



**University of
Zurich** ^{UZH}

University of Zurich
Department of Economics

Working Paper Series

ISSN 1664-7041 (print)
ISSN 1664-705X (online)

Working Paper No. 35

**Maintaining Efficiency While Integrating Entrants
From Lower-Performing Groups:
An Experimental Study**

Timothy C. Salmon and Roberto A. Weber

Revised version, July 2014

Maintaining Efficiency While Integrating Entrants From Lower-Performing Groups: An Experimental Study*

Timothy C. Salmon[†]
Southern Methodist University

Roberto A. Weber[‡]
University of Zurich

July 2014

Abstract

Efficiently growing a group or firm often requires integration of individuals from lower-performing entities. We explore the effectiveness of two policies intended to facilitate such integration, using a laboratory experiment that models production as a coordination game with Pareto-ranked equilibria. We initially create an efficient group and an inefficient one. We then allow individuals to move into the high-performing group and vary by treatment whether movement is unrestricted, limited to one entrant per period, or subject to an entry exam. We include two additional treatments that combine the two restrictions in different ways to help understand why the institutions are effective in maintaining coordination. We find that both restrictions work to maintain efficient coordination but they are effective for different reasons.

JEL Codes: C72, C92, M12

Key Words: Growth, Entry, Coordination, Experiments

*The authors would like to thank the National Science Foundation for research support in funding these experiments. We also gratefully acknowledge support from the research priority program at the University of Zurich "Foundations of Human Social Behavior."

[†]Department of Economics, Southern Methodist University, 3300 Dyer Street, Suite 301E, Umphrey Lee Center, Dallas TX 75275-0496; tsalmon@smu.edu; phone: (214) 768-3547; fax: (214) 768-1821.

[‡]Department of Economics, University of Zurich, Blümlisalpstrasse 10, 8006 Zurich; roberto.weber@econ.uzh.ch; phone: +41 (0)44 634 36 88; fax: +41 (0)44 634 49 07 .

1 Introduction

An important issue in the pursuit of large efficiently-performing groups is how to incorporate new entrants. For example, a growing economy might need to rely on immigration to provide a suitable labor supply. Similarly, profitable firms often expand by integrating individuals through external hiring.

In many cases, the growing group is one that has previously performed well on some important dimension, such as productivity or profitability, leading to a desire to maintain such high performance while undergoing growth. Incumbents in the growing entity have a history of working well together, abiding by efficiency-enhancing norms, expecting high performance from one another, and thereby obtaining coordination on an efficient set of behaviors. Conversely, entrants are often from lower-performing entities, such as poorer nations or less profitable firms, where expectations of others' behavior are lower and where the resulting behavior is coordinated on less efficient equilibria. This creates an important concern: when incorporating entrants from low-performing groups, how can growing entities maintain the high performance that has made them successful? Specifically, in order to grow successfully while incorporating new entrants, it must be that both (i) new entrants adapt their behavior to the higher standards of the better-performing group and also that (ii) incumbents in the high-performing group maintain the good behavior that has previously led to success.

In this paper, we use a laboratory experiment to explore two common policies for facilitating the successful entry of individuals from lower-performing groups into ones that have historically performed more efficiently. The first policy we examine is an entry exam, whereby individuals are required to demonstrate knowledge or skills prior to being granted entry. For example, countries often set skill-based thresholds for granting entry to immigrants, such as sufficiently high professional or educational standing, or may implement language proficiency requirements or naturalization exams as a requirement for citizenship. A club may require entrants to demonstrate knowledge of the club's history or traditions prior to entry. Finally, many firms use pre-employment testing as a way of screening potential job applicants for desirable qualities. The second policy is an entry quota, whereby the number of entrants in a given time period is regulated. For example, a country may limit the number of immigrants that are allowed entry in a given year, as the United States did with the 1924 Immigration Act. Similarly, a club may limit the number of new members it accepts. Or a firm may explicitly target a slow-growth strategy as a way to maintain a corporate culture or existing efficient practices.

Of course, as the examples above reveal, actual policies used to regulate entry into groups vary widely, as do the specific contexts to which they are applied. Our experiment necessarily abstracts away from the rich details of the analogous real-world mechanisms to study the broad effectiveness of these kinds of policies. We also abstract from the many ways in which the “successful” and “unsuccessful” groups differ — such as production technologies, resources and institutions other than mechanisms regulating entry. While such differences are surely important for the success of those entities, we simplify our experimental environment to focus directly on how the institutions for regulating membership may affect behavior. We use the control afforded by the laboratory to create two groups that differ primarily on a single dimension, based on coordination on different behaviors and expectations.

Specifically, we focus on a production context in which there is high action interdependence — using a “minimum-effort” or “weak-link” coordination game, in which efficiency (output) is determined by the lowest input by any group member. Thus, in order to perform well, a group needs *all* players to make high input choices.¹ This makes it particularly difficult to incorporate new entrants since efficiency can be harmed if either any entrant or any incumbent selects lower input. Growth in such environments may also be particularly difficult because larger group sizes have a harder time obtaining efficient coordination (see Van Huyck et al., 1990; Weber 2006).

In a standard version of this coordination game there is some number, N , of players who each select from an ordered set of increasingly costly strategies, $x_i \in (1, \dots, K)$, and produce some collective output that is an increasing function of the minimum choice among all the players. Players benefit from higher output levels, but selecting higher strategies is costly, so players only want to select a higher strategy if all other players do so as well. The set of pure-strategy Nash equilibria for the game consists of all outcomes in which players select the same strategy. The incentives are such that equilibria corresponding to higher effort choices generate higher welfare than those in which players make lower choices and the unique Pareto-dominant equilibrium corresponds to all players selecting strategy K .

Our version of the game is augmented by including an element that leads to the group output being an increasing, concave function of the number of players in the group. Thus,

¹While situations that exactly mirror this game — i.e., in which production is determined by the minimum input — are rare, the game is a useful model for contexts in which the level of interdependence is sufficiently high that output is sensitive to the actions of small numbers of individuals. Games of this form have been applied to study diverse contexts, such as firm production (Brandts and Cooper, 2006), corruption (Andvig and Moene, 1990), macroeconomic performance (Cooper and John, 1988), public goods provision (Hirschleifer, 1983), and shared security (Kunreuther and Heal, 2003).

holding the minimum choice in the group constant, output is greater when the group is larger. This creates a *ceteris paribus* efficiency gain from having larger groups. It is still the case, though, that growth in the size of the group decreases earnings if it leads to coordination on a lower-output equilibrium.

In our experiment, we divide each group into two endogenously-determined subgroups — one “*High*-performing” subgroup which is smaller than the other “*Low*-performing” subgroup. Our procedure reliably creates small, efficiently-coordinated subgroups and larger subgroups coordinated on inefficient levels of output.

We then allow participants from the Low subgroup to voluntarily move to the High subgroup. Given our payoff structure, such movement can lead to a Pareto improvement whereby both incumbents in the High subgroup and entrants from the Low subgroup can earn higher payoffs. But this improvement is only realized if entry does not lead to lower output levels. That is, growth is only successful if entrants from the Low subgroup adjust their strategy choices upwards while High-subgroup incumbents also persist in making high choices as the group grows.

Our subgroups can therefore represent, for example, very simple laboratory abstractions of a large country or firm coordinated on inefficient practices, and a smaller one that is more successful. Movement from the large subgroup into the smaller one can make everyone better off, but only if coordination on high output levels persists throughout the growth process. To our knowledge, ours is the first experimental design that repeatedly creates two such subgroups, and then uses them to explore the effects of movement from the inefficient group into the efficient one.

Our primary purpose is to explore the relative effectiveness of alternative mechanisms for regulating entry into the successful subgroup. We consider five treatments, based on two main forms of entry regulation. In a *Baseline* treatment, we impose no restrictions, allowing anyone who wants to move to do so. In a *Quota* treatment, only one participant is allowed to move per period, with multiple requests accepted or declined by a random lottery. In a *Quiz* treatment, there is no limit on the number of entrants, but all entrants must successfully complete a short quiz demonstrating both the ability to calculate game payoffs and knowledge of historical outcomes in the High subgroup.

We then combine these two restricted entry treatments in different ways to help us understand how each is effective. In one treatment we call the *Combined* treatment, both restrictions are simultaneously in place. This allows us to explore whether the combination of the two mechanisms yields any additional benefit and provides an indication of whether they operate through different underlying processes.

Finally, after initially conducting sessions with the above four treatments, we added an additional treatment to shed light on the mechanism underlying the effectiveness of the *Quiz* treatment. This *Quiz Based Quota* treatment implements a quota mechanism in which the number of entrants allowed in each period is based on the average number of subjects who were able to move between groups in the Quiz treatment, but those entrants need not pass the quiz in order to move. Hence, in comparison with the Quota treatment, which grows at a slower rate, the Quiz Based Quota treatment allows us to test whether a less restrictive quota is equally effective, or whether the rate of growth matters. Moreover, in comparison with the Quiz treatment, which grows at the same rate as the Quiz Based Quota but screens entrants to allow only those who pass the quiz to move, this treatment allows us to identify whether such a screening function is important for the effectiveness of the Quiz mechanism, or whether it is simply the fact that the quiz slows down growth relative to the Baseline that is sufficient.

Previous research in economics, organizations, and sociology explores the determinants of successful assimilation of new entrants into entities such as countries, firms, or work teams (Borjas, 1995; Cable and Judge, 1996; Card, 2005; Chatman and Flynn, 2001; Friedberg, 2001; Ortega, 2005; Rumbaut, 1997). Some of this work explicitly addresses the need for and effectiveness of policies like those we study here to regulate entry. For example, Warne (1921) argued for the importance of government policies to regulate and ensure the assimilation of immigrants into U.S. society, noting that, “If the volume of total immigration is likely at any time to become so large as to make ineffective our forces of assimilation, then this volume should be decreased by restrictive measures; if the elements comprising any particular group or race, no matter how few in number relatively, are unassimilable into our American life then these too should be restricted” (p. 185). More recently, Borjas (1994) discussed how changes to a points-based scoring system used to evaluate immigrants to Canada significantly affected the educational attainment of admitted immigrants. In an organizational context, Ryan and Sackett (1987) demonstrated that pre-employment screening can be useful for discriminating between honest and dishonest employees. O’ Reilly III, Chatman, and Caldwell (1991) demonstrated that greater correspondence between an individual’s characteristics and the organization’s culture yields more successful outcomes in hiring, such as longer employee retention.

Our work adds to this literature, though from a novel perspective, by using a laboratory experiment to study the effectiveness of different policies for regulating entry. While the laboratory necessarily sacrifices some degree of generalizability to these real-world contexts, it also allows us to avoid concerns that are present in the field. For example, in

evaluating the effectiveness of different policies for regulating entry in real-world groups, the endogenous selection of policies makes inference about whether outcomes are truly the result of a particular policy difficult. In the laboratory, however, we can avoid this concern by randomly-assigning groups to different policies.

Our work is more closely related to a small number of experimental studies that explore alternate methods for determining group composition, and their effect on efficiency in production games. For example, Weber (2006) shows how groups that start of small and grow can maintain efficient coordination. While closely related to our work, there are at least two important ways in which the experiment in Weber (2006) leaves open key questions about how to manage the growth of groups outside of the lab. First, entry in Weber (2006) drew from participants with no prior history of playing the game. This is unlike many external situations, in which entry occurs by drawing from people from lower-performing environments where there may be a history of failure or inefficiency. In our experiment, entry involves incorporating participants with a strong history of inefficient outcomes and behaviors. This difference in the prior experience of entrants is important since it may very well be the case that the slow growth in Weber (2006) proves ineffective when incorporating entrants from lower-performing groups, and that these entrants' negative history proves too strong to overcome. It is important to understand how such characteristics of entrants and their experience can affect their assimilation. Second, Weber (2006) does not consider alternative mechanisms for regulating entry, as we do here. Thus, the experiment in Weber (2006) serves primarily as a demonstration that slow growth *can* facilitate efficient coordination in large groups, at least large by laboratory standards, but does not study different ways to grow. Here, we begin to compare different ways to achieve the goal of efficient growth.² McCarter and Sheremeta (2013) also examine the question of how group behavior is affected by the introduction of new members. In that study, two members of a four person group are replaced to determine if a group can maintain coordination after the introduction of new members. While related to our work, their paper does not consider mechanisms for forming groups like those we study here, does not examine the effect of growing groups and also does not incorporate the key element of individuals moving from unsuccessful to successful groups that is necessary to examine our questions of interest.

In other related work, Ahn, Isaac and Salmon (2008, 2009) use a public goods envi-

²There are also other differences between our experiment and Weber's. For example, in Weber (2006), entry is exogenously determined, according to a growth schedule specified *ex ante* by the experimenter and commonly known by all participants. In contrast, in our experiment the movement of subjects across groups takes place more naturally: subjects decide whether they wish to move to the more efficient group, and then face any entry regulations prior to successfully doing so.

ronment to explore how endogenous group formation influences contributions, with players having the opportunity to enter and exit groups. Their experiments vary the degree to which both entry and exit from groups are restricted, finding that both restricted entry and exit significantly influence contributions. More closely related to our experiment, Riedl, Rohde and Strobel (2012) similarly study endogenous group formation, but in the context of a weak-link coordination game similar to the one we use. They also find that voluntary association can produce groups that are large and efficiently coordinated, and that the key mechanism is the ability of those acting efficiently to exclude those whose past behavior corresponds to an inefficient equilibrium.³ Viewed jointly, these studies all show that groups can grow efficiently when they can exclude specific individuals, conditional on their past behavior. Our experiment studies the effect of a different set of policies, which do not allow entry to be conditioned on past individual behavior, but rather on limitations imposed directly as part of the entry process. Our study is therefore informative for situations in which it is difficult to exclude individuals based on prior conduct, as when such behavior is difficult to observe.

2 Experimental Design

In our experiment subjects interacted through computers while playing a version of a weak-link coordination game over a total of 25 periods.⁴ Earnings from the game were denominated in Experimental Currency Units (ECU), which were converted to money at the end of the experiment. The composition of the group within which subjects played the game varied across periods.

2.1 The Game

In each period, subjects had 7 tokens that they could either keep or contribute to a group account. We will let $x_i \in \{0, 1, \dots, 7\}$ denote the amount contributed. For every token retained the subject received 3 ECU. The total return on the tokens contributed to the

³Other experimental studies identify alternative mechanisms that improve coordination by creating mutual confidence among players in coordination games similar to the weak-link game. Some of these mechanisms include including auctioning the right to play (Van Huyck, et al., 1993), strengthening group identity (Chen and Chen, 2011), or allowing leaders to recommend strategies (Brandts, Cooper and Weber, forthcoming). It is also worth noting that there is significant heterogeneity both in the effectiveness of different mechanisms, and in group outcomes observed within a particular mechanism. For example, many groups in Weber’s (2006) study that grow slowly nevertheless fail to maintain efficient coordination. For a review of this literature, see Devetag and Ortmann (2007).

⁴The software used for this experiment was programmed in Z-Tree (Fischbacher, 2007).

group account depended on the number of players in the group, N , and the minimum of the tokens contributed among the group (N). Let x_{-i} represent the vector of contributions of all of the other members of i 's group. Then the total return on the group contribution was $6 * N^{.25} * \min(x_i, x_{-i})$. We will refer to $6 * N^{.25}$ as the group account multiplier since it indicates the per token return for the minimum number of tokens contributed to the group account. Finally, contributing to the group account cost an individual $x_i^{1.5}$ ECU.⁵ The entire payoff function is contained in equation 1.

$$\pi(x_i, x_{-i}) = 3 * (7 - x_i) + 6 * N^{.25} * \min(x_i, x_{-i}) - x_i^{1.5} \quad (1)$$

The properties of this game are standard for weak-link games as described above. For any group size $N \geq 2$, any level of contribution is a Nash equilibrium if all players contribute the same amount. This is so because a player does strictly worse by contributing above the minimum contribution of others than by simply matching the minimum, while if all other players contribute a common amount above a player's contribution the player can do strictly better by matching the minimum contribution of others. All players contributing 7, or the maximum number of tokens, is the Pareto-dominant equilibrium. To make certain that subjects would be able to understand the incentives generated by this payoff function we provided them with multiple payoff tables corresponding to differing values of N and we provided the ability in the software to calculate prospective earnings by entering a hypothetical minimum for others as well as an own contribution to see the resulting payoff for as many combinations as they liked prior to making a decision in a period.

2.2 Stages

Each session involved 20 or 30 subjects who were split into separate 10-person groups. These groups were fixed for the duration of the session, which lasted 25 periods, and there was never any interaction across those groups. The experiment proceeded in three stages.

In Stage 1, which consisted of Periods 1-3, all members of a 10-person group played the weak-link game together. In the instructions for this stage, subjects were given a payoff table, corresponding to $N = 10$, to which they could refer in making their choices.⁶ At the end of every period in Stage 1, subjects saw their own contribution, the minimum contribution in the 10-person group, the group account multiplier for the group, and their

⁵Contributing also incurs the opportunity cost of 3 ECU for each token no longer retained. In the instructions, we described the "cost" as $x_i^{1.5}$, but subjects were also told of the value of retained tokens, allowing them to calculate the opportunity costs.

⁶Copies of all instruction scripts are available from authors.

own earnings for the period. This stage lasted for three periods because the point of the stage was to allow subjects to demonstrate some degree of heterogeneity in terms of willingness to choose high numbers but without those attempting to choose the high numbers becoming too affected by those who chose low numbers.

Stage 2 took place over Periods 4-10. At the beginning of Period 4, each 10-person group was split into two subgroups that we called the “High” and “Low” subgroups. The High subgroup consisted of the two members from the 10-person group who had contributed the most in Stage 1. The Low subgroup consisted of the other 8 members, who were the lowest 8 contributors during Stage 1.⁷ Subjects received additional verbal and written instructions about the new subgroups, including additional payoff tables corresponding to values of $N = 2$ and $N = 8$. The experiment software told subjects to which group they were assigned and gave them a complete Stage 1 contribution history of their fellow group members. This history was provided primarily as a benefit to the High group members, to make it easier for them to coordinate by allowing them to see the prior choices of the person with whom they were matched in Stage 2. To further facilitate separation between the two subgroups, the instructions used the labels “High” and “Low” to refer to the subgroups, and we were also explicit about how the subgroups were determined. Subjects then began Stage 2 of the experiment, playing the weak-link game with only those in their 2- or 8-person subgroup. After every period in Stage 2 subjects saw the same information as before about their own subgroup but they were also given information about what was happening in the other subgroup, including the number of group members, the group account multiplier and minimum group contribution.

Our expectation, based on previous literature on weak-link games, was that the High subgroups would be able to coordinate on a high level of contributions by the end of Stage 2, while the Low subgroups would fail to achieve high coordination levels and would instead play several periods with low minimum group contributions. This was necessary in order to set up Stage 3 of the experiment, in which we allowed subjects in the Low subgroup the opportunity to move to the High subgroup.

Stage 3 lasted from Period 11 through the final period, 25. Prior to each period, subjects still in the Low subgroup were asked if they wished to move to the High subgroup. Moves into the High subgroup were irreversible and subjects who began in the High subgroup were never allowed the option of moving into the Low subgroup. We chose this limited flexibility to allow us to precisely address our main research question of whether successful

⁷Subjects were not told prior to this event that they would be divided in this way so their Stage 1 contributions could not have been made strategically to try to get into the High (or Low) subgroup.

groups could maintain their level of efficient coordination during an influx from the lower performing group. Other than the ability to transition groups, this stage proceeded in the same manner and subjects received the same information at the end of a period as in Stage 2. Stage 3 lasted for 15 periods, as we wanted to allow enough periods for subjects to move between groups if they wanted. Our focus is on how groups behave during and immediately after the integration phase — as this is when the group formation institutions should have their most substantial effect.

2.3 Treatments

The mechanism underlying the group transitions varied by experimental treatment. At the beginning of Stage 3, complete verbal and written instructions were provided to members of both subgroups about the group transition mechanism that would be used. This is important, as it meant that members of the High subgroup knew what restrictions governed entry into their subgroup. Also, at the beginning of each period, following any movement from the Low subgroup to the High subgroup, all subjects saw the number of people in each of the two subgroups and had the opportunity to explore payoffs for their current (and possibly new) group size.

In the first treatment, Baseline, any subject indicating a desire to move from the Low subgroup into the High one was allowed to do so. At the beginning of a period, Low subgroup subjects wishing to move indicated this choice on the computer and were then moved into the High subgroup. Thus, there were no restrictions on movement from the Low subgroup to the High subgroup. The other four treatments involved some degree of restrictions.

In our second treatment, Quota, only one subject was allowed to move from the Low subgroup to the High subgroup per period. This was implemented by the software randomly selecting one subject to be allowed to move from among all those expressing a desire to change subgroups in a period. The unchosen subjects wishing to move were forced to remain in the Low subgroup for that period. As in the Baseline, at the beginning of a period subjects in the Low subgroup again indicated whether they wished to move to the other subgroup. On the next screen, they found out whether they had been able to do so.

In the third treatment, Quiz, subjects had to correctly answer a two question quiz in order to move between subgroups. Any subjects expressing an interest in moving and successfully answering the quiz questions were allowed to move to the High subgroup. One question asked subjects about the history of the High subgroup from the prior period (such

as the size of the subgroup or the minimum contribution level). The other question gave subjects a sample set of contributions and asked them to calculate the payoff for one of the players. We gave subjects extensive instructions on how to do this calculation and the software pre-calculated for them the group account multiplier ($6 * N^{.25}$) as well as the cost of contributing ($x_i^{1.5}$), which made the remaining calculation quite straightforward. We also provided subjects with access to the standard Windows calculator.

Our final two treatments combined elements of the previous two. The Combined treatment imposes both a quiz and a quota restriction. In this treatment, all subjects wishing to transition answered the same two kinds of questions as in the Quiz treatment but only one person among those who got the answers correct was allowed to move. The final treatment, called Quiz Based Quota, imposes a period-by-period quota that creates a similar growth trajectory to that produced, on average, by realized subject movements in the Quiz treatment. Hence, this condition was implemented after the other conditions, by taking the average number of movers in each period from the Quiz treatment and implementing a similar quota for each period of the Quiz Based Quota treatment.⁸

We chose our two base institutions — Quiz and Quota — in large part due to the fact that entry restrictions like these are widely used, as we discussed in our introduction. Both mechanisms can potentially facilitate integration, but possibly for different reasons. One might expect the Quota mechanism to be more successful in maintaining coordination than the Baseline because it allows the larger group to slowly assimilate new members. For example, with only one entrant at a time, entrants need not worry about how the behavior of other entrants may affect the minimum, alleviating one possible reason why rapid growth might fail. In the Quiz treatment, it is possible for large numbers of new members to migrate per round, but only subjects who can demonstrate that they understand something about the history of the large group and how the production technology works are able to do so. Such screening might allow a better maintenance of coordination than in the Baseline treatment, by selecting those more likely to be aware of the benefits of matching the historically high contributions in the High subgroup and by affecting the beliefs of the incumbent High subgroup members such that they are willing to take the risk and continue their high contributions.

⁸Specifically, for the first three periods of Stage 3, the average numbers of movers in the Quiz treatment were 2.86, 1.86 and 1.14, respectively. Therefore, we implemented quotas of 3, 2 and 1 in the Quiz Based Quota treatment for these three periods. This allows for slightly larger groups at the end of three periods (6) than in the Quiz treatment (5.86), in order to account for possible cases in which a lower number of subjects attempt to move than allowed by the quota. For all remaining periods in the Quiz treatment, average movement varied between 0 and 1; we implemented this in the Quiz Based Quota treatment by alternating the number of movers allowed between 0 and 1 for the remainder of Stage 3.

While it is possible for both of these mechanisms to be effective, the above discussion highlights that the reasons for their effectiveness could be different. We conducted the two additional treatments to understand better the relationship between the two mechanisms. The Combined treatment was designed to tell us if there is some additive effect of combining the mechanisms. That is, does having the single person who wins the quota lottery also be someone selected by the screening technology in the Quiz treatment lead to a better outcome than just having a single randomly selected person allowed in? The Quiz Based Quota institution was designed after we observed our results from the first four treatments, primarily to understand whether the implicit quota implemented by the screening process in the Quiz treatment — since not everyone who attempted a move was able to pass the quiz — could explain the effectiveness of that treatment. The Quiz Based Quota allows us to see if any mechanism that reduces the speed of entry to that observed with the Quiz mechanism is effective for maintaining coordination, or whether the nature of the Quiz itself contains additional mechanisms that facilitate integration and coordination. Furthermore, the Quiz Based Quota also allows us to test if the rate of restricted entry, which varies between the Quota and Quiz Based Quota, is important for integration. We will address each of these issues in the results section.

2.4 Procedural details

We conducted a total of 15 sessions of this experiment, with a total of 360 subjects in 36 ten-person groups. We obtained data on 7 independent 10-person groups for each of the Baseline, Quota and Combined treatments; 9 groups for the Combined treatment; and 6 groups for the Quiz Based Quota treatment. Sessions were conducted in both the United States and Switzerland.⁹ Earnings from the game, denominated in Experimental Currency Units (ECU), were converted to money at a rate of 60 ECU = 1 U.S. Dollar in sessions conducted in the United States and 16 ECU = 1 Swiss Franc in sessions conducted in Switzerland. These payment parameters were calibrated to match the average subject payment at each location.¹⁰ Subjects also received a participation payment (10 USD or 10 CHF). Average total earnings were 22.03 USD in the U.S. and 59.50 CHF in Switzerland,

⁹Eight of the sessions were conducted at the xs/fs lab at Florida State University, producing four independent 10-person groups for each of the first four treatments (Baseline, Quota, Quiz and Combined). We also conducted one session with two 10-person groups of the Combined treatment at the Pittsburgh Experimental Economics Laboratory (PEEL) at the University of Pittsburgh. In Switzerland, we conducted one additional session each of the first four treatments, each with three 10-person groups and two sessions with three groups each of the of the Quiz Based Quota treatment.

¹⁰The realized average hourly payments also roughly match the average payment for student workers at the respective universities.

for an experiment lasting approximately 2 hours.

3 Results

Recall that, in our experiment, each 10-person group remained fixed for the entire experiment. However, in Stages 2 and 3, the 10-person group consisted of a High and a Low subgroup of potentially varying size. For simplicity, in the rest of the paper we generally refer to these subgroups using the term “group,” and when we use the terms “group minimum” or “group size” we are referring to the subgroups. We explicitly note any instances in which we are referring to the entire 10-person super-group.

Our goal is to study the integration of individuals from lower- into higher-performing groups. Thus, a pre-condition for our analysis is that the 2-person High subgroup be coordinated on a higher minimum contribution level than the corresponding 8-person Low subgroup at the end of Stage 2. Our design was generally successful for inducing this directional difference; in 33 of 36 ten-person groups, the minimum contribution was higher in the final period of Stage 2 in the High subgroup than in the corresponding Low subgroup. Since it is difficult to interpret the effects of the entry mechanisms in the groups for which this was not the case, we remove the remaining three 10-person groups (1 each in the Baseline, Quota and Quiz Based Quota treatments) from consideration in the subsequent data analysis.

3.1 Summary Statistics

We begin by first presenting a series of summary statistics and graphical depictions of the data, before presenting the formal statistical results through a series of regressions. Table 1 presents a first look at the data, summarizing two individual-level variables (average overall contributions and average overall profits) and two group-level variables (average group size and average group minimum) for each treatment. The results are presented separately for Stage 2 (periods 4-10) and for the last 10 periods from Stage 3 (16-25). The reason to present the summary statistics from those ranges is that the Stage 2 numbers demonstrate how well the two kinds of groups perform prior to the experimental treatment, and thus indicates the extent to which our Stage 2 manipulation did indeed produce differentially efficient “Low” and “High” groups, while the last 10 periods of Stage 3 demonstrate the overall effect of the treatment after most of the group transitions have taken place.

There are a number of features that are readily apparent in the summarized data. First, we see that our design regularly induced Low groups coordinated on equilibria with

| | | Contribs. | | Group Min. | | Group Size | | Profits | |
|----------------|------------------|-----------|-------|------------|-------|------------|-------|---------|-------|
| | | 4-10 | 16-25 | 4-10 | 16-25 | 4-10 | 16-25 | 4-10 | 16-25 |
| Low Groups | Baseline | 1.5 | 2.2 | 0.7 | 2.3 | 8 | 2.3 | 21.6 | 25.7 |
| | Quota | 1.6 | 2.7 | 0.9 | 1.9 | 8 | 1.6 | 22.3 | 23.5 |
| | Quiz | 1.1 | 3.3 | 0.4 | 3.4 | 8 | 1.8 | 20.3 | 23.8 |
| | Combined | 1.8 | 0.8 | 1.3 | 0.6 | 8 | 2.4 | 26.4 | 21.9 |
| | Quiz Based Quota | 1.7 | 3.8 | 0.8 | 2.7 | 8 | 1.4 | 21.4 | 26.4 |
| High Groups | Baseline | 5.8 | 1.2 | 5.6 | 0.8 | 2 | 8.5 | 29.0 | 24.4 |
| | Quota | 5.8 | 3.8 | 5.6 | 3.3 | 2 | 9.6 | 29.4 | 35.9 |
| | Quiz | 5.7 | 3.4 | 5.3 | 3.2 | 2 | 9.0 | 27.9 | 35.0 |
| | Combined | 5.0 | 4.0 | 4.8 | 3.7 | 2 | 8.1 | 27.9 | 40.5 |
| | Quiz Based Quota | 6.0 | 2.3 | 5.6 | 2.0 | 2 | 9.3 | 28.3 | 31.7 |

Table 1: Summary statistics (means) from Low and High groups by period and treatment.

low contributions in Stage 2 (the average contribution level is generally between 1 and 2) while the High groups are able to coordinate on more efficient equilibria (the average contribution levels are between 5 and 6). Thus, our Stage 2 manipulation generally worked as we expected.¹¹ Second, in Stage 3, most of the Low group subjects choose to move into the High group, though in every treatment there are subjects who remain in the Low group.¹²

Next, we look at average contributions and minima in Stage 3 for the High groups, by treatment, which is the focus of our study. In the Baseline, contributions collapse to around 1, on average, indicating that unregulated entry has a strongly detrimental effect on efficient coordination. However, in the four treatments with entry restrictions, the High groups are generally able to maintain coordination on equilibria with higher contributions, though with some backsliding relative to their average contributions in Stage 2. The early indication, then, is that all mechanisms involving entry restrictions appear at least somewhat effective at sustaining more successful integration of entrants from the Low group into the High group than the baseline of no restrictions.¹³ We will later provide more precise statistical

¹¹We note that while we dropped three groups for which Stage 2 did not work as expected, were we to include these groups in the summary statistic calculations, the overall picture would not change much. The average for the High groups would still be above the average of the Low groups in all treatments by a significant amount.

¹²The average group sizes for the Low and High groups do not sum to 10 for Periods 16-25. This is because, for all calculations in the table, we omit cases in which there was an empty group (i.e., when all original Low group members had moved to the High group).

¹³The Low group contributions are not our main focus, but we interestingly find that the few remaining subjects in the Low groups by periods 16-25 are sometimes able to improve the efficiency of coordination (reflected both in higher average contributions and group minima, relative to Stage 2).

characterizations of this level of success, including controls for heterogeneity in Stage 2 outcomes, and attempt to determine if any of these mechanisms outperform the others.

We also see that subjects' profits generally follow the pattern indicated by the contributions and minima. That is, average profits for members of the High group near the end of the experiment are higher in the four restricted entry treatment conditions than in the Baseline. Moreover, profits are persistently better for High than Low group subjects in all treatments except in the Baseline. This further reflects the efficient integration that occurs under the treatments that regulate entry.

Interestingly, the aggregate measures reflect a generally consistent ordering between the four mechanisms that restrict entry. The Combined treatment does slightly better in contributions, minima and earnings than the Quota and Quiz mechanisms, which perform roughly equally well, particularly in minimum contributions and average earnings. The Quiz Based Quota, while better on these measures than the Baseline, fares worse than the other three restricted entry treatments. In the next subsection, we provide more detailed tests of these differences.

Figures 1 and 2 present the time paths that illustrate these results, showing the averages of each of the summary statistics for each period in the High groups and by treatment.¹⁴ The trends displayed in the figures display the same general characteristics indicated by the summary statistics. In the Baseline, average contributions and minima drop dramatically between Periods 10 and 11, while in the four treatments with restricted entry the corresponding decreases are more modest. Following the initial part of the growth process, the Combined treatment generally yields the highest individual contribution levels, minima and earnings, while the Quota and Quiz mechanisms perform roughly equally well. Finally, the Quiz Based Quota mechanism, while performing better than the Baseline, generally yields lower individual contributions, group minima and earnings than the other three mechanisms.

Looking at entry into the High group ("Group Size" in Figure 2), we see the time pattern of how group size evolves. The first thing to note is that the growth rate of the Quiz and Quiz Based Quota treatments are very similar, but this is of course by construction. They both end up with quite similar group sizes to those in the Baseline treatment after the first five periods of Stage 3, but the Baseline starts with a much more rapid initial growth rate. The Quota and Combined treatments lead to slower growth, with the latter naturally being slightly slower than the former. However, we observe convergence to similarly large group

¹⁴We do not provide similar figures for the Low groups to conserve space and due to the fact that their behavior following the transition to Stage 3 is not of primary interest.

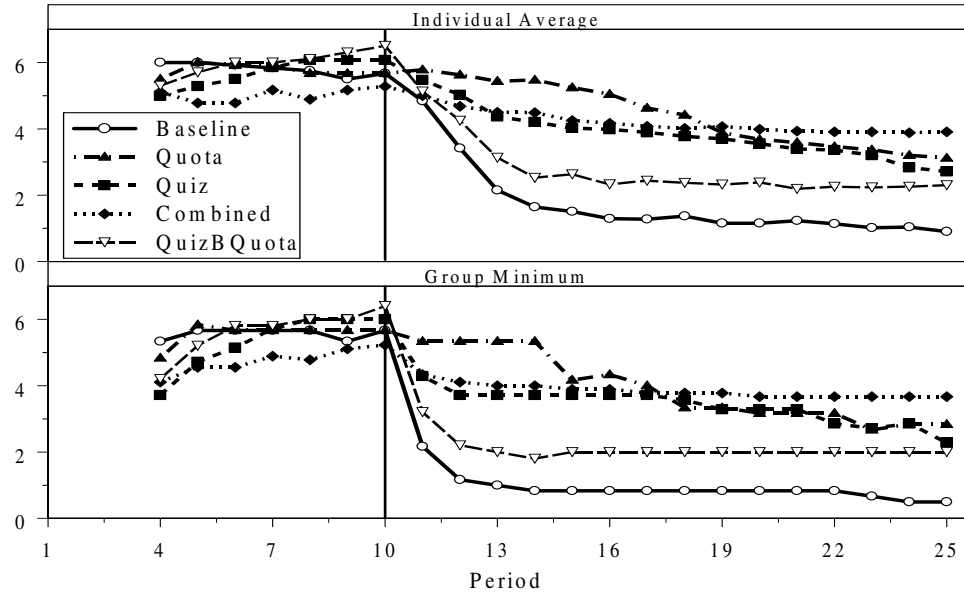


Figure 1: Averages of individual and group minimum contributions for High groups.

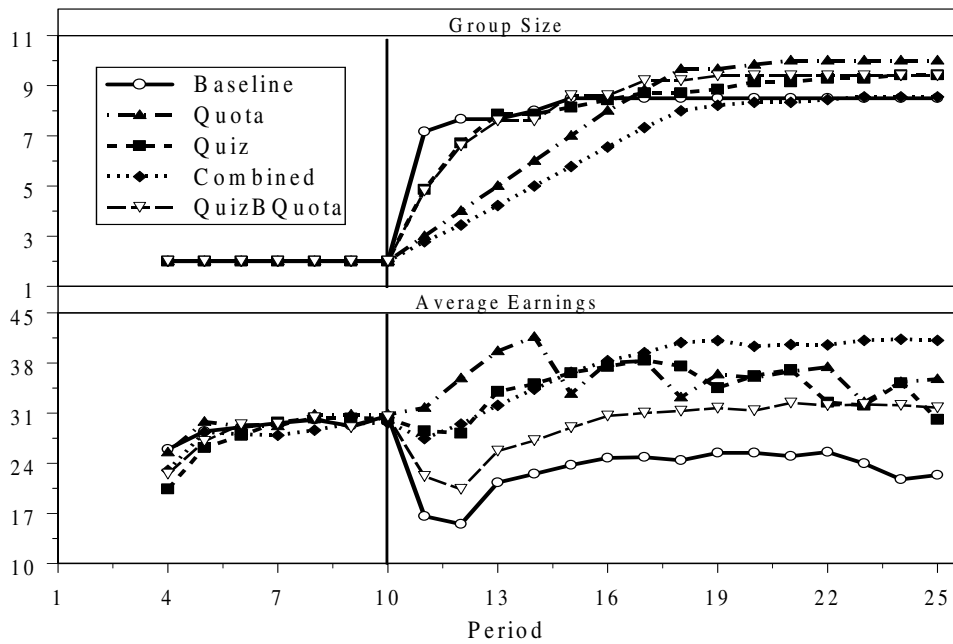


Figure 2: Average group size and individual earnings for High groups.

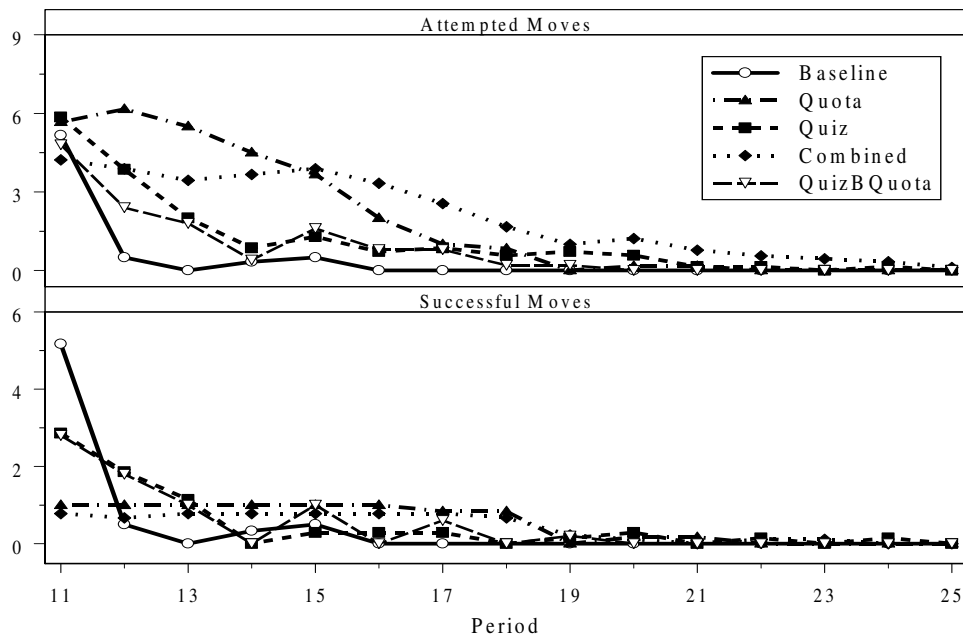


Figure 3: Average attempted and successful moves per period per group with attempted moves in that period.

sizes, between 8 and 10, in later periods across all treatments.

Figure 3 shows more detail on group movement. The graph presents the average number of attempted (top panel) and successful (bottom panel) moves, by treatment, for periods in which at least one move was attempted. In Period 11 there are approximately the same number of attempted moves in all four restricted entry treatments as in the Baseline. In the four treatments with restrictions, many of those initial attempts fail in Period 11, leading there to be substantially fewer successful moves in those treatments than in the Baseline. In particular, we see a clear ordering of the realized movement allowed by the distinct mechanisms: the actual movement in Period 11 of the Quiz and Quiz Based Quota treatments is roughly half of what it is in the Baseline, while in the Quota and Combined treatments it is roughly one-third of what it is in the Quiz and Quiz Based Quota.

After Period 11, there are generally more attempted moves in the four treatments than in the Baseline, at least partly due to the fact that there remain fewer subjects in the Low groups in the Baseline. However, after Period 11, the realized number of moves converge rapidly to similar numbers across all treatments. Therefore, while the treatments limiting movement produce a large difference in the number of successful moves in the first period in which movement is allowed, their effects on total movement in subsequent periods is not substantial.

| | Average of Individual Contributions | | Minimum of Group Contributions | |
|------------------|-------------------------------------|---------------|--------------------------------|---------------|
| | Periods 4-10 | Periods 16-25 | Periods 4-10 | Periods 16-25 |
| Quota | 0.700 | 0.018 | 0.846 | 0.012 |
| Quiz | 0.950 | 0.048 | 0.847 | 0.024 |
| Combined | 0.710 | 0.001 | 0.790 | 0.003 |
| Quiz Based Quota | 0.517 | 0.046 | 0.666 | 0.068 |

Table 2: P-Values from Wilcoxon Rank-Sum tests of differences in distributions of High group averages between treatments and baseline.

3.2 Statistical Tests Regarding Contributions and Earnings

For our first set of basic statistical results we provide a set of non-parametric Wilcoxon rank-sum tests on whether group-level averages differ between the treatments and the Baseline in each phase. We then conduct a more comprehensive statistical analysis, using a difference-in-difference approach to measure the marginal impact of each mechanism controlling for initial heterogeneity in Stage 2 group outcomes.

The p -values resulting from the rank-sum pairwise tests are in Table 2. For each test, an observation consists of the average contribution (or average group minimum) over a stage for each group, meaning we have between 5 and 9 total observations for each treatment. Given that we are aggregating across all periods in a stage for each group, this test represents a very conservative test of treatment differences. Examining the results of the tests in Table 2, we see that there are no differences between either average contributions or group minimum contributions in Stage 2, i.e. periods 4-10. This is important to establish, since the fact that treatments had not been introduced to the subjects by this point means there should be no differences in behavior across treatments. On the other hand, in the last 10 periods of Stage 3, Periods 16-25, we find that all of the restricted entry treatments are at least marginally significantly different from the Baseline with regard to both overall average contributions and group minimum contributions.¹⁵

We also note that the ordering of the significance levels for Stage 3 generally concord with what we observed earlier when considering the aggregate measures of behavior. The strongest significance levels obtain for the Combined treatment, while the weakest obtain for the Quiz Based Quota, which is sometimes only marginally significantly different from the Baseline.

For a more detailed regression analysis of treatment effects on contributions and earn-

¹⁵There are no statistically significant differences according to this test between the restricted entry treatments. We omit the test results to conserve space.

ings, we employ an estimation strategy that allows us to examine the treatment effects in Stage 3 conditional on behavior in Stage 2. We employ a standard differences-in-differences structure for panel data, in which we specify binary variables for the different treatment effects activated throughout the sessions, *Quota*, *Quiz*, *Comb* and *QBQ*; a binary variable corresponding to Stage 3, *S3*; and then interactions of those treatment effects with the Stage 3 dummy. These interaction terms measure any difference between behavior after the introduction of the group transition mechanism treatment has occurred (in Stage 3) but conditioned on the behavior in that population prior to the treatment being implemented (i.e., controlling for behavior in Stage 2). For the full specification see equation 2, which also includes variables to identify group size (GS_i) and period effects. In order to accommodate any subject- or group-level heterogeneity we use a random-effects panel structure with the random effect, c_i , on the individual or group level depending on the regression. The regressions include robust standard errors clustered at the subject or group level depending on the unit of observation we are using. Since our regressions are conducted separately for those in High groups and Low groups we begin this analysis after those group identities have been established. Consequently these regressions do not include data prior to Period 4.

$$\begin{aligned}
y_{it} = & \alpha_0 + \alpha_1 GS_i + \alpha_2 Period + \alpha_3 Quota + \alpha_4 Quiz + \alpha_5 Comb + \alpha_6 QBQ + \\
& \beta_0(LG*S3) + \beta_3(LG*S3*Quota) + \beta_4(LG*S3*Quiz) + \beta_5(LG*S3*Comb) + \beta_6(LG*S3*QBQ) + \\
& \gamma_0(HG * S3) + \gamma_1(HG * S3 * GS_i) + \gamma_2(HG * S3 * Period) + \\
& \gamma_3(HG*S3*Quota) + \gamma_4(HG*S3*Quiz) + \gamma_5(HG*S3*Comb) + \gamma_6(HG*S3*QBQ) + c_i + \varepsilon_{it}
\end{aligned} \tag{2}$$

Our first regressions look at individual contributions and earnings using data from each individual in a period as the unit of observation. To allow us to cleanly separate the behavioral treatment effects among those subjects who were initially in the High group from those who were initially in the Low group we run separate regressions for both types of subjects. For the subjects initially in the Low groups they have two sets of Stage 3 treatment variables. One set measures the effect of being in a particular treatment on behavior while still in the Low group and the other set measures the behavior while in the High group (if a move has occurred). Since those initially in the High group are restricted to be only in that group, they have only one set of Stage 3 interactions. These regressions can be found in the first four columns of Table 3.

| | Individual Contrib | | Individual Earnings | | Group Min |
|----------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | Initially in Low | Initially in High | Initially in Low | Initially in High | High |
| GroupSize, α_1 | -0.139*** (0.048) | -0.100** (0.050) | 0.730*** (0.202) | 0.512 (0.370) | -0.240*** (0.073) |
| Period, α_2 | -0.068*** (0.025) | -0.098*** (0.017) | 0.543*** (0.071) | 0.296*** (0.104) | 0.003 (0.021) |
| Quota, α_3 | 0.077 (0.151) | -0.048 (0.465) | 0.691 (0.913) | 0.431 (0.971) | >-0.001 (0.741) |
| Quiz, α_4 | -0.423*** (0.147) | -0.116 (0.513) | -1.296 (0.841) | -1.148 (1.227) | -0.245 (0.835) |
| Combined, α_5 | 0.224 (0.157) | -0.786 (0.560) | 4.779*** (0.877) | -1.128 (1.051) | -0.825 (0.887) |
| QuizBQuota, α_6 | 0.138 (0.177) | 0.176 (0.471) | -0.218 (1.163) | -0.686 (1.367) | 0.0571 (0.771) |
| Low Group * Stage 3, β_0 | 0.573 (0.495) | | 0.131 (2.055) | | |
| LGS3 * Quota, β_3 | -0.477 (0.464) | | -0.783 (1.952) | | |
| LGS3 * Quiz, β_4 | 0.0434 (0.517) | | -2.307 (2.057) | | |
| LGS3 * Combined, β_5 | -0.533 (0.469) | | -3.066 (1.923) | | |
| LGS3 * QuizBQuota, β_6 | -0.172 (0.645) | | 2.518 (2.215) | | |
| High Group * Stage 3, γ_0 | 0.715 (0.582) | -2.353*** (0.545) | 2.773 (3.347) | -13.79*** (2.895) | -3.196*** (0.755) |
| HGS3 * GroupSize, γ_1 | 0.082 (0.066) | | 0.175 (0.289) | | |
| HGS3 * Period, γ_2 | -0.061** (0.026) | | -0.502*** (0.095) | | |
| HGS3 * Quota, γ_3 | 3.034*** (0.488) | 2.844*** (0.531) | 12.500*** (2.623) | 13.17*** (3.263) | 2.941*** (0.769) |
| HGS3 * Quiz, γ_4 | 2.612*** (0.542) | 2.469*** (0.664) | 14.97*** (2.656) | 12.81*** (3.571) | 2.756*** (0.968) |
| HGS3 * Combin, γ_5 | 2.825*** (0.499) | 3.023*** (0.550) | 14.48*** (2.498) | 16.60*** (2.888) | 3.412*** (0.810) |
| HGS3 * QuizBQuota, γ_6 | 1.175** (0.596) | 0.981 (0.665) | 4.196 (2.582) | 7.147* (3.677) | 1.187 (1.032) |
| Constant, α_0 | 3.117*** (0.462) | 6.696*** (0.401) | 12.000*** (1.979) | 25.900*** (1.005) | 6.033*** (0.637) |
| Obs (Clusters) | 5,808 (264) | 1,452 (66) | 5,808 (264) | 1,452 (66) | 726 (33) |

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table 3: Regressions of Individual contributions and earnings as well as Minimum Contribution by group.

Across all regressions, the key variables for this study are the Stage 3 treatment variables for subjects in the High groups, γ_{3-6} . The fundamental question motivating this study is the degree to which any of the treatment mechanisms facilitate efficient coordination (higher contributions and earnings) in the High group once entry from the Low group is allowed in Stage 3. Thus, for example, we are interested in how the treatment mechanisms influence both whether subjects originally in the Low group make higher contributions once in the High group in Stage 3 (first column of results) and whether subjects originally in the High group make high contributions in Stage 3 (second column).

For the subjects initially in the Low groups we observe that for the Baseline treatment there is no significant change in their contributions as they move into the High group in Stage 3, as indicated by the lack of significance of the coefficient for the High Group * Stage 3 dummy, γ_0 . On the other hand, the High Group * Stage 3 coefficients for the restricted entry treatments, $\gamma_3 - \gamma_6$, are all positive and significant, though again we notice that the result is weakest in magnitude and significance for the Quiz Based Quota treatment.

For the subjects originally in the High groups, the results are slightly different. First, the Stage 3 dummy, γ_0 , is negative and significant indicating that, in the Baseline treatment, those in the High group contribute on average 2.35 less in Stage 3 than Stage 2, holding all other aspects equal. Further, the Stage 3 coefficients on the dummy variables for three of the restricted entry treatments — Quota, Quiz and Combined — are positive and significant, indicating that in those treatments the High group members at least do not decrease their contributions by as much in Stage 3 as their counterparts in the Baseline treatment. The coefficient for the Quiz Based Quota treatment is, however, not significant, indicating that this is not the case for that treatment. To examine the overall effect on Stage 3 contributions for each treatment we need to look at the linear combination of the Stage 3 binary variable (which measures the Stage 3 effect for the omitted treatment, Baseline) with the relevant interaction, $\gamma_0 + \gamma_{3,4,5 \text{ or } 6}$. These come out to be 0.491, 0.115, 0.670 and -1.37 for the Quota, Quiz, Combined and Quiz Based Quota treatments, respectively. The p -values on tests of whether these linear combinations are equal to 0 are 0.079, 0.800, 0.008 and 0.004. The interpretation of these tests is that for the Quota, Quiz and Combined treatments, the original members of the High groups do not exhibit a fundamental decrease in contributions between Stages 2 and 3 (and exhibit a very minor increase in Quota and a slightly larger increase in Combined) while they exhibit a statistically significant decrease in the Quiz Based Quota treatment. Hence, supporting our earlier observation, the Quiz Based Quota treatment appears less effective than the other three mechanisms. We also again find some evidence that the Combined mechanism does slightly better.

The regressions in the third and fourth columns of Table 3 show similar results for subjects' earnings. By examining the same coefficients again in the third column, we see that earnings for those that began in the Low Group increase substantially when they move to the High group in the first three restricted entry treatments compared to the Baseline, but there is no such relative increase for the Quiz Based Quota treatment.

For those that begin in the High groups, studied in Column 4, their earnings fall substantially in the Baseline treatment during Stage 3; but again this decrease is largely reversed in the Quota, Quiz and Combined treatment but only partially reversed in the Quiz Based Quota. By again testing linear combinations of the Baseline effect with the interactions, we see that earnings do not fall significantly during Stage 3 in the Quota and Quiz treatments, they increase slightly in the Combined treatment, and they decline in the Quiz Based Quota treatment.

Further evidence regarding the effectiveness of the treatments can be found in the last column of Table 3, which presents a similar regression using the minimum contribution in a group in a period as the unit of observation. We only conduct this regression for the High groups, since this is our focus. We see that the group minimum drops significantly during Stage 3 in the Baseline condition, measured by γ_0 in the last column, while minima in the first three treatments are significantly above the Baseline ($\gamma_{3,4}$ and 5); yet again, that is not the case for the Quiz Based Quota treatment. Tests on the linear combinations to identify the total effect on the group minima for the treatments again show that there is no statistically significant difference between Stage 2 and Stage 3 for the Quota, Quiz and Combined treatments, but for the Baseline and Quiz Based Quota the group minima are lower.¹⁶

The results above demonstrate that, in general, the performance of the Quiz Based Quota mechanism is better than the Baseline mechanism, but only slightly and often statistically insignificantly, while all other mechanisms perform significantly better than the Baseline. This leads to an immediate question regarding whether Quiz Based Quota is significantly different from the other treatments with entry restrictions. One could use linear combination tests using the results in Table 3 to test this but instead we provide a more straightforward approach to this where we replicate all the regressions from Table 3, but exclude the Baseline data and use the Quiz Based Quota treatment as the default case.

¹⁶Figure 1 suggests possible time trends in the group minima toward the end of the experiment, which make one question whether these time trends differed among the treatments. To test this, we conducted a regression similar to model 5 in Table 3 but restricting attention to only the last 10 periods and interacting treatment with periods. We find a few very minor differences in the time trends, but none of these are of substantial magnitude and none is significant at the $p < 0.05$ level.

We provide the full results of this analysis in the Appendix. This analysis reveals that the relevant coefficients for the restricted entry treatments tend to be large, positive and generally statistically significant, indicating that, indeed, the Quiz Based Quota performs significantly worse than the other three restricted entry mechanisms.

The finding that the Quiz Based Quota treatment performs less well than the other treatments is important for at least two reasons. First, it shows that it is not simply that *any* mechanism that restricts entry works equally well. The Quiz Based Quota implements slower growth than in the Baseline, and works slightly better than that treatment. But, it does not work as well as the more restrictive Quota mechanism, which regulates entry more strictly and obtains coordination on higher contributions and minima. Hence, characteristics of the mechanisms themselves — such as how quickly they grow the group — are important. The fact that the slower Quota treatment works better than the faster Quiz Based Quota tells us that the rate of growth matters. Second, the difference between the Quiz Based Quota and Quiz treatments helps us understand why the Quiz treatment is effective. It is not just that it slows down growth relative to the Baseline — the Quiz Based Quota treatment also does that to the same extent. Instead, the screening mechanism of the quiz itself seems important for its effectiveness.¹⁷ Hence, the limited effectiveness of the Quiz Based Quota, relative to the Quota and Quiz treatments, tells us that either strict restrictions on entry or a screening mechanism that requires entrants to demonstrate some knowledge can work, but they work for different reasons.¹⁸

3.3 Responses to Group Transitions

The results above show that in all treatments there is at least some decline in the level of coordination between Stages 2 and 3. Under some mechanisms this decline is negligible while in others it is substantial. In this section we examine how contribution levels differentially adjust to group transitions between the treatments, and how contributions by High group

¹⁷We can also test whether the Quiz performs a screening function by comparing the average contributions, from Stage 1, of those who pass the quiz in Period 11 with those who attempt the quiz but do not pass in that period and those who do not even attempt the quiz. Indeed, we find that those who pass in Period 11 tend to be higher contributors from Stage 1 than the other two groups in both the Quiz (2.00 vs 1.63 and 1.84) and Combined (2.42 vs. 2.23 and 2.21) treatments. While these differences are not significant, perhaps due to limited power, they are consistent with the Quiz playing a screening function.

¹⁸To a lesser extent, we can also observe this finding in the fact that the Combined treatment works slightly better than either the Quiz or the Quota mechanism, though the difference is not large enough for us to find it statistically significant. The indication is that adding one mechanism to the other has a (slight) beneficial effect, supporting the claim that they work for independent reasons. Given the lack of strong significance of the superiority of the Combined mechanism, we do not want to overstate this claim but note it as a possibility.

incumbents and new entrants differ in early rounds of group transitions, in order to better understand why coordination failed or succeeded in the different treatments and complement our earlier interpretations regarding why the Quiz and Quota mechanisms succeeded in maintaining coordination.

When thinking about the potential causes for coordination breakdown, one might initially think to blame the Low group members who come into a well coordinated High group. The new entrants might not contribute at the same level as the High group incumbents and thereby lead to a breakdown in coordination. We noted before that the quiz might help to alleviate this problem by only selecting individuals to transition between groups who will contribute at that higher level, leading to its success in maintaining coordination. Similarly, the strict quota may reassure entrants by eliminating concerns that other contemporaneous entrants may make low contributions, causing such entrants to be more willing to match the contributions of the High group. On the other hand, it is also possible that any coordination failure is due to the High group members (perhaps inappropriately) mistrusting the Low group members coming in and causing the decline themselves.¹⁹ We discussed before how the quota mechanism might assist with this problem — by limiting the number of new entrants to a small number, it may make incumbents more trusting of that single entrant than they would be of multiple random entrants. Similarly, the quiz may reassure incumbents that those entering understand the benefits of high contributions.

What we seek to do here is directly examine the degree to which success or failure is determined by the behavior of entrants or incumbents, to provide further evidence on why the mechanisms that succeed are effective. Tables 4 and 5 provide some detailed observations on group transition effects to help us address this question.

Table 4 shows summary statistics (average contributions) before and after movement between groups is allowed. This table focuses on the original High group members and on those who began in a Low group but moved into the corresponding High group in Period 11. The top panel presents mean contributions for Periods 10 and 11, i.e. the periods right before and right after group transitions are allowed. The bottom panel compares the contributions for these individuals for Periods 8-10 and 11-13, or three periods before and after transitions are allowed. While the contributions of both groups are not comparable in Periods 10 or 8-10, their contributions after group movement is allowed do inform us regarding what might be contributing to coordination failure or success.

For those originally in the Low groups who move in Period 11 the average contribution

¹⁹This would be consistent with McCarter and Sheremeta (2013), who find that it is distrust of new members by existing members of their fixed size groups that leads to coordination failure.

| | Original Low Group Members | | Original High Group Members | |
|------------------|----------------------------|---------------|-----------------------------|---------------|
| | Period 10 | Period 11 | Period 10 | Period 11 |
| Baseline | 1.19 | 4.71 | 5.67 | 5.17 |
| Quota | 1.33 | 6.17 | 5.67 | 5.58 |
| Quiz | 0.50 | 5.30 | 6.07 | 5.71 |
| Combined | 1.57 | 5.29 | 5.28 | 4.89 |
| Quiz Based Quota | 1.57 | 5.21 | 6.50 | 5.0 |
| | Periods 8-10 | Periods 11-13 | Periods 8-10 | Periods 11-13 |
| | Baseline | 1.30 | 3.30 | 5.64 |
| Quota | 1.40 | 5.64 | 5.67 | 5.53 |
| Quiz | 1.01 | 4.69 | 6.07 | 5.26 |
| Combined | 1.89 | 4.65 | 5.11 | 4.72 |
| Quiz Based Quota | 1.46 | 3.98 | 6.30 | 4.10 |

Table 4: Summary statistics of contributions before and after group transitions for original High group members and those from the Low groups who moved in period 11.

prior to moving is around 1 in all treatments. After a group transition, it increases in Period 11 in all treatments, but least so for the Baseline and Quiz Based Quota treatments. Hence, we have an initial piece of evidence that part of the effect of the successful mechanisms is on the immediate behavior of entrants following a move. For those originally in the High groups the table shows that, in all treatments, the onset of movement between groups leads to an immediate decrease in average contributions, even in the first period, and this difference is again largest for the Baseline and Quiz Based Quota treatments. Thus, simply knowing that entrants are moving into the High group leads some of the original High group members to slightly lower their contributions. As we hypothesized, this appears to be at least slightly larger when larger numbers of unscreened entrants enter. It is also useful to compare the High group contributions in Periods 10 and 11 to the Period 11 contributions of the Low group members joining them. What we see is that in 3 out of the 5 treatments, the entrants from the Low group actually contribute more on average than the High group incumbents in Period 11 and in 2 out of the 5 treatments they contribute above the High group level from Period 10 in Stage 2.

If we examine what happens over the 3-period windows, we see that for both High group incumbents and entrants from the Low groups, contributions generally stabilize in the three restricted entry treatments that are successful (Quota, Quiz and Combined). However, they drop more precipitously in the Quiz Based Quota and especially in the Baseline.

Table 5 reports regressions intended to allow for a better understanding for how responses to group transitions vary by treatments. The regressions use as the dependent

variable the average of the contributions a subject originally in the Low group made for the one or three periods after executing a group transition (Original LG members) or a subject originally in the High group made after group transitions are allowed (Original HG members). For this table we consider all subjects who began in a Low group and moved to a High group, not just those who did so in Period 11 as in the prior table. The independent variables are the subject’s own contributions in the previous one or three periods,²⁰ the minimum in the High Group from the prior period, as well as the treatment binary variables. For the subjects in the High groups we include an additional control, which is the minimum of contributions they experienced in the 10-person group in Stage 1.²¹ Note that given the fact that Low and High group members are responding to different events and beginning from different baselines, it is not meaningful to compare coefficient values between the Low group and High group specifications.

For those moving from Low groups into High groups (first two regressions), we see that the immediate (one-period) response is that subjects in the three successful conditions increase their contributions by more than in the Baseline, while the difference is not significant for the Quiz Based Quota treatment. The differences all become stronger for the 3-period regression. For those in the High Group, there appears to be a general decline from Period 10 contribution levels in Period 11, based on the fact that the coefficient on the prior period minimum is less than 1.²² There are however no significant differences across treatments. If we look at the 3-period regressions, there are now differences across treatments with the 3 successful mechanisms yielding higher contributions than Baseline and Quiz Based Quota treatments. This difference between 1- and 3-period results indicate that responses by the incumbents to the treatments did not occur prospectively, but rather primarily only after interacting with the new members. This suggests the success or failure of the mechanisms in ensuring efficient coordination was primarily driven by its effect on the behavior of the

²⁰Note that since the High groups are well coordinated at the end of Stage 2, the group minimum from $t - 1$ and High group members own individual contributions are perfectly correlated. Thus the individual’s own contribution from one period prior is dropped from the regression for the High group members. While the correlation is not quite as strong for the 3-period regression, the colinearity is still strong enough to affect the results and so we drop it from that regression as well.

²¹Though it turns out to be insignificant, the variable could have been important for the behavior of the members of the High groups because the observation of what happened at the beginning of the experiment may have helped to form their expectations of how new members from the Low group will behave.

²²Determining if the overall drop in contributions between Periods 10 and 11 is significant is difficult from the regression results alone. To test this directly we conducted both Wilcoxon signed-rank tests and paired t-tests and found equivalent results. There is insufficient data to make strong claims about the significance of individual treatments, but if we pool the data for all High Group incumbents, contributions are significantly different between the two periods, with Period 11 being lower, $p < 0.001$. In general, we find that the incumbents either kept their contributions constant (44/66) or dropped them (20/66), rather than increasing them (2/66).

| | Contributions After Move / Stage 3 Begins | | | |
|--------------------------|---|-----------|---------------------|-----------|
| | Original LG Members | | Original HG Members | |
| | 1 Period | 3 Periods | 1 Period | 3 Periods |
| Group Minimum from | 0.654*** | 0.616*** | 0.962*** | 0.808*** |
| High Group $t - 1$ | (0.054) | (0.039) | (0.082) | (0.068) |
| Avg of Own Contributions | 0.108* | 0.210*** | - | - |
| x Periods Before Move | (0.064) | (0.0501) | | |
| Stage 1 Minimum | - | - | -0.069 | -0.113 |
| | | | (0.155) | (0.116) |
| Quota | 1.247*** | 2.036*** | 0.440 | 1.760*** |
| | (0.310) | (0.230) | (0.369) | (0.410) |
| Quiz | 1.122*** | 1.646*** | 0.192 | 1.131** |
| | (0.335) | (0.277) | (0.505) | (0.488) |
| Combined | 0.867*** | 1.657*** | 0.161 | 1.294*** |
| | (0.290) | (0.227) | (0.393) | (0.420) |
| Quiz Based Quota | 0.667 | 1.010*** | -0.824 | -0.219 |
| | (0.410) | (0.281) | (0.556) | (0.452) |
| Constant | 0.581 | -0.427* | -0.251 | -0.714 |
| | (0.358) | (0.254) | (0.486) | (0.448) |
| Observations | 235 | 234 | 66 | 66 |
| R^2 | 0.487 | 0.637 | 0.711 | 0.713 |

Robust standard errors in parentheses, *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 5: Regressions on behavior before and after group transitions.

new members.

Taken together, these results allow us the basis for some speculative comments about what may have been important to the success/failure of groups maintaining a high level of coordination in Stage 3. In all treatments, there is some indication of initial distrust among the High group incumbents toward the new entrants. This distrust is perhaps justified in the Baseline treatment, as indicated by the new entrants in that treatment not increasing their contributions by as much in the others. This led to quick coordination failure in the Baseline treatment, consistent with Figure 1. In the Quiz Based Quota treatment, coordination also failed, but there it failed despite the new entrants on average contributing above the average of the incumbents in Period 11. This was the treatment, though, in which the incumbents indicated the least trust in new entrants and the new entrants appeared to have not sustained their contributions above the original High group members for those first three periods. In the other three treatments, the sustained good contributions of the new entrants seems to have been sufficient to foster mutual trust between both incumbent and new members, allowing for sustained coordination. These results provide additional support for our previous interpretations regarding why the quiz and quota mechanisms were successful. New entrants are generally willing to increase their contributions to match the incumbents but less so when they enter in large or randomly selected groups. Incumbent members also seem to be slightly more trusting of those who enter in either small or specially selected groups. Thus, at the level of individual responses to group transitions under the different mechanisms, we again find support for the effectiveness of the three mechanisms with the strongest entry restrictions.

4 Conclusion

This study explores the effectiveness of different mechanisms for allowing small, efficiently-coordinated groups to grow successfully by incorporating new group members from lower-performing groups. We investigated simple mechanisms, similar to ones commonly used outside of the laboratory to restrict entry. One was a simple Quota mechanism limiting the number of entrants in a period. The second was a Quiz through which entrants had to demonstrate a minimum level of knowledge of the growing group's history and of the incentive structure in our production environment. We also included two additional treatments that incorporated features of these mechanisms. Our Combined treatment imposed both restrictions while our Quiz Based Quota treatment imposed a Quota determined by the average number of people who successfully moved between groups in the Quiz treatment.

The results of the experiment show that all of the mechanisms for restricting entry have some degree of effectiveness for facilitating persistent coordination on high contributions while incorporating new entrants. The Baseline treatment, with no entry restrictions, rapidly collapses to very low contribution levels. However, the Quota and Quiz mechanisms both perform roughly equally well in allowing successful growth without a substantial breakdown in efficient coordination. Hence, either a mechanism that restricts entry to a limited number of entrants or one that screens entrants can be effective.

Our other two treatments help address why these two mechanisms might have worked. First, the Combined treatment performs slightly (though statistically insignificantly) better than the Quiz and Quota treatments. This is an initial indication, though hardly a conclusive one, that the two mechanisms may do something different. More compelling evidence of how the two mechanisms work is evident from the Quiz Based Quota treatment. This allowed groups to grow at the same rate as in the Quiz treatment but without the screening component. The results clearly demonstrate that while these groups might not have failed quite so badly as the Baseline of free entry, they performed worse than the Quiz and Quota treatments. These comparisons tell us something about why the other two mechanisms worked. First, by comparing the Quota treatment, which only allows one entrant per period, to the Quiz Based Quota, which also limits entry but allows more rapid expansion, we see that the former outperforms the latter. Hence, part of the reason the Quota works well is because of the number of entrants it allows — a less restrictive quota works less well. Second, the Quiz Based Quota implements the same growth trajectory as the Quiz treatment, but works much better. Hence, the Quiz works not just by slowing down entry, but also by providing a screening function that only permits entry by those able to demonstrate relevant knowledge.

Our results also greatly expand our knowledge of how small groups can grow successfully, even when entrants have a strong history of coordination failure and inefficiency. Importantly, in contrast with Weber (2006), we also show that groups can grow *quickly* while maintaining efficient coordination. While the size of High Groups in the Quiz treatment lagged that of the High Groups in the Baseline by a period or so, their growth rates were otherwise quite similar. Yet the High groups in the Quiz treatment substantially outperformed those in the Baseline. And there was only a negligible difference between the success of the High groups in the Quiz and the Quota treatment, even though the High groups in the latter treatment grew at a slower rate. This demonstrates that the selection properties of the entry exam can be as effective as the quantity limit imposed by the Quota.

The above results indicate a substitutability between limited entry without information

and information without limited entry. When only one person enters the efficient group at a time, everyone is sufficiently secure that the efficient coordination will not be harmed even if the incumbents do not know much about that person. Similarly, when multiple people enter but those people have demonstrated relevant knowledge, everyone is similarly sufficiently confident that growth will not harm the efficient coordination. Hence, what seems important is that there is *some* mechanism to ensure that everyone is sufficiently confident that entry will not overwhelm efficiency, allowing growth to be successful. As we note earlier, the ability to screen potential entrants based on past behavior can provide another such mechanism (Ahn, Isaac and Salmon, 2008 & 2009; Riedl, Rohde and Strobel, 2012). Our work contributes to this important issue by demonstrating the importance of alternative mechanisms.

References

- Ahn, T., R. M. Isaac, and T. C. Salmon (2008). Endogenous Group Formation. *Journal of Public Economic Theory* 10(1), 171–194.
- Ahn, T., R. M. Isaac, and T. C. Salmon (2009). Coming and Going: Experiments on Endogenous Group Sizes for Excludable Public Goods. *Journal of Public Economics* 93(1-2), 336–351.
- Andvig, J. C. and K. O. Moene (1990, January). How corruption may corrupt. *Journal of Economic Behavior & Organization* 13(1), 63–76.
- Borjas, G. J. (1994). The economics of immigration. *Journal of Economic Literature* 32(4), 1667–1717.
- Brandts, J. and D. Cooper (2006). A Change Would Do You Good An Experimental Study on How to Overcome Coordination Failure in Organizations. *The American Economic Review* 96, 669–693.
- Brandts, J., D. Cooper, and R. Weber (Forthcoming). Legitimacy, communication and leadership in the turnaround game. *Management Science* –, –.
- Cable, D. M. and T. A. Judge (1996). Person-organization fit, job choice decisions, and organizational entry. *Organizational Behavior and Human Decision Processes* 67(3), 294–311.

- Card, D. (2005). Is the new immigration really so bad. *Economic Journal* 115(507), 300–323.
- Chatman, J. A. and F. J. Flynn (2001). The influence of demographic composition on the emergence and consequences of cooperative norms in work teams. *Academy of Management Journal* 44(5), 956–974.
- Chen, R. and Y. Chen (2011). The potential of social identity for equilibrium selection. *American Economic Review* 101(6), 2562–2589.
- Cooper, R. and A. John (1988). Coordinating coordination failures in keynesian models. *The Quarterly Journal of Economics* 103(3), 441–463.
- Devetag, G. and A. Ortmann (2007). When and why? a critical survey on coordination failure in the laboratory. *Experimental Economics* 10(3), 331–344.
- Fischbacher, U. (2007). z-Tree: Zurich Toolbox For Readymade Economic Experiments. *Experimental Economics* 10(2), 171–178.
- Friedberg, R. M. (2001). The impact of mass migration on the israeli labor market. *The Quarterly Journal of Economics* 116(4), 1373–1408.
- Hirshleifer, J. (1983). From weakest-link to best-shot: The voluntary provision of public goods. *Public Choice* 41, 371–383.
- Kunreuther, H. and G. Heal (2003). Interdependent Security. *Journal of Risk and Uncertainty* 26, 231–249.
- McCarter, M. W. and R. M. Sheremeta (2013). You can't put old wine in new bottles: The effect of newcomers on coordination in groups. *PLoS One* 8(1), e55058.doi:10.1371/journal.pone.0055058.
- O' Reilly III, C. A., J. Chatman, and D. F. Caldwell (1991). People and organizational culture: A profile comparison approach to assessing person-organization fit. *The Academy of Management Journal* 34(3), 487–516.
- Ortega, F. (2005). Immigration quotas skill upgrading. *Journal of Public Economics* 89(9-10), 1841–1863.
- Riedl, A., I. M. T. Rohde, and M. Strobel (2011). Coordination in weakest-link games. IZA Discussion Paper 6223.

Rumbaut, R. G. (1997). Immigration and incorporation. *Sociological Perspectives* 40(3), 333–338.

Ryan, A. M. and P. R. Sackett (1987). Pre-employment honesty testing: Fakability, reactions of test takers and company image. *Journal of Business and Psychology* 1(3), 248–256.

Van Huyck, J. B., R. C. Battalio, and R. O. Beil (1990). Tacit coordination games, strategic uncertainty, and coordination failure. *The American Economic Review* 80(1), 234–248.

Van Huyck, J. B., R. C. Battalio, and R. O. Beil (1993). Asset markets as an equilibrium selection mechanism: Coordination failure, game form auctions, and tacit communication. *Games and Economic Behavior* 5(3), 485–504.

Warne, F. J. (1921). Assimilation and governmental regulation. *The Annals of the American Academy of Political and Social Science* 93, 181–185.

Weber, R. A. (2006). Manging growth to achieve efficient coordination in large groups. *The American Economic Review* 96(1), 114–126.

Appendix A: Comparison of Restricted Entry Treatments

| | Individual Contrib | | Individual Earnings | | Group Min |
|----------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | Initially in Low | Initially in High | Initially in Low | Initially in High | High |
| GroupSize, α_1 | -0.133*** (0.049) | -0.163*** (0.049) | 0.615*** (0.197) | 0.717* (0.419) | -0.264*** (0.083) |
| Period, α_2 | -0.071** (0.031) | -0.059*** (0.019) | 0.520*** (0.071) | 0.191 (0.141) | 0.020 (0.028) |
| Quota, α_3 | -0.061 (0.186) | -0.224 (0.384) | 0.908 (1.218) | 1.117 (1.278) | -0.057 (0.607) |
| Quiz, α_4 | -0.561*** (0.183) | -0.292 (0.441) | -1.079 (1.165) | -0.462 (1.482) | -0.302 (0.721) |
| Combined, α_5 | 0.086 (0.191) | -0.962* (0.496) | 4.997*** (1.191) | -0.442 (1.340) | -0.883 (0.780) |
| Low Group * Stage 3, β_0 | 0.448 (0.417) | | 2.253* (1.202) | | |
| LGS3 * Quota, β_3 | -0.319 (0.483) | | -3.108** (1.236) | | |
| LGS3 * Quiz, β_4 | 0.227 (0.543) | | -4.801*** (1.433) | | |
| LGS3 * Combined, β_5 | -0.369 (0.511) | | -5.330*** (1.282) | | |
| High Group * Stage 3, γ_0 | 1.910*** (0.558) | -1.398*** (0.485) | 6.503*** (2.326) | -6.831** (3.463) | -2.041** (0.856) |
| HGS3 * GroupSize, γ_1 | 0.030 (0.070) | | 0.494 (0.320) | | |
| HGS3 * Period, γ_2 | -0.0380 (0.0324) | | -0.587*** (0.110) | | |
| HGS3 * Quota, γ_3 | 1.863*** (0.465) | 1.835*** (0.519) | 8.167*** (1.766) | 6.121* (3.685) | 1.743* (0.925) |
| HGS3 * Quiz, γ_4 | 1.426*** (0.526) | 1.479** (0.646) | 10.75*** (1.861) | 5.691 (3.964) | 1.565 (1.094) |
| HGS3 * Combin, γ_5 | 1.597*** (0.490) | 1.935*** (0.540) | 10.31*** (1.677) | 9.803*** (3.415) | 2.184** (0.956) |
| Constant, α_0 | 3.229*** (0.514) | 6.722*** (0.334) | 12.87*** (2.036) | 25.54*** (1.456) | 6.019*** (0.506) |
| Obs (Clusters) | 4,752 (216) | 1,188 (54) | 4,752 (216) | 1,188 (54) | 594 (27) |

Robust standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1