.670	1.000	1.0375	0.835
	L	0.924	0.974	0.405	0.545	0.695	0.553
$B3$	H	1.128	1.128	0.905	1.256	1.256	0.874
	B	1.000	1.025	0.610	1.000	1.0375	0.708
	L	0.872	0.972	0.315	0.744	0.894	0.361

Table 2: MCCS contract parameters as a function of C^B.

Number of Centers	Variance of Cost Shocks	Cost Caps Inexper.	Cost Caps Exper.	MCCS Inexper.	MCCS Exper.
1	Low	5	10	30	5
	High	30	15	20	15
2	Low	20	5	25	5
	High	20	10	20	5
Total		75	40	95	30

Table 3: Number of periods collected per treatment.

	# of Centers	Cost Variance	Cost Caps Inexper.	Cost Caps Exper.	Cost Caps N.C.B.	MCCS Inexper.	MCCS Exper.	MCCS C.B.
Average HQ Payoff	1	Low	748	768	672	802	863	860
		High	730	778	672	800	874	860
	2	Low	746	524	672	805	883	860
		High	767	777	672	962	1000	860
	Average		745	744	672	836	894	860
Average IC Payoff	1	Low	1111	950	1160	1134	1297	1255
		High	1021	963	1160	1133	1345	1255
	2	Low	908	777	1160	1106	1398	1255
		High	997	1007	1160	1383	1569	1255
	Average		991	948	1160	1179	1383	1255

Table 4: Average period payoffs to the HQ and the ICs in each treatment versus the No-Carryover Benchmark (N.C.B.) and the Carryover Benchmark (C.B.).

	Number of ICs	Variance of Cost Shocks	Cost Caps		MCCS	
			Inexp.	Exper.	Inexp.	Exper.
Percent missions canceled	1	Low	17.5%	10.0%	4.7%	0.0%
		High	24.0	23.1	3.3	0.0
	2	Low	16.8	50.0	2.6	0.0
		High	25.6	25.0	0.0	14.3
	Average		22.1%	23.7%	2.9%	3.0%
Final cost vs. initial estimate (C_F/C_j^E)	1	Low	1.05	1.20	0.98	1.11
		High	1.06	1.04	1.12	0.80
	2	Low	1.17	1.34	0.99	1.01
		High	1.05	1.06	0.86	0.95
	Average		1.08	1.12	0.99	0.91
Reliability at launch (R_j)	1	Low	94.5%	99.3%	96.5%	98.1%
		High	99.2	95.9	94.0	98.2
	2	Low	96.3	95.4	92.7	97.1
		High	98.3	97.5	95.1	96.1
	Average		97.9%	97.1%	94.7%	97.7%
Number of innovations per period	1	Low	0.41	1.40	1.17	1.40
		High	0.97	0.73	1.41	1.60
	2	Low	0.60	0.80	1.00	1.00
		High	0.90	0.10	1.10	1.60
	Average		0.81	0.75	1.16	1.47

Table 5: Summary of results for all periods.

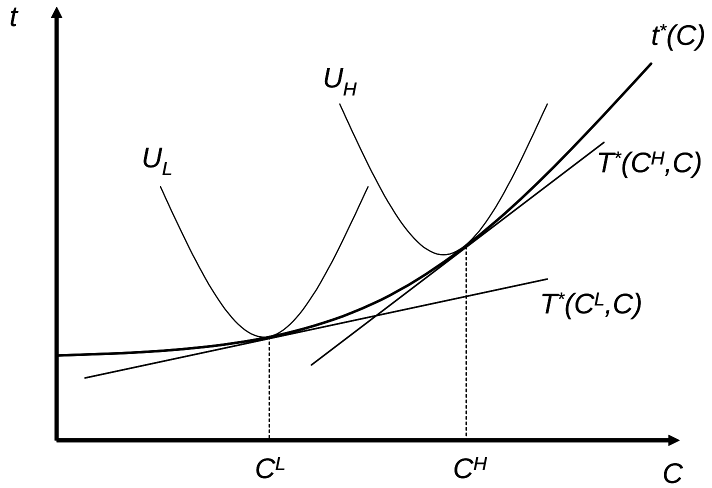

Figure 1: The convex contract $t^*(C)$ and each linear contract $T^*(C^E, C)$ are incentive compatible.

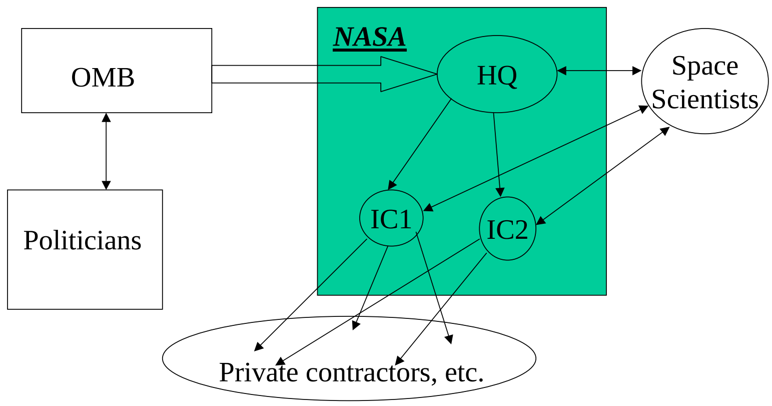

Figure 2: The NASA Mission Acquisition Environment.

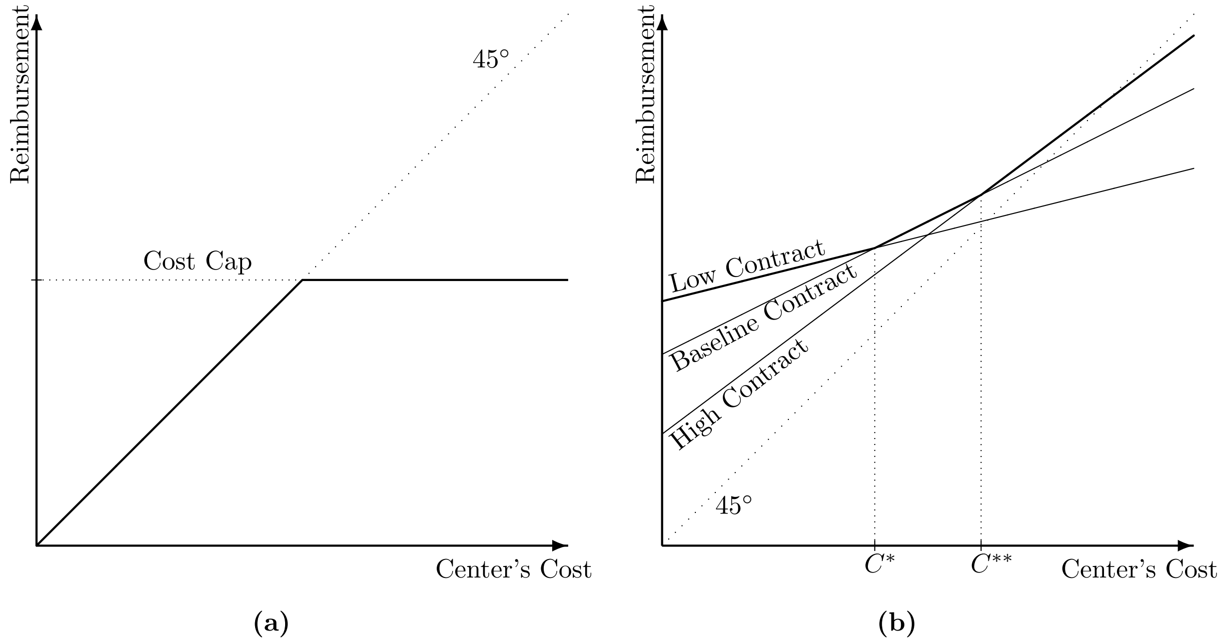

Figure 3: (a) Cost cap reimbursement schedule. (b) MCCS reimbursement schedule.

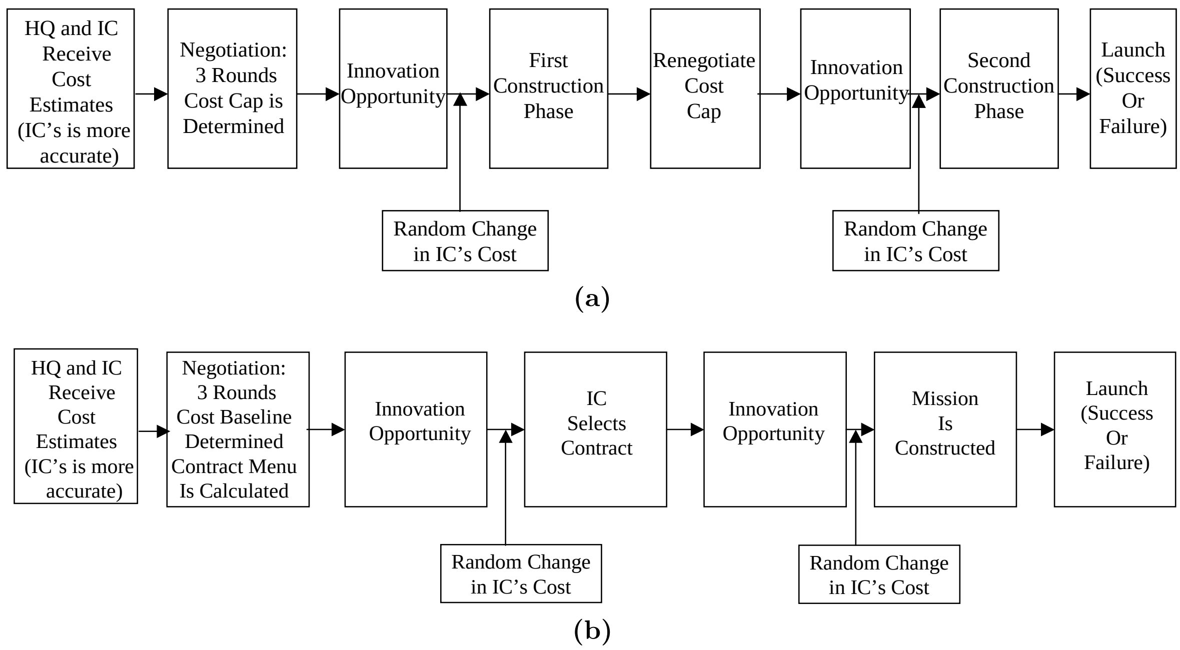

Figure 4: Timing of (a) the Cost Cap process and (b) the MCCS process.