Are Experienced Managers Experts at Overcoming Coordination Failure?

November 29, 2005

David J. Cooper[*]
Associate Professor of Economics
Weatherhead School of Management
Case Western Reserve University
11119 Bellflower Road
Cleveland, OH 44106
ph: 216-368-4294
email: david.cooper@case.edu

Abstract: This paper studies experiments set in a corporate environment where a manager attempts to overcome a history of coordination failure by employees using either financial incentives or communication. I compare the choices of subject managers drawn from a standard undergraduate population with subject managers drawn from the executive MBA (EMBA) program at Case's Weatherhead School of Management. The EMBA subjects are a group of experienced, successful managers; all of the EMBA subjects have at least ten years of work experience, including at least five years in a supervisory role, and have average annual earnings in excess of $120,000. The EMBA subject managers are able to overcome a history of coordination failure significantly faster than the undergraduate subject managers. This superior performance is driven neither by differences in the financial incentives offered to the employees nor by use of an inherently different communications strategy. Instead, EMBA subject managers are significantly more likely to use the same "good" communication strategy as is identified for undergraduate subject managers through systematic coding of manager's messages to employees.

Keywords: Coordination, Communication, Experiments, Executives, Expertise

JEL Classification Codes: C92, D23, J31, L23, M52

If getting employees to coordinate their efforts is an important issue in many workplaces, experienced, successful managers ought to be good at solving this problem. Both selection and learning arguments support this assertion. Given that promotion as a manager relies in large part on outperforming other managers who are in similar positions, managers who are inherently good at getting employees to coordinate should have an advantage in the promotion tournament and hence should be overrepresented in the population of successful managers. Even if individuals who become successful managers are not inherently better than the general population at solving coordination problems, their experience should give them an advantage. Repeatedly facing situations where success depends on getting employees to coordinate their efforts gives managers the incentive and opportunity to learn what approaches will work well for overcoming a history of coordination failure. Regardless of whether selection or learning is the cause, if seasoned managers are experts at fostering coordination among employees, studying their behavior provides a valuable window into how coordination problems are solved in corporate settings. This paper therefore presents artefactual field experiments[1] comparing the behavior of experienced, successful managers with that of naïve subjects, the standard undergraduate population, in a laboratory environment where success depends on fostering coordination among employees.

Specifically, I study experiments set in a corporate environment called the turnaround game. Managers' and employees' payoffs in the turnaround game depend positively on employees coordinating at high effort levels. The underlying game being played by employees, the weak-link game, represents an especially challenging coordination problem. As an archetypical example, consider a firm that produces output via an assembly line. One step of the production process cannot be started until the previous one has been completed. Production can therefore progress no faster than the slowest worker on the assembly line.[2] Even though all workers might benefit from a high level of productivity, no worker will be willing to work hard if he believes other workers will not. The resulting weak link game confronts a manager with a challenging problem. To overcome a history of coordination failure, she must find some way to get all workers to *simultaneously* increase their effort levels. Unilateral changes are likely to fail as they leave productivity unchanged while forcing employees to bear the costs of greater effort. Indeed, there exists an extensive experimental literature illustrating the difficulties of coordination in weak-link games.[3]

In an initial phase of the turnaround game, where managerial intervention is absent and incentives to coordinate are low, employees almost invariably slip into coordination failure. An additional subject is then introduced as a manager. This subject manager, faced with overcoming a history of coordination failure, can employ either financial incentives or communication as instruments of change.[4] I compare the choices of subject managers drawn from a standard

---

[1] Harrison and List (2004) introduce this term to describe standard economic experiments using non-standard subject populations.
[2] For examples from field settings, see Knez and Simester's (2002) study of Continental Airline's turnaround and Ichniowski, Shaw, and Prennushi's (1997) work on steel production.
[3] See Camerer (2003, ch. 7) for a recent summary of this literature.
[4] For other experiments studying the effects of pre-play communication see Cooper, DeJong, Forsythe, and Ross (1989 and 1992), Van Huyck, Gillette, and Battalio (1992), and Capra, C. M., T. Tanaka, C. Camerer, L. Munyan, V. Sovero, L. Wang, and C. Noussair (2005). These studies differ several ways from the work presented here. In all of these papers except Capra *et al*, the set of available messages is very limited. In none of these studies is there an external agent with a stake in the outcome, like the firm manager in my experiments. None of these studies use experts as subjects.

undergraduate population with subject managers drawn from the executive MBA (EMBA) program at Case's Weatherhead School of Management. The EMBA subjects are a group of experienced, successful managers; all of the EMBA subjects have at least ten years of work experience, including at least five years in a supervisory role, and have average annual earnings in excess of $120,000. The EMBA subject managers overcome a history of coordination failure significantly faster than the undergraduate subject managers. This superior performance is driven neither by differences in the financial incentives offered to the employees nor by use of an inherently different communication strategy. Instead, EMBA subject managers are significantly more likely to use the same "good" communication strategy as is identified for undergraduate subject managers through systematic coding of manager's messages to employees.

These results, confirming the hypothesis that successful, experienced executives will have expertise at solving coordination problems, are valuable for several reasons. The results of these experiments verify the external validity of earlier experiments with the turnaround game that only used undergraduate subjects (Brandts and Cooper, 2005b). The "good" communication strategy employed by successful managers is quite simple: ask for a specific effort level, stress the mutual benefits of coordination, imply that employees are well-paid, and don't implement any long-term plan that upsets the balance between managers and employees. EMBA subject managers are more likely to use this communication strategy, but don't have an alternative path to coordination that is unused by undergraduate subject managers. The list of common messages is almost identical for the two groups of managers and the good communication strategy identified in the current study is virtually identical to the successful communication strategy described by Brandts and Cooper (2005b). It follows that the "good" communication strategy identified in these experiments probably parallels strategies used in field settings to overcome coordination failure.

More broadly, the nature of managers' expertise underlines the importance of communication, consistent with the conclusions of Cooper and Brandts (2005b). When economists study how managers affect their employees' behavior, the focus is usually on factors like financial incentives and monitoring that fit easily into economic models. While these factors are doubtlessly important, the particular nature of coordination problems suggests that communication will play at least as important a role.[5] Coordination problems are fundamentally a problem of beliefs. In a weak link game, an employee will only be willing to increase his effort if he believes that others will increase their effort levels as well. The "good" communication strategy described above succeeds largely as a coordinating device for beliefs, convincing employees that others will be trying to coordinate at higher effort levels in the upcoming round. It is telling that the dimension along which EMBA subject managers differ from undergraduates is in the frequency of good communication strategies. To the extent that both selection and learning bias the population of successful executives towards strategies that are effective, it appears that good communication is a winning strategy in the field.

Finally, laboratory experiments using experienced managers as subjects allow us to infer characteristics about the managers' workplaces that cannot be observed directly. While it seems likely that coordination problems are quite common in workplaces, the existing evidence is episodic in nature. Our EMBA subjects are drawn from a wide variety of industries. The selection and learning effects outlined above only have force when individuals are faced with problems in their workplaces that are related to the coordination problems they face in the

---

[5] Beyond the following theoretical arguments, studies in organizational behavior suggest that communication is one of the crucial variables that influence change (e.g. Ford and Ford, 1995; Kotter, 1996; Foss, 2001).

laboratory. The strong performance of EMBA subject managers therefore suggests that they must frequently face coordination problems in their workplaces. Although general managerial expertise could also underlie the relative success of EMBA subject managers, the good communication strategy that underlies their advantage is quite specific. If the primary problem managers faced was determining how much to pay their employees to prevent shirking, for example, it is unlikely they would have learned to communicate in this particular way. The problems that experienced, successful managers are particularly good at solving are liable to be ones that are important to their success.[6]

The remainder of this paper is organized as follows. Section 1 introduces the corporate turnaround game and develops its theoretical properties. Section 2 describes the experimental design and procedures, emphasizing the differing nature of the undergraduate and EMBA subject pools. Section 3 presents the results. Section 4 synopsizes the paper, discusses the relationship between this paper and the literature on artefactual field experiments with managers, and explores directions for future research.

2. The Corporate Turnaround Game

An experimental firm in the corporate turnaround game consists of four employees and a manager. The structure of the game played by these five subjects is governed by the following pair of basic design choices.

1)   The firm's technology has a weak-link structure.

For many organizations the individual (or unit) doing the worst job – the "weak link" – determines the overall productivity of an organization (Kremer, 1993). Starting from this observation, Knez and Camerer (1994) argue that the game played within many firms takes on the form of the "minimum game" introduced by Van Huyck, Battalio, and Beil (1990). Knez and Camerer refer to these games by the more evocative term "weak-link games," a terminology I follow. These are a type of coordination game. Each player simultaneously chooses a strategy that can be thought of as an effort level. A player's payoff is a decreasing function of his own effort and an increasing function of the minimum effort chosen by the players in the group. Payoffs are set up so that it is worthwhile for a player to raise his effort level if *and only if* it will increase the minimum effort for the group. Coordinating on any of the available effort levels is a Nash equilibrium for a weak-link game, but results from earlier experiments suggest that play usually evolves towards the payoff dominated equilibrium where all players choose the lowest possible effort level. The weak-link game therefore represents a particularly challenging case for coordination.[7]

In each round of the corporate turnaround game, the firm's total surplus is determined by the minimum effort chosen among its four employees. In an initial phase the manager sets a

---

[6] As another example of this reasoning, Fehr and List (2004) find that CEOs are both more trusting and more trustworthy than undergraduate subjects in trust games. To the extent that either selection or learning cause this difference, their result implies that there must be some benefit to CEOs for being more trusting and trustworthy. It seems plausible that these characteristics allow CEOs to build the reciprocity necessary to implement potentially valuable incomplete contracts which otherwise would fail to be incentive compatible. See also Cooper, Kagel, Lo, and Gu (1999) for a related example; they use the behavior of Chinese managerial subjects to infer that the ratchet effect is an important feature of state-owned businesses where these managers worked.

[7] See Van Huyck *et al* (1990) and Knez and Camerer (1994 and 2000). Coordination failure is less likely in small groups – see preceding as well as Weber, Camerer and Knez (2004).

4

bonus rate which governs how the firm's surplus is split between himself and the employees. For all feasible bonus rates, the four employees then play a weak-link game among themselves. Hence, the turnaround game focuses on a worst-case scenario for coordination. If experienced, successful managers are better than undergraduates at overcoming coordination failure in organizations with a weak-link structure, this advantage should also be present in less extreme environments.

2)   Both managers and employees get limited information feedback.
        Specifically, I assume the firm manager observes the minimum effort selected (which is revealed by the firm's productivity), but cannot observe any individual employee's effort level. Likewise, employees observe their own effort and the minimum effort for the firm, but not the individual efforts of the other three employees. For the managers this implies that they lack the necessary information to tailor bonuses to the effort put forth by individuals and can only offer bonuses based on the minimum effort among all employees. In other words, limiting the information available to the firm manager restricts the tools available for overcoming coordination failure. Limiting employees' information gives managers a significantly more difficult task, as previous work demonstrates that overcoming coordination failure is substantially harder when employees cannot observe others' effort levels (Brandts and Cooper, 2005c).
        Limiting the manager's information about employees' choices implies that he, consistent with the spirit of most principal-agent models, has difficulty monitoring them. There is no particular reason to believe that this assumption has either more or less external validity than assuming the manager individually monitors the employees, as both cases occur in reality. The primary benefit of limiting the manager's information is a considerable simplification of the environment as well as confronting the manager with a more challenging task. Limiting the employees' information about each others' choices also accentuates the importance of leadership by the manager. When employees can see the choices of other employees, leading by example often takes place (Brandts and Cooper, 2005a). One or more employees make a large increase in their effort levels presumably in the hope of leading laggards to match this effort, thereby overturning a history of coordination failure. This sort of internal leadership works reasonably well with full feedback, but can't work with limited feedback as laggards can't see the effort choices of putative leaders. Limited feedback therefore leaves managers as the primary source of potential leadership within the firm.
        Turning to the specifics of the turnaround game, a round of the game starts with the manager setting a bonus rate (B) that determines how much additional pay each employee receives per unit increase in the minimum effort. Bonuses are based solely on the minimum effort, consistent with the assumption that the manager cannot observe individual efforts. Feasible bonus rates are restricted to the integers B $\in$ {6,7,8,9,10,11,12,13,14,15}. Given the other parameter values used in these experiments, this is the set of (integer) bonus rates that neither make the choice of positive effort a dominated strategy nor allow the manager to lose money. At the same time the manager selects a bonus rate, he can also send a message to his employees. With minor restrictions, described in Section 3, this message can consist of any text the manager wishes to type. Employees observe B and the manager's message and then simultaneously select effort levels. Intuitively, employees spend forty hours per week on the job, and effort measures the number of these hours that they actually work hard rather than loafing. Employee i's effort, E_i, is restricted to be in ten hour increments: $E_i \in$ {0,10, 20, 30, 40}. The

players' payoffs, denominated in "Experimental Currency Units" or ECUs, are given by the following functions:

(1) Manager payoff: $\pi_F = 100 + \left((60 - 4B) \times \min_{i \in \{1,2,3,4\}} (E_i) \right)$

(2) Employee i payoff: $\pi_e^i = 200 - 5E_i + \left(B \times \min_{j \in \{1,2,3,4\}} (E_j) \right)$

The manager's payoff depends on the minimum effort selected by his employees, consistent with the assumption that the firm's production technol ogy has the weak-link property. As can be seen in Equations 1 and 2, the bonus transfers a portion of the firm's revenues to its employees. Managers are limited to a linear bonus scheme to simplify the game. Any messages sent by the manager are cheap talk and do not enter the payoff function.

Table 1
Employee i's Payoff Table, B = 6

		Minimum Effort by Other Employees				
		0	**10**	**20**	**30**	**40**
	0	200	200	200	200	200
Effort	**10**	150	210	210	210	210
By	**20**	100	160	220	220	220
Employee i	**30**	50	110	170	230	230
	40	0	60	120	180	240

Employee i's Payoff Table, B = 14

		Minimum Effort by Other Employees				
		0	**10**	**20**	**30**	**40**
	0	200	200	200	200	200
Effort	**10**	150	290	290	290	290
By	**20**	100	240	380	380	380
Employee i	**30**	50	190	330	470	470
	40	0	140	280	420	560

To develop the basic theoretical properties of the turnaround game, consider first the proper subgame where employees choose effort levels after the manager has chosen a bonus rate. For all available values of the bonus rate the resulting subgame is a weak-link game. Coordinating on any of the five available effort levels is a Nash equilibrium. Given this multiplicity of equilibria in the subgames, subgame perfection has little predictive power for the

6

full game. In particular, any combination of bonus rate and minimum effort can be supported by a subgame perfect equilibrium.

To understand why overcoming coordination failure is so difficult in this environment, consider the game induced by a bonus value of B = 6 (see top panel of Table 1). Suppose that the employees have previously all chosen effort level 0. An employee who thinks about raising his effort from 0 to 10 faces a certain payoff reduction of 50 ECUs due to increased effort, while his maximum possible gain is only 10 ECUs beyond the 200 ECUs he gets without risk by choosing 0. For the proposed increase to have a positive expected profit, the employee must believe the probability of the three other employees *simultaneously* raising their efforts from 0 to 10 equals at least 5/6. Treating the other three employees as statistically independent, this translates into requiring a 94% chance of increased effort for each of the other three employees.[8] Given these grim incentives, overcoming coordination failure is unlikely with a low bonus rate.

Now imagine that a new manager takes over the firm. Determined to shake the firm out of its underperforming ways, she raises the bonus rate to B = 14, the highest bonus rate at which the firm manager can earn a profit. This yields the payoff table shown in the bottom panel of Table 1. The incentives for employees to increase effort are strengthened by this change. Again suppose we start with all four employees choosing effort level 0. An employee increasing his effort from 0 to 10 still faces a certain loss of 50 ECUs, but the potential gain is now 90 ECUs. The break-even probability that the other employees will simultaneously increase their efforts is now 5/14. Assuming the other three employees are independent, this translates into requiring a 71% chance that each employee raises his effort. These are better odds than we saw with B = 6, but still daunting. Although one can imagine employees now at least attempting to overcome coordination failure, there is clearly room for the manager to play a leadership role beyond increasing the bonus rate.

3. Experimental Design and Procedures

The focus of this experimental design is on how the choices and resulting performance of naïve and expert managers will differ in overcoming a history of coordination failure. As such, the nature of the subject population is of particular interest. For naïve managers I recruited from the undergraduate population at Case Western Reserve University.[9] The expert managers were all recruited from the executive MBA population at Case's Weatherhead School of Management. Throughout this paper, when I refer to a "manager," I am describing the subject's role in the game, not whether they were an undergraduate or an executive MBA student. The abbreviations "EMBA" and "UG" will be used to distinguish executive MBA and undergraduate subjects respectively.

While Case's undergraduate and EMBA programs are both highly ranked,[10] the characteristics of their students are quite different. These are summarized in Table 2. For undergraduates, statistics reported in Table 2 only reflect subjects who were assigned the role of

---

[8] To derive this probability, solve for p such that $200 = 150*(1-p^3) + 210*p^3$.
[9] Results from the undergraduate managed firms have been reported previously in Brandts and Cooper (2005b) as part of the dataset for their one-way communication treatment.
[10] US News and World Report ranked Case 35th in its 2005 rankings of America's best undergraduate programs (Sklaroff, 2004). In the most recent Business Week rankings of EMBA programs, done in 2003, Weatherhead's program ranked 21[rst] in the world (Merritt, 2003)

manager, as this is the relevant group for comparison with the EMBAs (who were all assigned the role of managers).

Table 2
Description of Subject Populations

	Undergraduate Manager	EMBA Manager
Number of Firms	20	19
Demographic Information		
Number of Women	13	3
Average Age	18.7	38.2
Number of Subjects with Post-graduate Degree	0	3
Features of the Data		
Firms with Missing Messages	4	0
Firms in Round 10 Coordination Failure	16	15

Case's undergraduate program does not admit students on a part-time basis and almost exclusively has traditional college students who enter out of high school and plan to complete their degrees within four or five years – for the most recent cohort with complete data (admitted in Fall 1998), 78% graduated within five years. The UG managers are all young with an average age of 18.7 years and a maximum age of 22.[11] The crucial feature of the UG managers, deriving from their youth and status as full-time students, is a lack of substantial work or managerial experience.

The EMBA subjects are quite different from the undergraduate population. Weatherhead's EMBA program is designed primarily to appeal to mid-career managers who want to move up into senior management or start their own firm.[12] Students come from a wide variety of industries, with manufacturing being by far the most common background, consistent with the Midwestern student body, but including heavy representation from the healthcare and service industries. As a condition for admission to the program, potential students must have a "minimum of 10 years relevant work experience, with at least five years involving substantial managerial responsibilities."[13] With rare exceptions, it is expected that EMBA students will be

---

[11] In recruiting experimental subjects at Case, we rely heavily on announcements made in large lecture classes. As these classes are taken disproportionately by first and second year students, the resulting subject population tends to be younger than the full undergraduate population.

[12] For the most recent academic year, only 6% of the EMBA students listed "CEO" or "Owner" as their current managerial position.

[13] According to Betty Vandenbosch (email, June 6, 2005), the associate dean for Weatherhead's EMBA program, the work experience requirement is relaxed only in rare cases for individuals with 8 or 9 years of work experience who "run companies or own them." The definition of "substantial" is somewhat subjective, but requires some sort of supervisory responsibility.

currently active managers. Unlike the UG managers, EMBA managers all have substantial work and managerial experience. Judging by their earnings, the EMBA managers are also successful; the average annual salary of EMBA students was $122,000 in the most recent academic year.

While it would be nice to attribute any differences between the performance of UG and EMBA managers to differences in work and managerial experience, several other possible causes must be noted.

1.   EMBA managers are almost twenty years older on average than the UG managers. Thus, any differences could be due to age independent of any managerial experiences – forty year old auto mechanics might perform just as well in the manager role as forty year old executives. A useful feature of the data is that EMBA ages vary widely from a minimum of 31 years to a maximum of 49 years. Assuming that the marginal effect of age on managerial performance is roughly constant, econometric controls can therefore be used to separate the effects of age from those of managerial experience.

2.   The proportion of men is much higher among the EMBA managers than UG managers. This scarcity of women reflects the male dominated EMBA population. For the cohort (class of '04) that EMBA managers were drawn from, only 9 of 42 students were women. Econometric controls will be used to separate gender effects from those of managerial experience.

3.   EMBA managers have taken classes that might affect their performance. The EMBA program at Case emphasizes leadership training with students required to take multiple courses on "leadership assessment and development." Given that leadership training is perfectly correlated with being an EMBA manager, the data cannot separate between effects due to managerial experience and effects due to leadership training. This is an interesting issue for future research.

4.   As noted previously, the EMBA students are not a randomly selected sample; by definition, all of the EMBA managers are individuals who chose a career in management, were sufficiently successful to qualify for the EMBA program, and chose to pursue an EMBA degree. Econometric controls can imperfectly account for the first stage of selection by using the undergraduates' majors as a proxy for how likely they are to enter management, but I have no means of controlling for the remaining types of selection. Selection is a secondary issue in analyzing this dataset since the focus is on studying whether (and how) EMBA managers act differently than UG managers, but it is of obvious interest to determine whether the observed differences are due to the inherent nature of the individuals or to their experiences as managers. In the conclusion I describe how these effects will be disentangled in future work.[14]

The recruiting of undergraduates was done in a standard fashion via posters, class sign-ups, and e-mails. The advertisement was for a generic economic experiment with no inducement offered for participation beyond the cash payoff. While this recruiting strategy generated a more than sufficient supply of undergraduate subjects, it would have been unlikely to work with EMBA students. All of the EMBAs have full-time jobs in addition to their classes. They spend a single day per week on campus, with all of this time devoted to classes or class-related activities. With average annual earnings in the low six-figure range, EMBA students don't

---

[14] On one potential confound I got lucky; only three EMBA managers have a post-graduate degree. It is therefore unlikely that any effects due to greater educational levels are driving the results.

especially need the average earnings of $25 featured in the advertisements. To put things plainly, EMBA students face severe time constraints and value ninety minutes of time at a far higher rate than any show-up fee I can possibly pay.[15] This is an extremely difficult population from which to recruit. Indeed, a second set of sessions in which I intended to run a follow-up treatment had to be abandoned when insufficient EMBAs agreed to participate.[16] In response to this challenge, I adopted both a non-standard recruiting strategy and some non-standard procedures. These are designed to make participation easier and more attractive for EMBA students. While my strong preference is to use completely standard recruiting techniques and experimental procedures, it is unrealistic to think these experiments could have been run without making some allowance for the difficulties of recruiting EMBA subjects.

All EMBA recruiting took place through an announcement in James Rebitzer's "Economics of Organizations" course. This is a required second year class for Case's EMBA program. It meets once per week. Students study incentive design in the first half of the course, but prior to the experiment the sole economic framework that students had discussed was a simple principle-agent model only loosely related to the turnaround game. Before I made my recruiting pitch, Dr. Rebitzer gave a brief introduction on game theory and experimental economics to the EMBA students.[17] I then gave a short presentation to the EMBA students describing the experiment in generic terms. To encourage participation, I stressed the value of their time in furthering research at the business school. This was intended to appeal to the strong esprit-de-corps among the EMBA students, although my impression is that it had little effect on their participation decisions. I also emphasized that they could gain valuable insights from participating and then joining in a feedback session to be held in class the following week.[18] In doing this, I tried to draw on their curiosity; contrary to stereotypes, EMBA students have a high degree of intellectual curiosity about issues relating to management. My feeling from talking with the subjects at the feedback session is that curiosity and/or the desire to learn were the primary motivations that led to participation. The monetary incentives were described, although highlighted less than is typical in recruiting undergraduate experimental subjects since financial inducements were (correctly) not expected to play an important role in the participation decisions of the EMBAs. I stressed that participation was easy (for reasons that will become clear when the procedures are described) and that most subjects enjoyed playing the games. Dr. Rebitzer encouraged the students to sign up for the experiments, but stressed that participation was *not* a required component of the course. Nineteen EMBA subjects agreed to participate, slightly less than half of the class.

The sessions with EMBA subjects were run in a single week between meetings of their class.[19] This was done to reduce the possibility of contamination across sessions. Sessions were

---

[15] It should be noted that the EMBA subjects seemed quite happy to be paid. Rather than thinking that EMBAs don't care about monetary incentives, it would be more accurate to state that the sunk cost of participating is generally larger than the purely monetary incentives I can offer.

[16] I wanted to run sessions in which the managers cannot communicate with their employees. Only three EMBAs in the class of 2005 indicated their willingness to participate. This failure was in part due to the class of 2005 being less than half as large as the class of 2004.

[17] The content of this introduction had little to do with the experiments. Students were given a broad non-technical description of what game theory studies and were told that experimenters are playing an increasingly important role in the development of game theory.

[18] I was very careful when describing the educational benefits of the experiments to not specifically link the experiments to any of their classroom materials.

[19] A single session with undergraduate managers only had been run in the previous week to test the software and a final session was run the following week with undergraduate managers only to balance the data set.

run in the evening to lessen work conflicts for the EMBA subjects. The software used to run the experiments is a web-based application. Undergraduate subjects came to a temporary laboratory on campus to participate. EMBA subjects participated via the internet from their homes. This was done to save them the time of coming on to the campus, a major issue in recruiting sufficient EMBA subjects. Instructions on logging into the system and using the software were sent to EMBA subjects the afternoon of their session. The use of participation via the internet raises several issues with the experimental controls. First, a number of the EMBA students were using unreliable dial-up connections. All EMBA subjects were given a helpline number they could call if they had technical problems or questions about the instructions, and several had extended conversations with my programmers about how to restore their connections to the internet. Two EMBA subjects had to quit the experiment early when their internet connections failed irretrievably. For other EMBA subjects who experienced problems of a less serious nature, their poor connectivity may well have distracted them and reduced their performance relative to other managers. Second, because the EMBA subjects were not under direct observation, it cannot be stated definitively that they participated alone. However, this seems unlikely to have been a serious problem. The EMBA subjects were invariably connecting from home offices and none of my follow-up discussions with them indicated that anyone else had been involved in their decision making. A more serious problem is that EMBA subjects were often multi-tasking while participating in the experiment. For example, one subject was packing for a business trip while playing the game. The distractions resulting from multi-tasking probably reduced the performance of EMBA managers as compared to UG managers. To the extent that I find that EMBA managers outperform UG managers, having EMBAs participating from remote locations likely reduced this effect.

At the beginning of a session, UG subjects were randomly assigned to either the role of an employee or a manager. All EMBA subjects were used as managers. As this implies, all subjects in the employee role were undergraduates. I aggressively over-recruited UG subjects to guarantee that any EMBA subject would have four UG subjects available to act as his/her employees. To make the experimental conditions parallel for UG and EMBA managers, UG managers were always physically separated from the undergraduates assigned to the employee role.[20] This makes it impossible for employees to identify who is their manager, regardless of whether the manager was participating on campus or via the internet.[21] Separating UG managers and employees also eliminates the possibility of non-verbal communication by UG managers that is unavailable to the EMBA managers. All sessions with EMBA managers also included UG managers. For all sessions, regardless of whether any EMBA managers were participating, all subjects were told that some subjects might be participating from remote locations. Subjects were not told anything about the characteristics of subjects who were participating via the internet.

Prior to the beginning of play, subjects received instructions (see appendix) and payoff quiz questions for both roles. Writing the instructions raised a delicate methodological issue. It is quite possible that EMBAs' expertise is context dependent. The literature on artefactual field

---

[20] When possible, UG managers were seated in a separate room from their employees. In some cases, the experiments were being run in large halls without any small room available nearby. In this case I seated the managers separately from the employees with a space of about fifty feet between the two groups and avoided as much as possible letting employees have a direct line of sight to the managers.

[21] When managers and employees are seated in close proximity, employees can guess who their manager is by seeing who is typing just before they receive a message.

experiments contains a number of examples in which the effect of expertise depends on whether experiments are framed in a way that mimics the relevant field setting. For example, Cooper *et al* (1999) find that the ability of textile and pharmaceutical plant managers to play strategically in a ratchet effect game was substantially improved by the use of a naturalistic context as opposed to an abstract context.[22] Since I wanted EMBAs to draw on their expertise as much as possible, this argued for instructions that attempt to mimic a workplace setting. However, hewing too close to the field setting could unintentionally draw on the strong emotional connotations of certain terms and lead to demand induced effects. For example, this experiment could have been framed in terms of employees "working hard" or "being lazy," but these terms would have strongly suggested to subjects how they ought to behave. To balance these concerns, instructions were framed in neutral but naturalistic terms. Subjects were told that in the role of employees they must allocate 40 hours a week, in ten hour blocks, between Task A and Task B. Task A is equivalent to effort, but this term, due to its strong connotations, was never used when communicating with subjects through the instructions or otherwise. Subjects knew the number of rounds in the experiments and knew in advance when the bonus would be set by the computer and when it would be set by another subject acting as manager. Subjects were reminded of this information throughout the course of the experiment.

In the course of running sessions, I discovered that many of the EMBA managers had substantial difficulty making it through the instructions and payoff quiz. This caused a long delay for a number of subjects between the time the experiment began and the time that actual play commenced. Although only one EMBA subject quit early because the experiment was running long, I was concerned that data from the later rounds for EMBA managers was unreliable because subjects began rushing to get the experiment finished.[23] Initially I responded by sending the EMBA subjects more complete instructions on how to use the software as well as urging EMBA subjects to call the hotline when they were confused. This improved matters, but to further speed up the process, for the final session (with five EMBA subjects) I emailed the EMBA subjects a copy of the instructions the afternoon of the experiment so they could review them and ask questions in advance. This completely eliminated the delays. Looking at the data, there is no obvious difference between EMBA data collected in early sessions and that collected in later sessions. This indicates that the changes in procedures designed to speed up the sessions had no impact on the choices of EMBA managers.[24]

The experimental design assumes that managers must overcome a history of coordination failure. Therefore, the experiments are designed so that employees will usually fail to coordinate prior to the possibility of managerial interventions. Subjects played thirty rounds of the turnaround game in fixed groups ("firms") of five: one manager and four employees. For the first ten rounds of the experiment the manager was strictly an observer; managers could see the same round by round feedback that they normally received, but could neither control the bonus rate nor communicate with employees. Managers were not paid for these rounds, although employees and managers were both shown the payoffs that the manager would have earned. The bonus rate was fixed at B = 6 for the first ten rounds, this being the lowest (integer) bonus rate

---

[22] See Harrison and List (2004) for a lengthy discussion of this issue.

[23] UG managers tended to play faster than EMBA managers, and no firms headed by UG managers faced serious time pressure in the later rounds.

[24] Given that the EMBA subjects knew each other, I was concerned about contamination between sessions. As noted above, one of the primary reasons for running the experiments in a single week was to limit the possibility of contamination, as most interactions between EMBAs take place when they are on campus for classes. Given the stability of the data over time, it seems unlikely that any contamination took place.

that does not make choosing positive effort levels a dominated strategy. The goal was to get most firms coordinated on the inefficient outcome with minimum effort equal to zero.

For the remaining twenty rounds the manager actively managed his firm. The employees were informed when the manager took over control of the firm. In all treatments the manager was then responsible for choosing a bonus rate in each round and received payoffs as shown in Equation 1. At the same time that they were asked to choose a bonus level, managers were given a text box in which they could type a message. Subjects were given no instructions about the content of the messages except that they could not identify themselves or use offensive language. Although we did not monitor messages in real time, our ex post reading of them indicates that subjects universally abided by these restrictions. Subjects were given no time or length limit on entering messages. Indeed, some of the messages were quite long and took a lengthy time to type. Once the manager was finished, whatever message he wrote was sent to all the employees in the firm at the same time they saw his choice for the bonus rate.

Payoffs were converted from ECUs to dollars at a 500:1 ratio. Average payoffs, including a $10 show-up fee were $23.89. A concern in designing these experiments was that the payoffs would seem far more salient to UG managers than their wealthy EMBA counterparts. In attempt to make the payoffs more salient for the EMBA subjects by drawing on their competitive nature, all managers, UG and EMBA, were told in the instructions that subjects playing as managers would be given access to a list giving the payoffs for all participants in this role. This list would be ordered from the highest to the lowest payoff, but would not include any identifying information about the subjects. Subjects were told that this information would only be made available after all the relevant sessions were completed. The existence of this list was not mentioned in the recruiting materials. My strong impression from post-experiment surveys and the feedback session is that the prospect of learning their relative performance had little effect on either UG or EMBA managers' choices.[25] In retrospect, this is unsurprising for several reasons: the instructions spend pages talking about the monetary payoffs while the list is only mentioned in one brief paragraph, all of the feedback subjects receive during the experiment focuses on the immediate payoffs, and the rankings are private information. It therefore appears likely that in spite of our best efforts the payoffs were more salient for the UG managers than the EMBA managers.

As part of subjects' post-experiment instructions I strongly encouraged subjects to not discuss the specifics of the experiment with others. This was intended to limit the possibility of contamination across sessions. In a post-experimental survey I asked subjects if any previous participant had discussed the details of the experiment with them. The lack of positive responses indicates that the vast majority of subjects complied with my request to not share their experiences. Data from the one firm where the UG manager had talked with an earlier participant about how to successfully play as a manager has been excluded from the analysis.

For the first session, containing four firms with UG managers, the messages were not saved due to a software bug. The software otherwise worked correctly for this session, so all the data on subjects' choices that was saved has been included in our dataset, but obviously data from this session is not reflected in any content analysis of managers' messages.

As part of the experimental design for Brandts and Cooper (2005b), sessions were run using a no-communications treatment. This treatment is identical to the experiment described

---

[25] This was especially dramatic for the UG managers. I delayed posting the list for UG subjects by months in case more sessions were needed. In all this time, none of the UG managers ever asked about how they had done relative to others, indicating that they didn't care.

above, except the firm manager was not allowed to send messages to the employees. In other words, the only instrument available to managers for altering employees' effort levels was changing the bonus rate. Ten firms with UG managers were run through this no-communication treatment at Case and a n additional eighteen firms with UG managers ran through the no-communication treatment at Universitat Autònoma de Barcelona. While the no-communication treatment is not part of the current experimental design per se, this data provides a useful point of comparison in explaining the results.

4. Results

Before getting into the data, it is useful to reiterate my terminology. A "firm" refers to a fixed grouping of four employees and a manager. "Firm-level" data consists of a single observation per round per firm, including the bonus rate, the messages sent by the manager, and the *minimum* effort chosen by the four employees of the firm. Throughout this paper, unless otherwise specified, I am using firm-level data. When I refer to a "manager," I am describing the subject's role in the game, not whether they were an undergraduate or an executive MBA student. The abbreviations EMBA and UG will be used to distinguish executive MBA and undergraduate subjects, respectively.

For the experiment to be interesting, it is necessary that a history of coordination failure be established in the first ten rounds. The data clearly meets this pre-condition. The average minimum effort falls from 7.18 in round 1 to a paltry 3.08 in round 10. The minimum effort is zero in round 10 for 31 out of a total of 39 firms. When subject managers took over in round 11, the need for a turnaround is almost always present.

Initial conditions differ little between the two treatments. In 16 of 20 firms with UG managers and 15 of 19 firms with EMBA managers, the minimum effort in round 10 was zero. The average minimum efforts are also almost identical: 3.50 for firms that will be managed by an UG and 2.63 for those to be managed by an EMBA.

4.A. Performance of EMBA vs. UG Managers

Figures 1 – 3 compare the outcomes for UG managers and EMBA managers in rounds 11 – 30, the rounds where the subject manager controls the bonus rate and is able to send messages. Round 10 data is included as well to allow for comparison of initial conditions. Figure 1 displays average minimum efforts, Figure 2 shows average bonus rates, and Figure 3 shows average profits. In each figure the top panel shows the average across all firms in each treatment while the bottom panel only uses data from firms that had a minimum effort of zero in round 10. In my opinion, the lower panels give the truest comparison of how UG and EMBA managers performed. By eliminating firms that represent outliers from the sample, the bottom panels only consider firms with identical initial conditions allowing for cleaner comparisons. Moreover, firms with a minimum effort of zero in round 10 represent the case of greatest interest. Unless noted otherwise, the discussion of Figures 1 – 3 refers to data taken from these firms that have an initial state of coordination failure.

The primary result of the paper can be seen in the lower panel of Figure 1. EMBA managers who take over firms with a history of coordination failure overcome coordination failure faster than UG managers in the same situation. In round 11, the average minimum effort in firms with EMBA managers is already 11.7 as compared with 4.4 for firms with UG
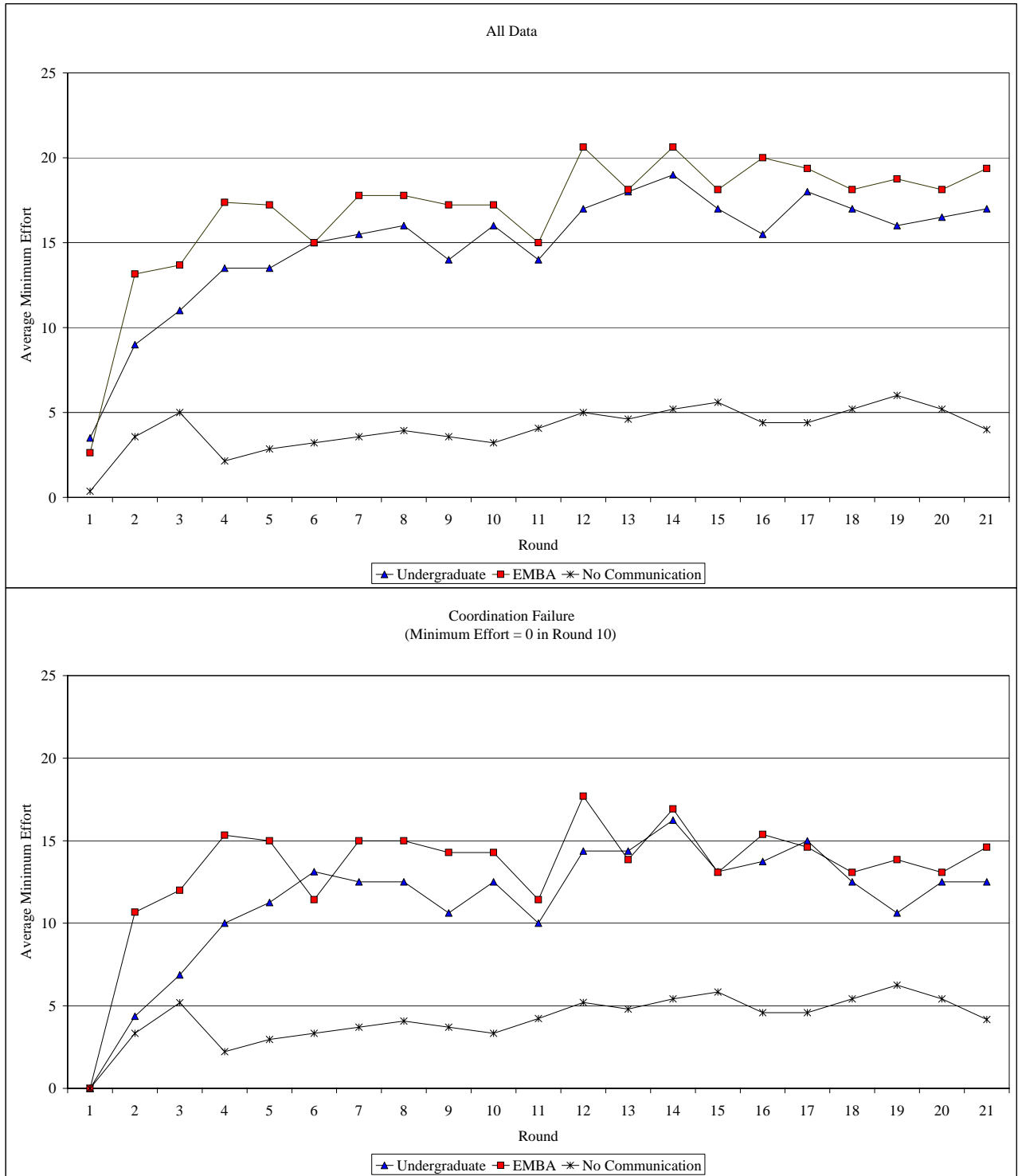
Figure 1

Comparison of Average Minimum Efforts

All Data

Coordination Failure
(Minimum Effort = 0 in Round 10)

Figure 2

Comparison of Average Bonus Rates

All Data

Coordination Failure
(Minimum Effort = 0 in Round 10)
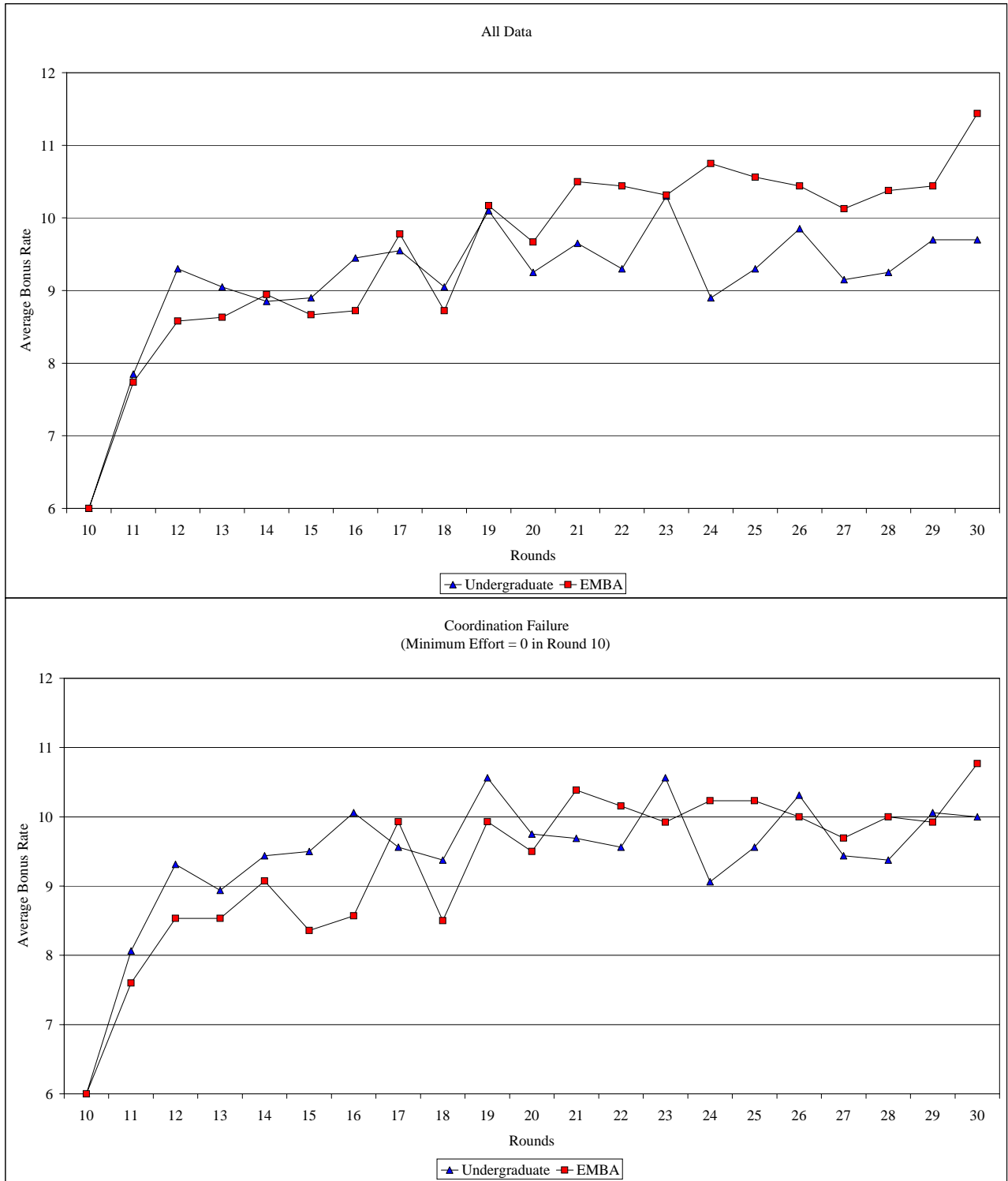
Figure 3

Comparison of Average Manager Profits

All Data

Coordination Failure
(Minimum Effort = 0 in Round 10)
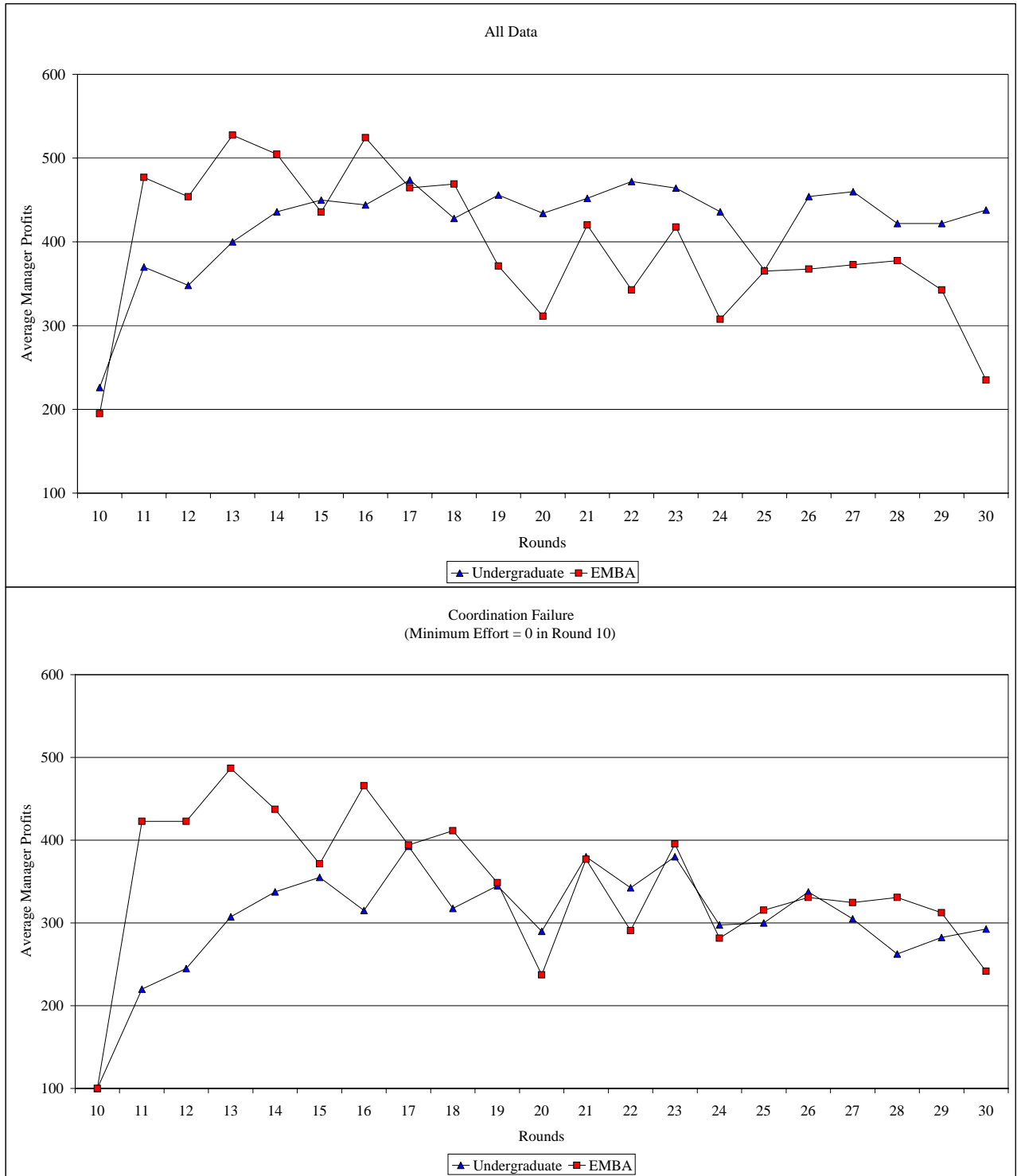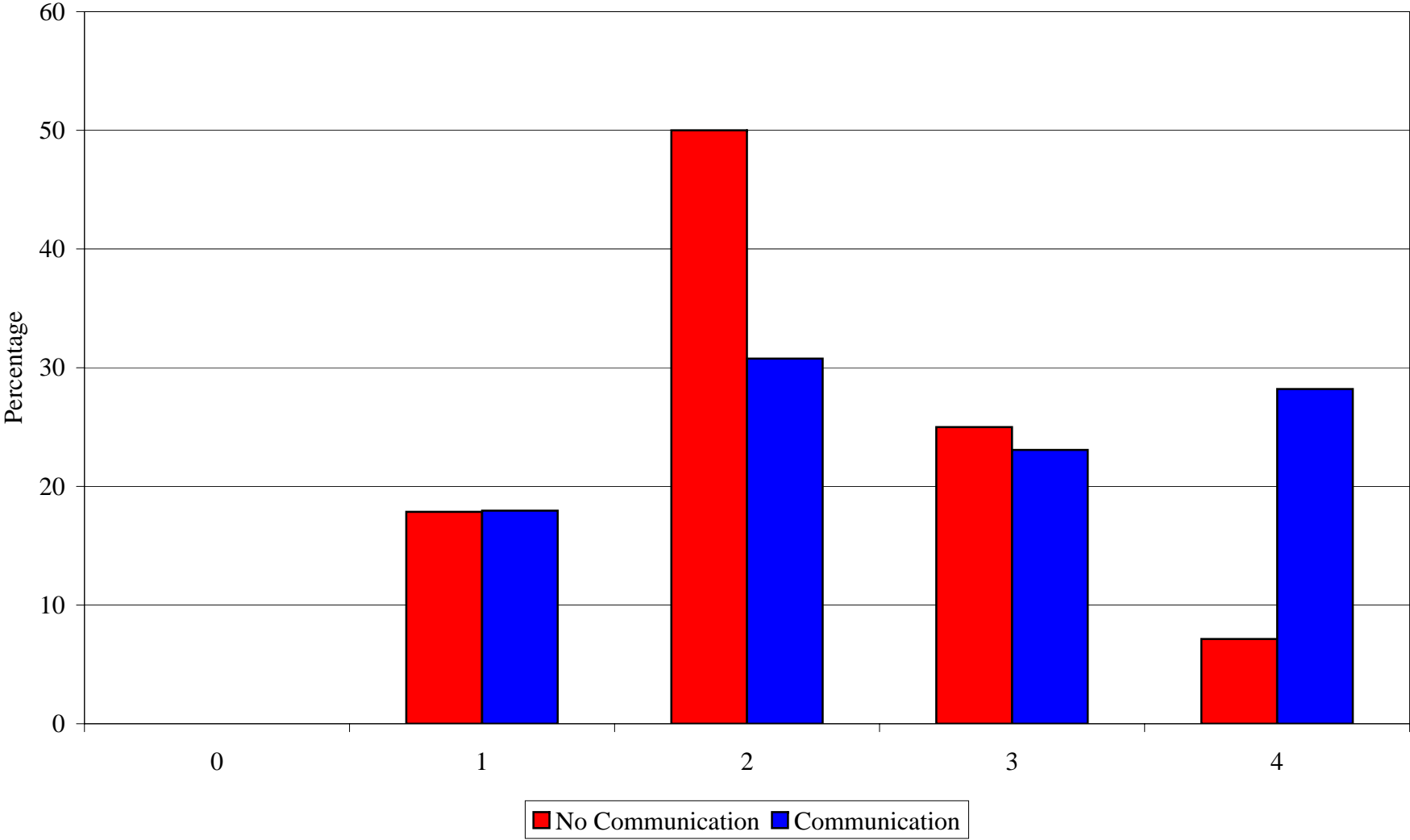
Figure 4
Distribution of Number of Employees Increasing Effort, Round 11

managers. This difference in minimum effort narrows over time as the UG managers catch up with the EMBA managers.[26] Over the final ten rounds, average minimum efforts are largely indistinguishable between the two groups.[27] This convergence for rounds 21 – 30 does not imply that UG managers have learned to play the turnaround game as well as EMBA managers. Faced with a new firm that has a history of coordination failure, the EMBA managers might continue to achieve faster turnarounds than the UG managers.

While neither group is completely successful at overcoming coordination failure, both UG and EMBA managers do well compared with data from the no-communication treatment of Brandts and Cooper (2005b) as shown in Figure 1.[28] To see why it is difficult for managers to overcome coordination failure even with communication, consider the incentives faced by employees. The highest bonus rate chosen for round 11 by any of the 39 managers in our sample is only 10. Given this bonus rate, for an employee to be willing choose an effort level above the previous round's minimum, he must believe that there is a 50% chance that all of the other three employees will choose an effort level above the previous minimum. Treating the employees as independent, this translates into requiring a 79.4% chance of each employee choosing an effort level above the previous rounds minimum. Empirically, in round 11 79.5% of employees choose an effort level above the previous round's minimum.[29] In other words, there is virtually no incentive (and substantial risk) for an employee to increase their effort level even with the highest bonus rate selected. The surprising thing isn't that managers do so poorly, but rather that they don't do even worse!

Figure 4 provides a clue as to how managers with communication are able to achieve a modicum of success. It shows the distribution of how many employees in a firm increase their effort between rounds 10 and 11, the critical moment in the experiment when the subject manager has just taken over the firm. With communication, managers are slightly more able to induce an increase in effort between rounds 10 and 11: 65% of the employees increase their effort with communication as opposed to only 55% without. More striking is how differently shaped the two distributions are. The distribution of how many employees increase their effort between rounds 10 and 11 is sharply peaked with no communication, but quite flat with communication. It appears that communication allows managers to induce correlation between employees' decisions to increase effort. It follows that the likelihood that the other three employees in a particular group will increase their effort levels may be far higher than the unconditional probability calculated above, making it easier to overcome coordination failure. This can cut both ways – higher correlation among employees' choices also makes it more likely

---

[26] Over all twenty rounds with subject managers, the average minimum effort is slightly higher for EMBA managers than UG managers (14.0 vs. 11.9) among firms that start with a minimum effort of 0 in round 10.

[27] The probability of a firm changing its minimum effort level is only slightly lower in rounds 21 – 30 than in rounds 11 – 20 (32% vs. 37%). However, in rounds 21 – 30 increases are only slightly more frequent than decreases (17% vs. 15%) while in rounds 11 – 20 increases are almost twice as frequent than decreases (24% vs. 13%). Thus, the changes in rounds 11 – 20 tend to represent consistent movement to higher effort levels while the changes in rounds 21 – 30 are more random.

[28] To use a larger sample, the data displayed here and in Figure 4 is drawn from both the U.S. and Spanish sessions of the no-communication treatment. Minimum efforts were slightly higher in the U.S. data (5.1 average minimum effort for rounds 11 – 30) than in the Spanish data (3.5 average minimum effort for rounds 11 – 30).

[29] This proportion is only slightly higher (82.5%) for the four groups where the bonus rate equals 10.

that few employees increase their effort – but given the difficulty of overcoming coordination failure the overall effect is positive.[30]

Returning to our comparison of performance by UG and EMBA managers, one possible explanation of how EMBA managers achieve success more rapidly than UG managers is that EMBAs pay their employees more. If true, this would suggest that the early achievement of coordination by EMBA managers is a pyrrhic victory as the additional revenue is dissipated on higher wages. The lower panel of Figure 2 dismisses this possibility. Bonus rates set by UGs are actually slightly higher in the early rounds and are almost identical over the course of the experiment. The average bonus rate for rounds 11 – 30 is 9.61 for UG managers and 9.45 for EMBA managers. Such small differences seem unlikely to drive any treatment effects.[31]

The data displayed in the lower panel of Figure 3 is consistent with our observations on average minimum efforts and bonus rates. Initially average profits are almost twice as high for EMBA managers as for UG managers; in round 11, the average profit for EMBAs is 423 ECUs as compared with 220 for UG managers. This difference is even more impressive when you recall that managers are guaranteed a fixed payoff of 100 ECUs. The gap narrows over time, with average profits being largely identical over rounds 21 – 30.[32]

To summarize, EMBA managers hold a temporary advantage over UG managers due to their ability to more rapidly overcome coordination failure. Figure 5 focuses directly on this feature of the data. The data in this figure comes from firms that have a minimum effort of zero in round 10. The variable being graphed is the cumulative distribution of first coordination times (e.g. the first round in which the firm's minimum effort is greater than zero). The difference between the two populations is transparent. Not only do a higher proportion of EMBA managed firms escape coordination failure in round 11 (10 of 15 vs. 6 of 16), but by round 14 all EMBA managed firms have achieved first coordination (e.g. had at least one round with minimum effort greater than zero) while 5 of 16 UG managed firms have not. Even at the end of the experiment there are still a couple of UG managed firms that haven't achieved coordination a single time.

---

[30] Greater correlation between employees' choices with communication has an interesting implication – in a setting where virtually all groups were successful without communication, the effect of communication might be negative rather than positive as increased correlation could only hurt rather than helping.

[31] It is worth noting that over the entire dataset, including firms with minimum effort greater than zero in round 10, the EMBAs set clearly higher bonus rates in rounds 21 – 30. This is driven entirely by EMBA managers whose firms had a minimum effort greater than zero in round 10. Such firms form a small minority of the sample, so I am reluctant to read too much into this difference. Looking at the eight firms involved, no obvious cause emerges. One outlier among the EMBAs chose the highest possible bonus rate, 15, for the last seven rounds. This individual was playing extremely fast in these late rounds and stopped sending messages, suggesting that he may have become bored with the game and was trying to finish as quickly as possible. Several other EMBAs choose a bonus rate of 15 for only the final round and sent a thank you note at the same time. Given that these EMBAs had made high profits, this apparently represents a case of pure positive reciprocity.

[32] Over all twenty rounds with subject managers, the average managerial profits per round (362 vs. 315) and total surplus per round (1323 vs 1265) are higher with EMBA managers than with UG managers for firms that start with a minimum effort of 0 in round 10.

Figure 5

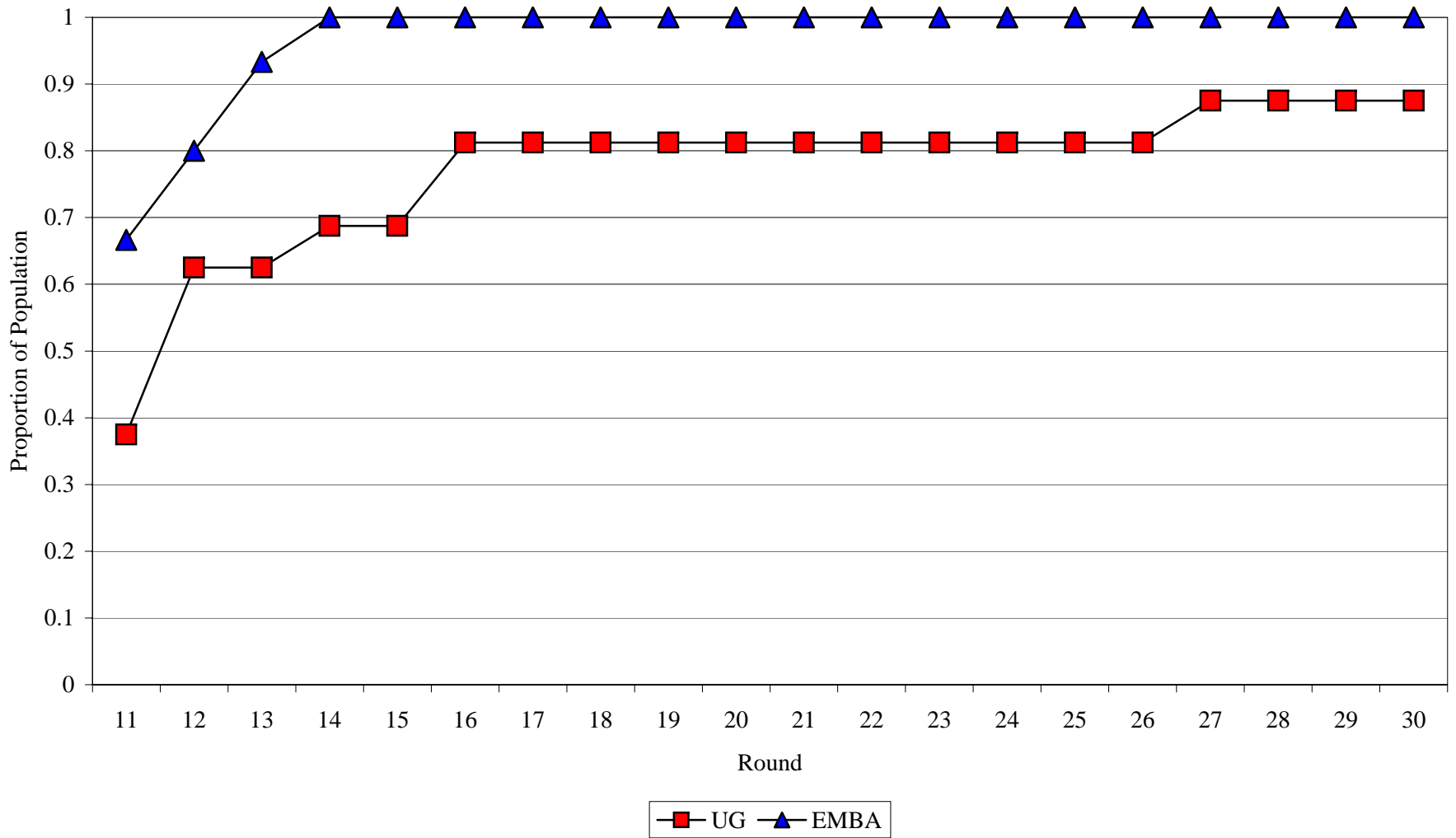Cumulative Distribution of First Coordination Times

Table 3

Ordered Probit Regressions
Are EMBA Managers Significantly Faster Than UG Managers?

Variable	Model 1	Model 2	Model 3	Model 4
Data Set	All Data Rounds 11 - 30		Lagged Minimum Effort = 0 Rounds 11 - 30	
Lagged Minimum Effort	.077*** (.016)	.077*** (.016)		
Twice Lagged Minimum Effort	.039*** (.010)	.039*** (.010)	.031* (.017)	.028* (.017)
Current Bonus	.035 (.034)	.036 (.035)	-.050 (.050)	-.044 (.049)
Change in Bonus	.152*** (.039)	.150*** (.039)	.216*** (.052)	.215*** (.051)
Minimum Effort Round 10	.014 (.009)	.014 (.010)	.046 (.039)	.073*** (.024)
Rounds 21 – 30	-.165* (.088)	-.162* (.089)	-.151 (.176)	-.142 (.187)
Executive MBA	.208* (.118)	.620** (.289)	.622*** (.193)	1.856*** (.527)
Gender (0 = Male, 1 = Female)		-.043 (.278)		.230 (.385)
Age		-.022 (.014)		-.052** (.023)
Log-likelihood	-672.34	-671.14	-184.16	-181.90

Note: Models 1 and 2 use 743 observations from 39 firms. Models 3 and 4 use 254 observations from 33 firms. Standard errors are corrected for clustering at the firm level.

***   Significant at 1% level
**   Significant at 5% level
*   Significant at 10% level

Up to this point I have relied on a purely visual examination of the data, ignoring any issues of statistical significance. To put the preceding discussion on a firmer statistical footing, Table 3 reports results from the estimation of ordered probit models. My use of an ordered probit specification is driven largely by the structure of the data. The dependent variable, a firm's minimum effort for the round being observed, is a categorical variable.   This makes use of an ordered probit model more apt than OLS (or one of its variants) which assumes the dependent variable is continuous. Given that the central question in this analysis is whether EMBA managers are able to overcome coordination failure more quickly than UG managers, Cox's proportional hazard model (Cox, 1972) is a plausible candidate as an alternative statistical model. Indeed, I have analyzed proportion hazards models equivalent to the ordered probits

reported here and find qualitatively similar results.[33] The use of a proportional hazards model has the advantages of focusing on the data of greatest interest as well as relying on weaker assumptions about functional forms. However, only a small fraction of the data can be used in fitting the proportional hazards model: all data from firms that had a positive minimum effort in round 10 as well as all data after a firm first achieved a positive minimum effort is discarded. I feel that using more data tips the balance in favor of the ordered probit analysis.

For all of the regressions reported in Table 3 the dependent variable is the firm's minimum effort. Models 1 and 2 in Table 3 include all observations from all 39 firms in the data set for rounds 11 – 30. Judging by Figures 1 and 5, the effectiveness of EMBA managers is most evident when firms are facing coordination failure. Models 3 and 4 therefore replicate Models 1 and 2 with the dataset limited to observations in which the firm's lagged minimum effort is zero. Since firms appear multiple times in the data set, the standard errors are corrected for clustering at the firm level (see Liang and Zeger, 1986). In specifying these regressions, the central question is whether EMBA managers are better able to increase the minimum effort levels of their firms. Because of this focus on change, the regressions include lagged dependent variables – the lagged minimum effort and the twice-lagged minimum effort. The interpretation of positive (negative) parameter estimates, given that lagged dependent variables are present in the specification, is that the variable in question makes the minimum effort more (less) likely to increase from the previous period's value. In other words, the regression model studies changes rather than levels. The inclusion of lagged dependent variables also eases the introduction of communication measures later in this section. Unlike usual, the use of lagged dependent variables does not require us to drop any data since the first data being included is from round 11, not round 1. All regressions reported in Table 3 include controls for the current bonus rate as well as the change in the bonus (current minus lagged bonus rate). Given that bonus rates are similar for UG and EMBA managers, these variables are included largely to absorb some of the noise in the data. The minimum effort for round 10 is included as an explanatory variable, but generally fails to achieve statistical significance (reflecting the presence of lagged dependent variables). Finally, all of the regressions include a dummy for rounds 21 – 30. This variable is always negative and often statistically significant. Given that the regressions capture *changes* in minimum effort levels, these negative estimates reflect the lack of movement of firms' average minimum effort levels over the final ten rounds.

Turning to the explanatory variables of primary interest, Model 1 includes a dummy for whether the firm manager was an EMBA. The parameter estimate for this variable is positive and statistically significant at the 10% level. Model 2 adds two additional covariates: a dummy for the manager's gender (0 = male, 1 = female) and the manager's age. No significant gender effect is detected in the data. The parameter estimate for the manager's age is negative and narrowly misses statistical significance at the 10% level. The parameter estimate for the EMBA dummy is greatly increased and edges up to significance at the 5% level in Model 2. The increased magnitude is due to an interaction with the "Age" variable – within a type of manager, older individuals do worse at overcoming coordination. Since EMBA managers are older than UG managers, the size of the associated parameter estimate must increase to compensate for the

---

[33] The proportional hazards model was fit only for firms that had a minimum effort of zero in round 10. The duration used for fitting was first time with minimum effort greater than zero. Qualitatively, the results were similar to those reported in Table 3 and 6. In particular, the dummy for EMBA managers was statistically significant even after controlling for other characteristics of the managers. Levels of statistical significance for the EMBA dummy tend to be higher in the proportional hazards model than in the ordered probits reported in Tables 3 and 6.

inclusion of "Age." The negative effect of age provides evidence, albeit weak, that the superior performance of EMBA managers cannot be attributed solely to their greater age but instead lies in something about the particular nature of their experiences. The results of Models 3 and 4 are similar to those of Models 1 and 2 but stronger. In Model 3, the EMBA dummy parameter estimate is about three times larger than in Model 1 and is statistically significant at the 1% level.[34] The gender effect in Model 4 is once again not statistically significant while the age effect now achieves statistical significance at the 5% level. Matching the impressions given by Figures 1 and 4, the primary advantage of EMBA managers lies in cases where coordination failure must be overcome.

As noted previously, two alternative explanations of why EMBA managers have expertise in overcoming coordination failure are selection and learning effects. The data cannot definitively separate these causes, but it does provide a hint. Five of the UG managers had majors in the business school (accounting, economics, or management).[35] While there is no guarantee that these individuals will go on to careers in management or that the numerous engineers in the sample won't, it seems reasonable to assume that choosing a major in a school of management is a leading indicator of a career in management. The performance of undergraduates who chose a major in the business school is virtually identical to the performance of those who don't. Their average minimum effort and payoff per round are 15.5 and 436 as compared with 15.4 and 430 for non-management majors. When a dummy for business school majors is added to Model 1 in Table 3, the parameter estimate is negative but tiny and statistically insignificant. This evidence isn't definitive, as the sample is small and choice of a management major is only a proxy for eventually having a career in management, but these results suggest that selection into a career in management isn't driving the main result.

Conclusion 1: **EMBA managers overcome a history coordination failure faster than UG managers. This difference is statistically significant and cannot be attributed to differences in the ages or genders of these two populations.**

4.B. Communication and Differences in Managerial Performance

Given that bonus rates are virtually identical for the two types of managers, something else must explain the superior performance of EMBA managers. The only other way in which managers interact with their employees is via the messages they send. This implies that differences in how EMBA managers and UG managers communicate with their employees probably explain the observed differences in performance.

Some way of quantifying the content of messages is necessary to get at these issues. Brandts and Cooper (2005b) developed and implemented a systematic scheme for coding the content of messages using methods largely identical to those employed by Cooper and Kagel (2005). In this section I report the results of applying this method to the current dataset.

In coding the content of messages, the goal is to systematically quantify any communication that might be relevant to play of the game without pre-judging which sorts of messages are important and which sorts are not. The process began by randomly selecting ten firms to serve as a test sample. Jordi Brandts and I as well as two research assistants

---

[34] Comparing marginal effects rather than parameter estimates, I find that the marginal effect is also roughly three times larger Model 3 than in Model 1.

[35] This includes individuals who choose one of these majors after the time the experiment was run.

independently developed coding schemes for the test sample. In a series of meetings we reconciled these individual efforts into a single coding scheme. The two research assistants then independently coded all messages. No effort was made to force agreement among coders – the goal was to have two independent readings of each message so that any coding errors were uncorrelated. At no point in the process of developing or implementing the coding scheme was either RA informed about any hypotheses I had about the messages. The RAs were repeatedly and explicitly told that their job was to capture what had been said rather than why it was said or what effect it had. Coding was binary – a message was coded as a 1 if it was deemed to contain the relevant category of content and zero otherwise. There was no requirement on the number of codings for a message – a coder could check as many or few categories as he or she deemed appropriate. A number of the categories have sub-categories. For example, Category 1 for managers is "Ask for Effort." Under this are three sub-categories, "Polite," "Rude," and "Specific effort level." A coder was free to check as many or few sub-categories as they desired when a category was checked off.[36]

For this paper, all analysis of the codings uses averages across coders. In averaging across coders, I am implicitly assuming that errors are independent across coders so that averaging reduces the total error. Considering the most common categories, defined as those that were coded for at least 2% of all possible observations, the mean cross-coder correlation was .605 with a maximum of .789 and a minimum of .383. These statistics are similar to the results reported in Cooper and Kagel (2005) and Brandts and Cooper (2005b). Note that all discussion of the codings only uses data in which the messages were recorded. The four UG led firms in which the messages were accidentally not saved are not included. Table 4 gives a brief description of the categories and summarizes the coding results for the managers. For a lengthier description of the common message categories along with exemplar quotes, see Brandts and Cooper (2005b).

---

[36] A sub-category could not be coded without coding the corresponding category. Although the coders weren't explicitly instructed to not check mutually exclusive sub-categories like "Polite" and "Rude," they never did so.

Table 4

Summary of Manager Codings, Rounds 11 - 30
Frequency of Codings

Category	Description	Undergraduate	EMBA	All Managers
1	Ask for Effort (code appropriate sub-categories as well)	0.269	0.251	0.259
1A	Polite	0.073	0.019	0.045
1B	Rude	0.016	0.000	0.008
1C	Specific effort level	0.120	0.163	0.143
2	Negative response (code appropriate sub-categories as well)	0.073	0.067	0.070
2A	Encouraging	0.014	0.016	0.015
2B	Hostile	0.016	0.003	0.009
2C	"Singling" our an employee	0.008	0.009	0.008
3	Positive response (praise, thanks, appreciation, etc)	0.172	0.210	0.192
4	Discuss monetary benefits of high effort (code appropriate sub-categories as well)	0.084	0.095	0.090
4A	Benefits for manager	0.011	0.004	0.008
4B	Benefits for employees	0.027	0.052	0.040
4C	Mutual Benefits	0.050	0.042	0.046
5	Implicit Contracts (code appropriate sub-categories as well)	0.061	0.071	0.066
5A	More effort today -> higher bonus tomorrow	0.042	0.054	0.048
5B	Lower effort today ? lower bonus tomorrow	0.006	0.003	0.005
5C	High bonus today, request higher effort in response	0.016	0.017	0.017
6	Laying out a plan (code appropriate sub-categories as well)	0.042	0.023	0.032
6A	Alternating plan	0.016	0.007	0.011
6B	Ratcheting up effort	0.002	0.007	0.005
7	Surprising employees (code appropriate sub-categories as well)	0.008	0.010	0.009
7A	Choosing higher bonus than specified by plan	0.005	0.007	0.006
7B	Choosing lower bonus than specified by plan	0.000	0.000	0.000
8	Encouragement (should not specifically refer to effort)	0.067	0.122	0.096
9	Use of humor	0.011	0.006	0.008
10	Emphasizing the bonus (includes explicitly stating what the bonus will be)	0.089	0.138	0.115
11	Comments about time (code appropriate sub-categories as well)	0.008	0.000	0.004
11A	Need to hurry to get finished	0.000	0.000	0.000
11B	Will be able to leave sooner if cooperate	0.003	0.000	0.002
12	Explicit reference to fairness	0.005	0.006	0.005
13	Explicit references to trust	0.006	0.042	0.025
14	Explicit references to reciprocity	0.005	0.012	0.008
15	Attempts by managers to appear sympathetic	0.011	0.006	0.008
16	Expressing confusion about the rules	0.000	0.000	0.000
17	Clarifying the rules	0.000	0.000	0.000
18	Misunderstanding rules	0.000	0.013	0.007

Table 5

Common Codings, Comparison of Frequencies for Rounds 11 - 15

Category	Description	Undergraduate	EMBA	Ratio
1	Ask for Effort (code appropriate sub-categories as well)	0.313	0.323	1.03
1A	Polite	0.063	0.038	0.60
1C	Specific effort level	0.069	0.172	2.50
2	Negative response (code appropriate sub-categories as well)	0.050	0.075	1.51
3	Positive response (praise, thanks, appreciation, etc)	0.169	0.258	1.53
4	Discuss monetary benefits of high effort (code appropriate sub-categories as well	0.131	0.113	0.86
4B	Benefits for employees	0.019	0.022	1.15
4C	Mutual Benefits	0.100	0.102	1.02
5	Implicit Contracts (code appropriate sub-categories as well)	0.094	0.140	1.49
5A	More effort today -> higher bonus tomorrow	0.063	0.108	1.72
5C	High bonus today, request higher effort in response	0.031	0.038	1.20
6	Laying out a plan (code appropriate sub-categories as well)	0.038	0.022	0.57
8	Encouragement (should not specifically refer to effort)	0.075	0.167	2.22
10	Emphasizing the bonus (includes explicitly stating what the bonus will be)	0.100	0.194	1.94
13	Explicit references to trust	0.013	0.086	6.88

To begin understanding the role of communication in the differing performance of UG and EMBA managers, Table 5 compares the frequency of the most common codes for the two types of managers. Once again, common codes are defined as all categories that were coded for at least 2% of all possible observations.[37] Given that most of the change in average minimum effort takes place over the first few rounds after the human managers take control, Table 5 reports frequencies for rounds 11 – 15.[38] The final column reports the ratio of coding frequency for EMBA managers as compared to UG managers. Categories which are at least 50% more likely for one type of manager than the other have been highlighted.

Scanning Table 5, several differences between UG and EMBA managers are apparent. First, EMBA managers tend to be much more communicative. The frequency of all codings over all rounds is 18% higher for EMBA managers. If we restrict our attention to common codings over the first five rounds with communication, this difference climbs to 40%.[39] This greater frequency of communication helps explains some puzzling features of the data; for example, EMBA managers are over 50% more likely in rounds 11 – 15 to be coded for both positive *and* negative responses to the employees (Categories 2 and 3).

Even given the talkative nature of EMBA managers, it is clear that what they are saying differs from the messages of UG managers. For example, EMBA managers are far more likely to request a specific effort level (Category 1A), less likely to lay out a long term plan (Category 6), and more likely to offer encouragement to employees (Category 8). A striking difference is that EMBA managers are over six times as likely to make explicit references to trust (Category 13). The latter example represents a rarity; for the most part, the difference between EMBA and UG manager messages lies not in what categories are common (as defined above) but rather in how frequently these common categories are coded. Explicit references to trust (Category 13) is the only category that is a common coding for EMBA managers in rounds 11 - 15 but not for UG managers. Likewise, use of humor (Category 9) is the only category that is a common coding for UG managers in rounds 11 – 15 but not EMBA managers.[40]

---

[37] I initially used a cutoff of 5%. However, robustness checks showed that a category with substantial explanatory power, Category 6, was excluded by this cutoff. I therefore choose a looser standard to avoid excluding any clearly important categories.

[38] The differences between messages sent by UG and EMBA managers are greatly diminished in rounds 16 – 30.

[39] These numbers do not include any sub-categories to avoid double counting.

[40] This difference is an anomaly, reflecting a single UG manager who was coded for use of humor very frequently.

Table 6

Ordered Probit Regressions
Communication and the Effect of EMBA Managers

Variable	Model 1	Model 2	Model 3
Lagged Minimum Effort	.071*** (.016)	.073*** (.017)	.073*** (.017)
Twice Lagged Minimum Effort	.041*** (.010)	.043*** (.010)	.043*** (.010)
Current Bonus	.105*** (.024)	.114*** (.024)	.112*** (.023)
Change in Bonus	.118*** (.040)	.118*** (.039)	.116*** (.041)
Minimum Effort Round 10	.024*** (.007)	.020*** (.007)	.021*** (.007)
Rounds 21 – 30	-.239*** (.083)	-.230*** (.077)	-.239*** (.078)
Executive MBA	.269** (.125)	.191 (.133)	.191 (.129)
Category 1c		.482*** (.158)	
Category 2		-.024 (.228)	
Category 3		.107 (.137)	
Category 4		.241 (.147)	
Category 5a		.238 (.252)	
Category 6		-.683* (.374)	
Category 8		-.273 (.183)	
Category 10		.473** (.227)	
Category 13		.380 (.516)	
Good Communication (Cat 1c + Cat 10 – Cat 6)			.570*** (.129)
Log-likelihood	-591.56	-573.83	-576.99

Note: Data from rounds 11 – 30 in sessions where messages were recorded (663 observations, 35 firms). Standard errors are corrected for clustering at the firm level.

*** Significant at 1% level
** Significant at 5% level
* Significant at 10% level

Having established that EMBA managers and UG managers have different frequencies for using common types of messages, the next issue is whether these differences explain their relative speeds in overcoming coordination failure. Table 6 reports the results of ordered probit regressions (with standard errors corrected for clustering) addressing this question. The data set is all observations from rounds 11 – 30 from firms where the messages were recorded. As in Table 3, all of the regressions on Table 6 include the lagged minimum effort, twice lagged minimum effort, current bonus, change in bonus, minimum effort for round 10, a dummy for rounds 21 – 30, and a dummy for whether the manager was an EMBA. The lagged dependent variables are especially important here as they make it possible to identify causality when the codings are introduced as independent variables. The issue is that the lagged minimum effort might have a causal relationship both with the current messages and with the current minimum effort. For example, suppose that a negative response (cat. 2) is more likely to be coded in round t if the minimum effort level is low in round t - 1. Imagine that this coding has no causal relationship with employees' behavior in round t but the lagged minimum effort is a good predictor of the current minimum effort. If the lagged minimum effort isn't included in the regression, I might falsely infer that a negative response causes a lower minimum effort.

To select which categories would be included in the regressions, I first limited my attention to common codings (e.g. categories coded for at least 2% of observations). To prevent co-linearity, I didn't include multiple sub-categories under a category. Instead, since the goal is to determine whether differences in communication strategies explain the greater effectiveness of EMBA managers, I chose the category or associated subcategory that had the greatest difference in relative frequency between the two types of managers in rounds 11 – 15. For example, comparing Categories 1, 1A, and 1C, EMBA managers were 3% more like to be coded for Category 1, UG managers were 66% more likely to be coded for Category 1A, and EMBA managers were 150% more likely to be coded for Category 1C. I therefore included Category 1C, the case with the greatest relative difference, as an independent variable in Model 2.

Model 1 in Table 6 replicates the results from Model 1 on Table 3 with the four firms where messages weren't recorded being dropped. The dummy for whether the manager was an EMBA is again positive and statistically significant. Model 2 adds average codings for the current round as independent variables. Categories 1c (requesting a specific effort level) and 10 (emphasizing the bonus rate) significantly increase the likelihood that the minimum effort level improves while Category 6 (laying out a plan) significantly reduces this likelihood.[41] Critically, the dummy for whether the manager was an EMBA is reduced by 29% and is no longer statistically significant.[42] It is striking that the message categories with a statistically significant effect (Categories 1c, 6, and 10) represent three of the five categories with the greatest relative difference between EMBA and UG managers. It stretches credulity that EMBA managers *by coincidence* would be especially different from UG managers in exactly the categories of messages which are most effective in overcoming coordination failure.

The ability of the communication variables to explain the difference between EMBA and UG managers is especially impressive because the codings only provide an imprecise measure of what is actually being communicated. There are qualitative differences between UG and EMBA manager messages that are obvious in reading through the transcripts but cannot be easily quantified and captured through the coding process. For instance, messages from EMBA

---

[41] Category 4 barely misses statistical significance. If Category 4c is substituted for Category 4, this achieves statistical significance. The results of interest are little affected by this substitution.

[42] I am comparing parameter estimates here, but marginal effects give similar results.

managers tend to be longer and more detailed than messages from UG managers. This allows messages with exactly the same coding to be more persuasive. Given this errors in measurement problem, the true impact of communication is probably being underestimated in Model 2 and, by extension, the EMBA coefficient is likely overestimated in Model 2.

It is important to note that the "good" communication strategy identified from the current dataset (ask for a specific effort level, point out the mutual benefits of higher effort, emphasize the bonus rate, and don't suggest a long-term plan) is almost the same as Brandts and Cooper (2005b) identify using undergraduates only.[43] This isn't a replication, since the UG manager data in this paper is a subset of the data used by Brandts and Cooper (2005b), but instead demonstrates that the previous results are not driven by the use of naïve subjects.

To understand why this particular communication strategy works, recall that coordination is largely a problem of beliefs. Communication correlates beliefs, leading to the correlation in actions shown in Figure 4. With a good communication strategy, the manager creates common beliefs that most employees will be choosing high effort levels. This is most obvious when a manager asks employees for a specific effort level (Category 1C).[44] More subtly, pointing out the *mutual* benefits of successful coordination (Category 4C) creates expectations that *all* employees will select high effort levels in order to enjoy higher payoffs, making it safer for any one employee to increase his effort level.[45] Even emphasizing the bonus rate (Category 10) [46] can change beliefs about others choices. If an employee receives a message pointing out that the benefits of effort have just gone up, *he knows that others are also seeing this message*. He should therefore believe that other employees are more likely to raise their effort levels. Many common types of messages, such as giving positive feedback (Category 3), negative feedback (Category 2),[47] explicitly referring to trust (Category 13), or suggesting a quid pro quo (Category 5A) have little impact on employee performance. Presumably this is because these categories make poor coordination devices. For instance, Category 5A is effective only if employees believe the manager is trustworthy (e.g. will raise the bonus rate as a reward for higher minimum effort) and that the other employees trust the manager. This is a less direct means of coordinating beliefs than simply telling everyone to choose the highest available effort level. Oddly, laying out a plan (Category 6) for how bonus rates and effort levels will change over time tends to retard the development of coordination. This feature of the data becomes more

---

[43] Brandts and Cooper (2005b) did not consider sub-categories (so for example, Category 1 is identified as part of a good communication strategy rather than sub-category 1C). Category 4 is statistically significant is Brandts and Cooper, but barely misses here with a smaller dataset. As noted in fn. 41, Category 4C is statistically significant both here and in Brandts and Cooper (2005b).

[44] The following sample message, sent by an EMBA manager, was coded under category 1C: "Now we're headed in the right direction again... another round of [effort level] 40's?" After receiving this message, all employees have a common expectation that effort level 40 is likely to be chosen by others.

[45] The following message, by an EMBA manager, illustrates the point. This was coded under categories 1C and 4C: "Hey team - let's work together. We all win if everybody works 40 hours at task A. No downside!"

[46] Most messages coded under category 10 imply that the employees are being paid well and/or underline that the bonus rate has just increased. For example, the following message was sent by an EMBA manager in a round when the bonus rate increased from 7 to 10: "Let's see if this bonus amount can help a little on task A."

[47] Typically, category 3 messages praise employees for their effort (e.g. "Good job you guys, you are truly amazing!!") while category 2 messages criticize low efforts (e.g. "Don't do that again! Boooo!").

understandable after looking at the specific plans suggested by managers – these generally involve an increase in the share of surplus being received by the manager.[48]

Model 3 replaces the average codings with a simpler measure of whether a manager is using a "good" communication strategy: the average codings for Categories 1c and 10 are added together and then the average coding for Category 6 is subtracted. In other words, this good communication measure is higher when managers use the two significantly helpful categories of messages and lower when the lone significantly harmful category is used. The results of Model 3 closely mirror those in Model 2, particularly vis-à-vis the EMBA coefficient estimate. The advantage of EMBA managers over UG managers in overcoming coordination failures is largely due their greater frequency of using a good communication strategy. If Model 3 is run separately for the UG and EMBA managers, the good communication measure is statistically significant at the 5% level for both populations. This supports the external validity of studies run using undergraduates, since the successful strategy identified from UG manager data also applies for the EMBA managers.

Table 7

Ordered Probit Regressions
Differences Between Communication by Undergraduate and EMBA Managers

Dependent Variable	Z-score for EMBA Dummy	Statistical Significance
Category 1c	0.74	---
Category 2	-0.07	---
Category 3	0.19	---
Category 4	0.80	---
Category 5a	0.80	---
Category 6	-0.72	---
Category 8	1.56	---
Category 10	1.18	---
Category 13	2.78	***
Good Communication	1.94	*

Note: Data from rounds 11 – 30 in sessions where messages were recorded (663 observations, 35 firms). Standard errors and the resulting z-scores are corrected for clustering at the firm level.

***   Significant at 1% level
**   Significant at 5% level
*   Significant at 10% level
---   Not significant at the 10% level

---

[48] For example, consider the following message from an EMBA manager who had recently gotten his employees to coordinate at effort level 10 by raising the bonus rate to 12, "I'm going to $11 this round to average my return and yours. Back to $12 next round. Your best strategy is to stay at 10, minimize your risk, I presume." This particular manager tried this approach several times and was consistently rewarded with lower effort levels.

A final point to be established is whether the difference in communication strategies between UG and EMBA managers is statistically significant. To answer this question, I ran a series of ordered probit regressions similar to those reported in Tables 3 and 6. There is one regression for each category of message included on Table 6 as well as a regression for the good communication measure. Specifically, the dependent variable in a regression is the average coding of the category in question for the current round. The data set again includes all observations from rounds 11 – 30 from firms where the messages were successfully recorded. All of the regressions reported in Table 7 include the lagged minimum effort, twice lagged minimum effort, lagged bonus, lagged change in bonus, minimum effort for round 10, a dummy for rounds 21 – 30, and a dummy for whether the manager was an EMBA. Note that the current bonus is not included since this is chosen by the manager simultaneously with the choice of a message.

Table 7 summarizes the findings of these regressions. The second column reports the z-score of the estimated coefficient for the EMBA dummy and the third column indicates the level of significance for this estimate. The z-scores have been corrected for clustering. Individually, only for Category 13 is there a significant difference in the likelihood that EMBA managers send a particular type of message. For the most part, the EMBA dummy does not even approach statistical significance. However, when codings are combined to create the measure of "good" communication, the EMBA dummy is positive and barely misses significance at the 5% level. The difference between UG and EMBA managers lays not so much in greater (or lesser) use of one particular category, but rather in greater usage of a combination of messages that jointly yield faster coordination.

Conclusion 2: **The significantly greater usage by EMBA managers of a "good" communication strategy explains much of their advantage over UG managers in overcoming coordination failure.**

5. Conclusion

This paper illustrates both the benefits and pitfalls of using experienced managers as experimental subjects. A common criticism of experiments using standard undergraduate subjects is a lack of external validity. Using EMBAs in the manager role allows me to address this criticism. The experimental results reported above verify that the conclusions of studies with the turnaround game using undergraduate subjects, particularly Brandts and Cooper (2005b), are likely to generalize to relevant field settings. EMBA managers are faster than UG managers in turning their firms around because they use the same good communication strategy identified in Brandts and Cooper (2005b) more frequently than UG managers. EMBA managers are not succeeding via a distinctly different communication strategy or incentive scheme than UG managers.

My confidence in these conclusions is strengthened by drawing on other papers that use artefactual field experiments to verify external validity. In some cases, the results with undergraduate subjects hold up well. In a famous early example, Lichtenstein and Slovic (1973) were able to generate preference reversals using gamblers in casinos.[49] In other cases, such as Cooper *et al's* (1999) results described previously, the behavior of expert subjects does not

---

[49] Smith, Suchanek, and Williams (1988) and Dyer and Kagel (1996) are good examples of using managers as subjects to verify that anomalous behavior observed with undergraduate subjects (bubbles and the winner's curse, respectively) extends to an expert population.

mirror that of standard undergraduate subjects.[50] Comparing these studies, some general patterns emerge for when the behavior of expert subjects will differ from the behavior of undergraduates. First, expert subjects' field experiences must be relevant for the tasks they are asked to perform in the laboratory. For example, the poor performance of construction executives in common value auctions reported by Dyer and Kagel (1996) can be attributed in part to a poor match between the problem faced by executives in the field and the task facing them in the lab. In the field setting there exist escape mechanisms that both help protect contractors from losses due to the winner's curse and their customers from non-performance due to contractors walking away from unprofitable projects. As these mechanisms are missing in the laboratory, the expertise of construction executives is rendered less useful.[51] Second, the context of the experiment must allow expert subjects to recognize that their expertise is relevant. Thus, the performance of plant managers in Cooper et al (1999) is significantly affected by whether the experiments are framed in an abstract or meaningful context. Strategic play is faster to develop with meaningful context, presumably because this eases the task of drawing on relevant field experience.

Both of these issues affect the experiments described above. While firm managers are allowed wide latitude in what messages they can send, it isn't difficult to imagine productive strategies which aren't allowed.[52] The instructions attempt to evoke a corporate setting, but the need to avoid emotionally loaded terms leads to obscure terminology. For example, the difference between "making an effort" and "slacking off" is natural, but the distinction between "Task A" and "Task B" may take some experience to decipher. This lack of a perfect match between the context and action sets in the field and laboratory works against finding differences between UG and EMBA managers, making Conclusion 1 more compelling.

The nature of the differences between UG and EMBA managers also fit well with the differences previous studies have found between student and manager subjects – the differences tend to be in intensity of behavior rather than type of behavior. For example, both Hannan et al (2002) and Fehr and List (2004) find that manager's behavior in trust games departs even further from the predictions of subgame perfection than student's behavior.[53] It appears that important phenomena observed in the laboratory are likely to exist in the field as well, but may differ in magnitude.

The success of EMBA managers underlines the importance of communication as claimed by Brandts and Cooper (2005b). The inclusion of EMBA subjects also makes it possible to infer characteristics of the environment in which EMBAs normally operate. Experienced, successful managers are not a random sample from the population. These individuals achieved their success either because they are naturally well suited to their jobs (selection) or because they have learned to do them well. If seasoned managers are particularly good at a specialized task and their success cannot be attributed to general managerial expertise, it is unlikely that their skill is a coincidence. Instead, it must play some role in the success that allows them to become experienced managers. EMBA managers are faster at achieving coordination in the turnaround

---

[50] See also Burns (1985), Potters and van Winden (2000), Hannan, Kagel, and Moser (2002), Fehr and List (2004), Montmarquette, Rulliere, Villeval, and Zeiliger (2004), and Alevy, Haigh, and List (2005) all of which find differences between undergraduate subjects and subjects with relevant business expertise.

[51] See McDaniel and Rutström (2001) for another good example of this problem.

[52] For example, managers might do better if they could commit to multi-round contracts.

[53] In other cases manager subjects differ less from the theory than undergraduates. For example, in both Cooper *et al* (1999) and Potters and van Winden (2000) the game theoretic predictions perform better with manager subjects than with students. The point is that in neither case do the manager subjects perfectly conform to the theoretical predictions, and the patterns of how subjects depart from the theory are similar across the two populations.

game. Their path to success is a highly specific communication strategy that is not obviously applicable to a wide variety of other problems faced by managers. Together, these two observations imply that coordination problems are prevalent in many workplaces. It is difficult to directly observe the interactions managers in a wide variety of industries have with their employees, but experiments using experienced managers as subjects provide an indirect method of studying these interactions.

The difficulties of using experienced managers as subjects are largely methodological. The most obvious revolve around getting managers to participate. In designing these experiments, I went to great lengths to make participation as painless as possible for the EMBA subjects. Even doing so it was difficult to recruit a sufficiently large subject pool − I did not have enough EMBA subjects to run a second treatment. In evaluating experiments using non-standard subject pools, experimental economics as a field must make some decisions on how much departure from ideal methods is permissible to gain access to an interesting population. For example, I generally don't allow subjects to participate from remote locations in experiments which use solely undergraduate subjects, but I felt this was a reasonable modification to have access to EMBA subjects who certainly would not have participated otherwise.

The more subtle difficulty of using experienced managers as subjects requires recognizing that this population differs from the standard undergraduate population on more dimensions than just having experience as managers. Some of these differences, such as the greater age of the EMBA subjects, can be controlled for econometrically. Other differences raise thornier problems. For example, the EMBA managers had more difficulty with the instructions than the UG managers. While this probably didn't affect our main conclusions, it may have reduced the size of the measured difference as well as making the EMBA data noisier. In comparing data from a non-standard population with data from the standard undergraduate population, it is important to remember that the *obvious* way in which the two populations differ is not necessarily the cause of differing behavior.

An important question not answered by this study is *why* the EMBA managers behave differently than UG managers; is this due to selection, learning, or leadership training received as part of the EMBA program? Separating out these possible explanations is a promising topic for further research. Eliminating leadership training as a possible cause of the observed differences is relatively easy, as the experiment can be replicated with incoming EMBA students prior to any leadership training. If leadership training is causing the superior performance of EMBA managers, it would be evidence that business-school education has a positive impact on managerial performance.

Separating selection from learning is more difficult. This is an important issue for many artefactual field experiments using expert populations. Experimenters have been sensitive to the possibility that individuals in the standard undergraduate subject population do not represent a random selection from the undergraduate population (e.g. Rutström, 1998; Philipson and Hedges, 1998), but less attention has been paid to the selection inherent in using experts as subjects. The experimental design presented above cleanly identifies differences between the standard undergraduate population and the population of successful managers. To some extent I can also identify differences between undergraduates who intend to become managers and successful managers by using information about the undergraduates' majors. With a larger sample and the use of an appropriate post-experimental survey, this identification could be made even stronger. Where I clearly fail is in identifying the difference between undergraduates who will *eventually* become successful managers and individuals who are *currently* successful

managers. This difference, if cleanly identified, make it possible to measure learning effects separately from any selection effects.

The problem, of course, is that I need to know (at least probabilistically) which undergraduates will become successful managers. The optimal way to solve this problem would be a longitudinal study in which experiments were run on undergraduate students and then their career paths were tracked over an extended period of time. For the strongest possible controls, follow-up experiments would be run after the subjects had been in their respective careers for a lengthy period of time (at least a decade to allow for multiple rounds of promotions). Unfortunately, such a study is far from practical.

As an alternative, I suggest taking advantage of the vast database available for students in MBA programs. Students entering full-time MBA programs (like Case's) provide an interesting contrast to those entering the EMBA program. Like EMBA students, MBA students intend to achieve success as businessmen, primarily as managers. Unlike EMBA students, MBA students are young and haven't yet either acquired substantial experience or faced significant selection beyond the admission process. A wealth of information is gathered about students for MBA programs like Case's. Virtually all programs have standard information like undergraduate GPA, GMAT scores, and previous work experiences. At Case, a number of non-standard psychological measures are gathered as well when students enter the MBA program. During their careers in the MBA program, students make numerous decisions about what classes to take and what areas to specialize in. All of this information (along with grades) is kept by the school. Virtually all business schools track of their alumni for fund-raising purposes. For example Case gathers information about the alumni's current employment (firm name, job title, and line of work).[54] While not perfect, using this information gives us a reasonably good measure of whether an alumnus has become a successful manager.

The extensive information available about MBA students suggests an empirical strategy. Rather than using the standard undergraduate population as naïve subjects, MBA students with three years or less of work experience could be used (at Case this would include most of the full-time MBA students). Using data on MBA alumni, a model can be estimated for how likely an individual is to become a successful manager conditional on information that can be observed for MBA subjects. We can use this information to estimate a likelihood that each MBA subject will eventually become a successful manager. This provides us with a set of weights that can be used to infer from experimental data gathered using MBA subjects the most likely distribution of outcomes for MBA subjects *who eventually become successful managers*. Comparing this inferred distribution of outcomes with the observed outcomes for successful managers makes it possible to cleanly identify the effects of learning as opposed to selection.

This suggested approach – using other data available about MBAs to model the selection process that differentiates EMBAs from naïve subjects – cannot be applied to the current dataset. Even if it was possible to retroactively gather the necessary information about the UG managers (which it isn't), our sample is too small and doesn't contain enough individuals who were likely to pursue careers in management. However, I feel that this approach could be quite fruitful in future studies.

---

[54] This data must be used with some care as the information is self-reported. This introduces problems both in term of errors in measurement and selection (as some alumni do not respond).

Appendix

The pagination of the instructions in this appendix corresponds to how the screens would have been broken up for subjects.

Introduction: Please read these instructions carefully. Take as much time as you need to complete the instructions. If you have a question, please raise your hand and one of us will come to your desk to answer it.

All payoffs during the experiment are denominated in an artificial currency, Experimental Currency Units. At the end of the experiment, Experimental Currency Units will be converted at a rate of 500 Experimental Currency Units equal one dollar.

You can view these instructions during the game by clicking on the help button.

Parts, Rounds, and Firms: This experiment will have two parts. In the first part there are ten rounds and in the second part there are twenty rounds. For this experiment you will be randomly assigned a firm with four other participants. You will be grouped with the same four participants for all thirty rounds.

Roles and Tasks: Each firm has 1 manager and 4 employees. We have already selected which participants will be managers and which will be employees. If you do not know if you are a manager or an employee, please raise reference the left side of your screen. The managers and employees are logged into the system from separate locations. You will be given the instructions for both roles.

Each round of the experiment can be thought of as a workweek. Prior to the start of the week, the firm manager selects a bonus rate as explained below. Each employee spends 40 hours per week at their firm. After seeing the bonus rate, employees have to choose how to allocate their time between two activities, Activity A and Activity B. They can choose to spend 0, 10, 20, 30, or 40 hours on activity A. Their remaining hours will be spent on activity B. For example, if an employee spends 30 hours on Activity A, the remaining 10 hours will be spent on Activity B. Weekly payoffs for both managers and employees depend on the bonus rate and the number of hours allocated to Activity A by the employees.

Firm Manager Payoff: The firm manager's payoff for a round is determined by the bonus rate (B) he/she chooses and the *minimum* number of hours an employee in his/her firm spends on Activity A. (The minimum is the smallest number of hours chosen for Activity A. For example, suppose the four employees choose to spend 20, 10, 40, and 40 hours on Activity A. The minimum hours spent on Activity A is 10 hours.) The manager's revenue is 60 Experimental Currency Units multiplied by the *minimum* number of hours an employee spends on Activity A. As a cost deducted from this revenue, the manager must pay each employee the bonus rate multiplied by the *minimum* number of hours an employee spends on Activity A. The manager also automatically gets a flat payoff of 100 Experimental Currency Units in each round. For example, suppose the firm's employees spend 30, 20, 40 and 40 hours on Activity A and the bonus rate equals 8. The minimum hours spent on Activity A is 20 hours. The firm manager's payoff equals $100 + 60*20 - 4*(8*20) = 660$ Experimental Currency Units. The first term in this equation is the firm manager's flat payoff. The second term, $60*20$, gives the manager's revenue. The third term, $4*(8*20)$, gives the manager's cost in bonuses per employee (160 experimental currency units) multiplied by the number of employees (4) in the firm.

These payoffs are summarized by the formula below. The variable B gives the bonus rate and the function $min(H_A)$ is the smallest or minimum number of hours an employee of the firm spends on Activity A. If you do not find this formula useful, don't worry about it. It is given to you as an additional way to understand the payoffs. For firm managers, the computer will show your payoff table at any point where you need to make a decision. These tables include all the information you will need to make your decision – a sample will be shown later in the instructions. For employees, the program displays the firm manager's payoff function any time you make a decision. For both roles, the computer calculates the firm manager's payoffs for you as part of the feedback you receive after each round.

$$\text{Firm Manager's Payoff} = 100 + (60 - 4B) * Min(H_a)$$

For the first part (Rounds 1 – 10) of the experiment the bonus rate is selected by the computer, not by the firm manager. For the second part (Rounds 11 – 30) of the experiment, the firm manager will select the bonus rate. *The firm manager will only be paid for the part of the experiment where they select the bonus rate.*

Employee Payoffs: The payoff for an employee of the firm is determined in each round by what bonus rate (B) the firm manager chooses, how many hours the employee spends on activity A, and the minimum number of hours employees in his/her firm spend on Activity A. The employee's payoff is reduced by 5 Experimental Currency Units per hour he/she spends on Activity A. The employee also receives the bonus rate multiplied by the minimum number of hours an employee in his/her firm spends on Activity A. Each employee also automatically gets a flat payoff of 200 Experimental Currency Units in each round. For example, suppose an employee spends 10 hours on Activity A. Suppose the other three workers in his/her firm spend 20, 40, and 40 hours and the bonus rate equals 8. The minimum hours spent on Activity A is 10 hours. The employee's payoff equals 200 - 5*10 + 8*10 = 230 Experimental Currency Units.

These payoffs are summarized by the formula below. The variable $H_A(i)$ gives the number of hours spent on Activity A by employee i. Once again, the variable B gives the bonus rate and the function $min(H_A)$ is the smallest or minimum number of hours an employee of the firm spends on Activity A. If you do not find this formula useful, don't worry about it. It is given to you as an additional way to understand the payoffs. For firm managers, the program always displays the employees' payoff function when you make a decision. For employees, the computer always shows your payoff table at any point where you need to make a decision. These tables include all the information you will need to make your decision – a sample will be shown later in the instructions. The program calculates your payoff for you as part of the feedback you receive after each round.

$$\text{Employee i's Payoff} = 200 - 5*H_A(i) + B*min(H_A)$$

Playing a Round as a Firm Manager: When it is time for you to set a bonus rate, the computer will display a screen like the one below. Choose a bonus rate using the buttons on the top of the screen. When you have made your final decision, click on the button labeled "Next". You will then be asked to confirm your choice.

Manager Payoff Table

		Bonus Rate									
		◐	◐	◐	◐	◐	◐	◐	◐	◐	◐
		6	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Minimum Hours Spent On Activity A By Employees	**0**	100	100	100	100	100	100	100	100	100	100
	10	460	420	380	340	300	260	220	180	140	100
	20	820	740	660	580	500	420	340	260	180	100
	30	1180	1060	940	820	700	580	460	340	220	100

$$\text{Employee Profit} = 200 - (5 * Ha) + (B * Min(Ha))$$

Note that when you choose a bonus rate you will not know how many hours the firms' employees will devote to activity A.

You can also type a message for the employees in a box that will appear below the table. This message will be sent when you confirm your choice of a bonus rate. Except for the following restrictions, you can type whatever you want in this box.

Restrictions on Messages

1. Please do not identify yourself or send any information that could be used to identify you (e.g. age, race, professional background, etc.).
2. Please refrain from using obscene or offensive language.

Playing a Round as an Employee: For each round the computer will display a screen like the one shown below. The payoffs shown in the table as well as the formula for the manager's profit will be adjusted for the changing values of B. For the example below we set B = 8. Notice that this is displayed below the table.

To choose a number of hours to spend on Activity A use the button on the left hand side of the screen. When you have made your final decision, click on the button labeled "Continue". You will then be asked to confirm your choice.

Employee Payoff Table

			Minimum Hours Spent by Other Employees on Activity A				
			0	**10**	**20**	**30**	**40**
My Hours On Activity A	◖	**0**	200	200	200	200	200
	◖	**10**	150	230	230	230	230
	◖	**20**	100	180	260	260	260
	◖	**30**	50	130	210	290	290

$$B = 8$$
$$\text{Firm Manager Payoff} = 100 + (60 * \min(H_A))$$

Note that when you make your choice you will not know the choices of other employees in your firm.

Feedback as a Firm Manager: At the end of each round you will receive a summary of what happened in the round including the bonus rate, the minimum number of hours spent on Activity A, your profit, the amount you paid to employees in bonuses, and your accumulated payoffs through the current round. The computer also provides a summary of this information for preceding rounds. For rounds where the computer selects the bonus rate, it displays a "sample profit" rather than a profit. You are not being paid in these rounds, but the sample profit lets you know what you would have earned.

Feedback as an Employee: At the end of each round you will receive a summary of what happened in the round including the number of hours you spent on Activity A, the minimum number of hours spent on Activity A, the bonus rate, the firm manager's payoff, your payoff for the latest round, and your accumulated payoffs through the current round. The computer also provides a summary of this information for preceding rounds.

Confidentiality and Payoff: At no point in time will we identify the firm manager or any of the employees in the firm. In other words, the actions you take in this experiment will remain confidential. At the end of the experiment, your accumulated payoffs will be converted from Experimental Currency Units to dollars at a rate of 500 Experimental Currency Units equals one dollar. You will be paid these converted earnings in cash along with the show-up fee of $10. You will be paid privately and we will not disclose your payoff to the other participants in the experiment. If you are participating on-campus you will be paid immediately; otherwise, you will be paid in your next scheduled economics class.

All subjects playing in the role of *firm managers* will be given access to a list giving the payoffs for all participants in this role. This list will be ordered from the highest to the lowest payoff, but will not include any identifying information about the subjects. This information will only be made available after all the relevant sessions have been completed.

References

Alevy, J., M. Haigh, and J. List (2005), "Information Cascades: Evidence from a Field Experiment with Financial Market Professionals," forthcoming *Journal of Finance*.

Brandts, J. and D. Cooper (2005a), "A Change Would Do You Good. An Experimental Study on How to Overcome Coordination Failure in Organizations," forthcoming *American Economic Review*.

Brandts, J. and D. Cooper (2005b), "It's What You Say Not What You Pay: An Experimental Study of Manager-Employee Relationships in Overcoming Coordination Failure," mimeo.

Brandts, J. and D. Cooper (2005c), "Observability and Overcoming Coordination Failure in Organizations. An Experimental Study," mimeo.

Camerer, C. (2003) *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton, NJ: Princeton University Press.

Capra, C. M., T. Tanaka, C. Camerer, L. Munyan, V. Sovero, L. Wang, and C. Noussair (2005), "The Impact of Simple Institutions in Experimental Economies with Poverty Traps", mimeo.

Cooper, D. and J. Kagel (2005), "Are Two Heads Better Than One? Team versus Individual Play in Signaling Games," *American Economic Review*, **95**, 3, 477 - 509.

Cooper, D., J. Kagel, W. Lo, and Q.L. Gu (1999), "Gaming Against Managers in Incentive Systems: Experimental Results with Chinese Students and Chinese Managers," *American Economics Review*, **89**, 4, 781 - 804.

Cooper, R., D. DeJong, R. Forsythe, and T. Ross (1989), "Communication in the Battle of the Sexes Game: Some Experimental Results," *RAND Journal of Economics,* **20**, 568-587.

Cooper, R., D. DeJong, R. Forsythe and T. Ross (1992), "Communication in Coordination Games," *Quarterly Journal of Economics*, **107**, 739-771.

Cox, D. R. (1972), "Regression Models and Life-Tables (with discussion)", *Journal of the Royal Statistical Society*, **34**, 187-220.

Dyer, D. and J. Kagel (1996), "Bidding in common value auctions: How the commercial construction industry corrects for the winner's curse," *Management Science* **42** 1463-1475.

Fehr, E. and J. List (2004), "The Hidden Costs and Returns of Incentives-Trust and Trustworthiness Among CEOs," *Journal of the European Economic Association*, **2**, 5, 743-771.

Ford, J. and L. Ford (1995), "The role of conversation in producing intentional change in organizations," *Academy of Management Review*, **20**, 541-570.

Foss, N. (2001), "Leadership, Beliefs and Coordination: An Explorative Discussion," *Industrial and Corporate Change*, **10**, 357-388.

Hannan, R., J. Kagel, D. Moser (2002), "Partial gift exchange in an experimental labor market: Impact of subject population differences, productivity differences, and effort requests on behavior." *Journal of Labor Economics*, **20**, 4, 923-951.

Harrison, G. and J. List (2004), "Field Experiments," *Journal of Economic Literature*, **42**, 4, 1009 – 1055.

Ichniowski, C., K. Shaw, and G. Prennushi (1997), "The Effects of Human Resource Management Practices on Productivity: A Study of Steel Finishing Lines," *American Economic Review*, **87**, 3, 291-313.

Knez, M. and C. Camerer (1994), "Creating Expectational Assets in the Laboratory – Coordination in Weakest-Link Games," *Strategic Management Journal*, **15**, 101-119.

Knez, M., and C. Camerer (2000), "Increasing Cooperation in Prisoner's Dilemmas by Establishing a Precedent of Efficiency in Coordination Games," *Organizational Behavior and Human Decision Processes*, **82**, 194-216.

Knez, M. and D. Simester (2002), "Form-Wide Incentives and Mutual Monitoring At Continental Airlines", *Journal of Labor Economics*, **19**, 4, 743-772.

Kotter, J. (1996), *Leading Change*, Boston: Harvard University School Press.

Kremer, M. (1993), "The O-Ring Theory of Economic Development," *Quarterly Journal of Economics,* **107**, 551- 575.

Liang, K., and S. Zeger (1986), "Longitudinal Data Analysis Using Generalized Linear Models," *Biometrika*, **73** , 13-22.

Lichtenstein, S., Slovic, P. (1973), "Response-induced Reversals of Preference in Gambling: An Extended Replication in Las Vegas," *Journal of Experimental Psychology*, **101**, 16-20.

McDaniel, T. and E. Rutström (2001), "Decision Making Costs and Problem Solving Performance," *Experimental Economics*, **4**, 2, 145–161.

Merritt, J. (2003), *BusinessWeek Guide to The Best Business Schools 8th. Edition*, New York: McGraw-Hill.

Montmarquette, C., J-L Rulliere, M-C Villeval, and Zeiliger R, "Redesigning teams and incentives in a merger: An experiment with managers and students," *Management Science*, **50**, 10, 1379-1389.

Philipson, T. and L. Hedges (1998), "Subject Evaluation in Social Experiments," *Econometrica*, **66**, 2, 381–408.

Potters, J. and F. van Winden (2000), "Professionals and students in a lobbying experiment - Professional rules of conduct and subject surrogacy," *Journal of Economic Behavior and Organization*, **43**, 4, 499-522.

Rutström, E. (1998), "Home-Grown Values and the Design of Incentive Compatible Auctions," *International Journal of Game Theory*, **27**, 3, pp. 427–41.

Sklaroff, S. (2004), *US News and World Report 2005 Edition America's Best Colleges*, New York: William D. Holliber.

Smith, V., G. Suchanek, and A. Williams (1988), "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets," *Econometrica*, **56**, 1119-1151.

Van Huyck, J., R. Battalio, and R. Beil (1990), "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review*, **80**, 234-248.

Van Huyck, J., A. Gillette, and R. Battalio (1992), "Credible Assignments in Coordination Games," *Games and Economic Behavior*, **4**, 4, 606-626.

Weber, R., C. Camerer and M. Knez (2004), "Timing and Virtual Observability in Ultimatum Bargaining and 'Weak Link' Coordination Games," *Experimental Economics*, **7**, 25-48.