Contests as a Tool for Development:

Evidence from Peru[1]

A Thesis

submitted by

Mikal S. Davis

In partial fulfillment of the requirement for the degree of

Master of Science

in

Economics

TUFTS UNIVERSITY

May, 2013

Advisor: Dr. Kelsey Jack

[1] A special thanks to Pachamama Raymi and Innovations for Poverty Action, for hosting me and being so helpful during the data collection and interview process. Also a special thanks to my secondary advisor, Dr. Margaret McConnell, for providing the funding that allowed me to travel to Peru and also helped formulate ideas for this research. Also many thanks to my friends and family who have supported me throughout this process.

Abstract

To investigate the effectiveness of contests in a development setting, I use a case study from Peru where contests are implemented by giving households numeric scores based on how well they have adopted improved agricultural, sanitation and agribusiness technologies. This paper analyzes the incentive effects of these contests by testing three hypotheses generated by theory: (1) Increasing the number of contestants will decrease average performance; (2) Performance is increasing in the prize spread; (3) High ability contestants will adopt low-risk strategies while low ability contestants will adopt high risk strategies. I find evidence consistent with theory for hypotheses (2) and (3), but the results for hypothesis (1) are ambiguous in regard to theoretical predictions. Together, the results suggest that contests influence behavior in predictable ways and therefore have the potential to be an additional policy option for development organizations.

Contents

1 Introduction

Development is not easy. If it were, we would not see the degree of poverty that currently persists throughout the world. The challenge is one shared by development organizations everywhere, and at the heart of the problem lays the basic question of how to incentivize behavior change. For poor people, changing behaviors by adopting new practices comes at a high cost; they have limited resources and even the smallest investment can be very risky. Often, programs are designed to help mitigate this problem by correcting market failures and providing missing incentives, which intends to spur new behaviors that promote long-term economic growth. These programs, like conditional cash transfers (CCT's), financial donations and savings groups have all seen varying degrees of success throughout the developing world, but can still be made obsolete by the negative effects of economic shocks. Contests provide a unique solution to this problem and therefore have the potential to be a good way to design incentives in this setting; they offer the advantage of reducing the effects of negative shocks common to all contestants by using relative performance (as opposed to absolute performance) to determine compensation. This means that a contestant's compensation is no longer adversely affected by draughts, plagues or other agricultural shocks, which can be painfully frequent for rural farmers. Although organizations like the one I study in Peru have begun to use contests to incentivize behavior change, the extent to which contests can be a useful tool in rural development has yet to be seen.[2]

The contests I study take place in rural communities in southern Peru and

[2] Although the use of contests has not been studied specifically, competition as a tool for development has begun to garner attention in the development literature. Linardi et al. (2011) look at the role of competition as an incentive device for savings groups. They find that savings levels increase in the short-run due to the competition, but converge back to normal levels in the long run. This data does not offer the advantage of long-run analysis, but suggests it as a promising avenue for further research.

incentivize households to compete with each other in adopting new agricultural, sanitation and agribusiness behaviors by rewarding contest winners with cash prizes. Households volunteer to participate and are given six months to complete a list of pre-determined improvements to their homes, agricultural practices and family hygiene.[3] At the end of the contest, each household is given a numeric score grading their success in each of the categories and the top households in each community receive cash prizes. The contests are administered by a local non-profit organization, Pachamama Raymi (PMR), and are coupled with educational workshops designed to assist the adoption of improved practices. PMR also attempts to incentivize teamwork within the same communities by creating inter-community contests which are based on the aggregate number of points a community earns. The contests take place in 3 regions in the provinces south of Lima and each region's communities participate in two sequential contests at both the household and community levels.

The question at the heart of the analysis is whether or not the behavior of PMR's contestants is consistent with theoretical predictions. If it is, the result would suggest that contests can offer another policy option for development organizations by providing a framework that produces predictable behaviors. On the other hand, if contestants are not behaving as theory predicts, it is possible that contests are still effective, but provide incentives through mechanisms different than those predicted by existing theory.[4] This would indicate that further research is needed to develop a body of theoretical literature for contests in more complex social settings.

To test this question, I use three of the most commonly tested hypothe-

[3]For a complete list of the pre-determined activities, see appendix.

[4]It is also worthy to note that finding behaviors that contradict theory might indicate that the participants do not fully understand the rules or prize structure.

ses generated by the existing theoretical literature (Lazear and Rosen (1981), Nalebuff and Steiglitz (1983), Gibbs(1996)) in order to either verify or refute the power of contests in this setting. The hypotheses tested are:

1. Increasing number of households from n to $n+1$ while holding the number of prizes fixed will decrease performance

2. Performance is increasing in the prize spread

3. High ability households will adopt low-risk strategies while low ability households with adopt high-risk strategies

These three hypotheses have been studied extensively, but their ability to be tested empirically is limited to strict assumptions of unobservables imposed by theory. The limitations have encouraged the development of a large body of lab experiments (List et al. (2010) Harbring and Irlenbusch (2002), Hvide (2002)), but due to their design, these experiments are generally uninformative about contests in a development setting. Another branch of the literature attempts to connect theoretical predictions with real-world applications by using institutional data, but is severely constrained by data availability and has consequently been restricted to using data from sporting events or large corporations (Knoeber and Thurman (1994) Casas-Acre and Martinez-Jerez (2009) Boudreau et al (2012) Becker and Hustled (1992)). This paper is the one of first empirical studies of contests in a development setting and therefore attempts to widen the scope of the non-experimental evidence by extending the analysis of contests to rural development.

After testing the three hypotheses in this setting, I find that in general, households do behave predictably in contests, especially after the first contest. This is a logical result seeing as the first contest serves as a chance for households to learn the rules and confirm that PMR will fulfill its promise of rewarding prizes.

The results from the first hypothesis help to clarify the mixed results found in the literature.[5] I find that there is a positive relationship between contest size and performance levels during the second contest while controlling for prize amount. One explanation for this could be that households tend to work as a team and more teammates leads to higher productivity. This is the most convincing explanation considering the additional incentive to co-operate generated by the extra layer of inter-community contests. Another explanation could be that households misunderstood the rules and expected a prize increase with every additional household, when in reality, the prize only increased after every additional 15 households. Observing behavior over more consecutive contests as well as analyzing the incentive effects of the inter-community contests would help clarify this conclusion, but was not available using this data.

The results from the second hypothesis conform to theoretical predictions: increasing the prize spread by 50 Peruvian Soles (∼$19) leads to almost a full standard deviation increase in performance scores, on average. This has interesting implications for contests used as a tool for development in that there is a definite mechanism through which effort can be manipulated. Thus, the result represents a potentially powerful way to incentivize effort in rural communities through contest design.

The third hypothesis also conforms to theoretical predictions. I find a slightly negative relationship between contestant quality and variability of performance (risk) in each region, indicating that high and low ability households adopt strategies as predicted by theory. I also test the presence of a sort-

[5] See Dechenuax et al. (2012) in section 3 for a detailed discussion of the ambiguous results the literature has produced so far.

ing effect; theory holds that in order to reduce the inefficiencies created by heterogeneity, sorting households into groups according to their ability can increase overall contest effort levels.[6] After sorting the winners from the first round into a separate league, the relationship between quality and variability of performance changes slightly, but is very small in magnitude. By design, the contests not only experience a sorting-out effect from separating the winners, but also a sorting-in effect from new households entering after the first contest. After taking the additional sorting effect into account, I find that the sorting-in effect dominates during the second contest. This provides evidence that the composition of the sorted-in group of households must be considered when trying to level the playing field, meaning larger groups of winners should be sorted out or separate leagues need to be formed through a different mechanism.

As a whole, the results offer evidence that contests can be an effective tool for motivating rural community members to adopt new technologies. This implies that PMR's contests are increasing effort levels and concentrating them on specific activities meant to help the households break the cycle of poverty and promote economic growth. Some caveats to the results are the lack of controls, which could cause omitted variable bias, the lack of variation within key variables in the dataset, which limits the analysis to narrowly defined tests and the lack of data from communities that participate in more than two contests. These potential hindrances are controlled for as much as the data permits.

The remainder of the paper is structured as follows: section 2 provides the

[6]A number of papers have addressed this sorting effect. The basic theory laid out by Lazear and Rosen (1981) predicts that mixed contests suffer from a general under-investment of effort due to the heterogeneity of ability. This reallocation of effort reduces the efficiency of contests compared to piece-rates or symmetric contests. Lazear and Rosen predict that this inefficiency can be overcome by sorting players according to their ability. This effectively creates separate, symmetric contests and therefore increases efficiency.

background of PMR and its practices in the field; section 3 discusses a simple model of multi-person contests and lays the theoretical foundations for the 3 hypotheses tested; section 4 reviews the relevant literature; section 5 describes the data; section 6 outlines the empirical approach for each hypothesis; section 7 summarizes the results; and section 8 discusses the results.

2 Background

Rural communities in the developing world pose several challenges to development organizations trying to improve the standard of living for poor households. These organizations often go to great lengths to find donors, design effective policies and implement them in the most sustainable way possible, but find that poverty persists despite their best efforts. The disincentive effect created by poverty is an obstacle for all organizations trying to create policies to alleviate poverty. Tradition and culture can also exaggerate this disincentive effect. When farming techniques have been used for generations, poor families are hesitant to switch to more productive techniques because of the uncertainty and therefore risk associated with the change.[7]

Thus, the challenge as a development organization is to design policies that can incentivize behavior change that will ultimately increase productivity or improve quality of life. Contests as a tool for development are a fairly new concept, but aim to side-step this incentive problem by using cash prizes to serve as the financial incentive for households to adopt more economically productive behaviors or make investments that can raise their standard of living in the long-run. Pachamama Raymi (PMR) is a non-profit organization that uses contests to do just that. It aims to promote local infrastructure investment in rural communities in Peru by offering the contest winners cash prizes large enough to make an economic difference in their lives. Their mission aims to use peer learning and contests as a mechanism to confront poverty from multiple angles; they incentivize families to invest in their households and the environment in order to create sustainable sources of income generation.

[7] See Chapter 8 of the 2000/2001 World Development Report for a more detailed discussion of poor people and risk. The authors suggest that poor people are inherently more vulnerable and therefore are less likely to engage in high-risk high-reward activities that could potentially get them out of the poverty trap.

Pachamama Raymi's Methodology

One of the biggest challenges of providing the necessary incentives is the fact that during the contests, PMR does not offer financial assistance to families to adopt these better practices. As a result, families are required to invest their own resources in improving their houses, agricultural techniques, etc. Given their high exposure to risk and short-run needs, poor families do not have much incentive to engage in activities that PMR considers to be necessary to improve their standard of living in the long-run. PMR is one of the first organizations to use contests as a tool to realign incentives in order to make these "better" practices a practical choice for poor families. PMR's methodology primarily consists of two mechanisms to incentivize rural community members to adopt better practices: Peer Learning and Contests.

Peer Learning represents about 39% of PMR's expenditures, and is a integral part of the PMR methodology. All of PMR's field staff come from rural communities and speak Quechua. This is meant to foster a socio-economic understanding and to also provide a means to communicate with community members in their first language. Many communities are very isolated and have few spanish-speaking members. They are also very close-knit and hesitate to trust foreigner-run NGO's, who have a history of not fulfilling commitments made to communities in the region.[8] Having facilitators propose the idea of a contest to community leaders is meant to help clarify the rules as well as the

[8] This is an important observation to make in regard to the analysis. Knowing that the region suffers from a general distrust of NGO's is necessary to take into account when observing effort levels across contests. Given this distrust, I would expect to see an increase in participation and effort across contests, ceterus paribus, due to the fact that after the first contest, skeptical community members have seen that the NGO fulfilled its promise in awarding prizes and will therefore participate and be willing to invest more of their resources to win the second contest. In a similar vein, apart from learning trust, there is also a certain amount of technical learning that takes place; people understand the rules better, have seen their neighbors participate, etc. This type of learning also needs to be taken into account.

long-term objectives of PMR, which helps the contests run more effectively. The facilitators are also experts in various agricultural practices and organize educational workshops that are meant to bring together farmers to share ideas and best practices. This type of peer learning allows for households to obtain the necessary knowledge to adopt new technologies and succeed in the contests.[9]

Contests are the primary motivational tool used by PMR. Contests take place on two levels: inter-community and intra-community. The intra-community contests are 6-month contests that put participating households in competition against one another. During the contest, elected "jury members" (who are also members of the same community) visit each household to asses progress, clarify rules and give the families feedback on their progress. This feedback is recorded as a binary score by the jury member. 1 indicates progress made and 0 indicates no progress at all. During the final visit at the end of the contest, each household is awarded points based on their progress relative to the beginning of the contest. Progress is measured by a predesignated list of activities that are designed to promote investment in participants' households, agricultural practices, entrepreneurship and the environment. Each activity is then assessed on a 1-10 point scale. After totaling the final scores of each participating household, the 6-7 households[10] with the most points win cash prizes. These prizes typically represent about 15-20% of the households annual income and therefore are economically significant incentives for families. Prizes are always awarded at a public ceremony, which adds additional incentive for

[9] An important caveat to this qualitative observation is that culturally, it is not polite for people in the Andean region to cite money as a motivator, even though it represents a significant economic value. Thus, even though most cited peer-learning as a prime motivator, I suspect the financial incentives play an equal if not more important role in motivation.

[10] The number of prizes depends on the number of people competing. For contests with more than 45 contestants, there are 7 prize winners, otherwise there are 6.

households due to the improved social standing that comes with winning.[11] The prize structure depends on how many families are participating, and typically increases after the first contest. The number of contests conducted in a given community depends on funding, but due to data restrictions, this paper will look at communities that have had only two contests.

The inter-community contests are competitions between communities within the same region and are designed to promote team-work amongst community members who are competing with each other on the intra-community level. The contest is executed in the same way as above, but uses a separate list of activities that are designed to promote community-wide infrastructure development as well as community organization. Prizes are awarded to the top 3 communities and the winnings are required to be spent on a mutually beneficial project for the community. The presence of the inter-community contests have to potential to generate an incentive for households to cooperate by sharing best practices, materials and knowledge of specific activities.

For both types, after the first 6-month contest, the second begins immediately and has the exact same structure with a different list of activities designed to build on the first contests' progress. One interesting detail of the second contest is that PMR automatically sorts the winners into a separate "league" called the Champions League. The purpose behind sorting out the winners is to avoid any disincentive effect created by having the same households win contest after contest. PMR's contest design, prize structure and automatic sorting provide a good setting to test hypotheses from tournament

[11] Status and recognition as incentives have been studied in competitive settings. Kosfeld and Neckermann (2011) conduct an experiment where students work on a database project and are awarded with a card for the best performance. Despite the fact that the card was purely symbolic and had no monetary value, the authors found that students in the treatment group had, on average, a 12% increase in effort compared to the control group. Their results suggest another source of motivation in this setting, but unfortunately any analysis of social effects are beyond the scope of this data.

theory. The contest design resembles a basic rank-order tournament, which has been studied in experimental and business contexts and is often used for its simplicity. The prize structure varies across community and contest, which provides a setting in which it is feasible to tests effects of this variation, while the automatic sorting element allows for tests on the effect of sorting on effort. These hypotheses will be formally outlined in the following section.

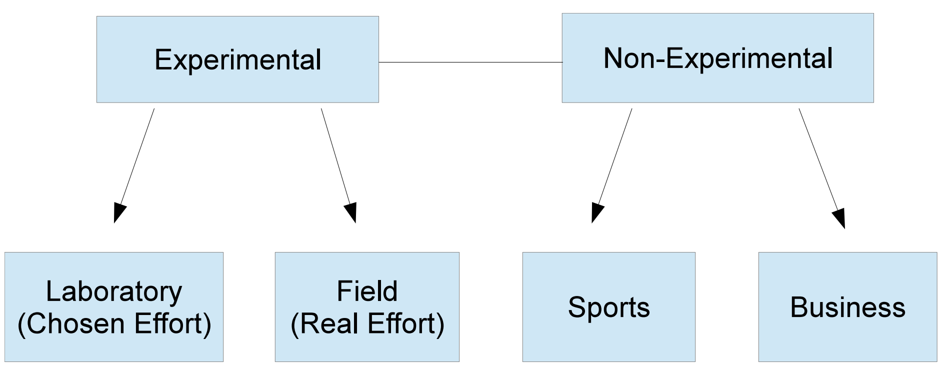
3 Literature Review

Contests are competitions used as incentive devices where agents are rewarded fixed prizes according to their relative performance compared to other participating agents. The theoretical foundations of behavior in contests were supplied by Lazear and Rosen (1981), who used a two-person model to outline the implications of contests and their incentives. More specifically, they outline how prize structure and contest structure can affect the effort levels of agents. They also argue that contests mimic the efficiency levels of piece-rate compensation schemes while avoiding the high costs of measuring output. Thus, the implicit argument is that under certain conditions, contests can provide sufficient incentives for individuals to perform better than traditional compensation schemes like hourly wages or salaries.

Their work has been extended by relaxing assumptions of heterogeneity and risk neutrality and examining different contest structures. Green and Stokey (1983) analyze rank-order contests in the presence of multiple risk-neutral agents and examine the implications of different random shock distributions. They show that contests can be a useful tool to eliminate the presence of a common shock, but are not, in general, "optimal" contracts due to the fact that a Nash equilibrium of effort is not always attainable. Nalebuff and Stiglitz (1983) further deepen the analysis by introducing risk-aversion and multiple prizes. They show how introducing risk aversion into the theoretical model with multiple prizes can affect incentives. Risk-averse agents, however, respond to changes in prize structure. The authors find this response depends on the distribution of the random component of output, but only formally test a uniform distribution. List et al (2010) elaborate on this observation and formally outlines the implications of having an increasing, decreasing or uniform density of uncertainty in output. They show that each type of distribution elicits a different effort response from individuals.

The above theoretical contributions generated a large body of empirical literature testing the theory in a variety of contexts. Over time, the literature has separated itself into two distinct branches, one that uses data from experiments, where the ability to control tournament inputs allows for researchers to compare the incentive effects of tournaments with various degrees of risk and heterogeneity. The main drawback to the experimental approach is that they use simplified settings to test theory and therefore may not predict behavior in more complex social settings. The other branch uses non-experimental data, usually from sporting events or businesses, in order to test whether or not the theoretical predictions hold in real-world settings. Figure 1 depicts the general structure of the literature.

Figure 1: Organization of the Empirical Literature



Both branches test whether or not specific theoretical predictions fit with actual behavior, the most common hypotheses tested are those that deal with effort. Understanding what aspect of contests directly affect effort levels can allow for the optimal design of contests and allow for them to be implemented as incentives devices to improve efficiency. The theoretical models show that contests can be more efficient in terms of effort, thus the empirical literature aims to identify the specific determinants of effort under contests. This paper addresses 3 of the most commonly tested hypotheses generated by theory:

1. The "n-effect" hypothesis: Increasing number of households from n to $n+1$ while holding the number of prizes fixed will decrease performance

2. The "spread effect" hypothesis: Performance is increasing in the prize spread

3. The "heterogeneity effect" hypothesis: High-ability households will adopt low-risk strategies while low-ability households will adopt high-risk strategies

These 3 hypotheses have been studied in a variety of contexts, both experimental and non-experimental, but due to lack of data, have not been as exhaustively studied in field settings. The following outlines the relevant literature for each.

The N Effect

Gibbs (1996) provides the theoretical foundation for this prediction. He shows mathematically that the effect of the number of participants in a contest on individual effort is dependent on the distribution of the random component of output, e. He determines that with a uniform distribution, the is no effect on effort because the marginal propensity to exert effort (MPE) is independent of both the number of contestants as well as the number of winners. He continues by showing that if the distribution is symmetric and unimodal, then adding one contestant while holding the number of winners fixed will decrease the MPE as long as the number of winners is less than half of the total number of contestants. He adds that if the number of winners is exactly equal to the half of the contests then no change in effort will be observed, and if the number of winners is greater than half of the contestant pool, then MPE will increase. Another prediction generated by his construction of the theory is that increasing the number of contestants, while holding the ratio of winners to contestants fixed, will decrease MPE for probabilities close to 0.5 and will decrease MPE for extreme p-values close to 0 or 1. He adds that this economies of scale effect is more likely to hold if the distribution is less peaked (has more variance) or the contest size is smaller.

Non-experimental evidence is relatively limited for this hypothesis. Casas-Acre and Martinez-Jerez (2009) explicitly test the predictions of the Gibbs model by using institutional data from a firm that began implementing contests as incentive devices in sales of commodities. The firm is located in a developing country and decided to use contests not only as incentives but to also negate common shocks in commodity markets. The setting differs slightly in that the contests are all two-period, dynamic competitions. The dynamic element is due to the fact that performance in both period contributes to each contestants final score and the contestants are informed of their relative position between periods. Most of their analysis focuses on the incentive effects created by the dynamic component, but they also test the effect of increasing contest size on effort. Their analysis uses a difference in differences approach and supports the theoretical implications proposed by Gibbs; effort decreased as the contest size was increased (holding the proportion of prizes constant at approximately 10%.) The implications of this result are that Gibbs' model seems to hold, but more empirical work is needed to verify its various implications.

Boudreau et al (2012) was the first to call this hypothesis the "n-effect" and tests the effect that increasing the number of competitors has on average effort in a computer programmer's tournament. The authors begin with the observation that in prior research, individuals in performance-based contests perform worse in the face of increased competition. Using a fixed effects estimation strategy, they find a small but significant negative relationship between number of competitors and effort. As a robustness check, the authors use a variety of alternate methods for controlling for heterogeneity in contests, they use OLS with no controls, contest fixed effects, competitor fixed effects, etc. They find that low-ability competitors have a significant and negative reaction to increased competition, but the majority of the "n-effect" comes from high-ability players. All in all, the authors find that the n-effect does exist and is in line with theoretical predications.

Experimental evidence surrounding this hypothesis has been mixed; both positive and negative effects of increased contest size on effort have been found in recent literature.

Dechenaux et al. (2012) provide a detailed survey of all the current experimental research being done on contests and auctions. They review studies that have analyzed the relationship between contest size and performance and find that the literature is mixed in regard to finding a well-established conclusion. Many of the papers surveyed conclude that the relationship is expected to be negative, while others conclude that the relationship depends on the distribution of noise, in which performance can be increasing, decreasing or remain unchanged as the number of contestants increases.

List et al. (2010) bridges the gap between these two findings by using both a laboratory and field experiment to highlight the reasons why evidence surrounding this prediction have not been consistent. They first outline a theoretical base that predicts an ambiguous effect of a change in number of contests. The ambiguity is due to the fact that the distribution of the error term is unobservable and often assumed to be normal. The authors show that the error term (uncertainty) dictates the direction of the effect of increasing the number of competitors. They show: (i) If the form of uncertainty has decreasing density, increasing the number of contestants will decrease equilibrium effort levels. (ii) If the form of uncertainty has uniform density, increasing the number of contestants will leave effort levels unchanged. (iii) If the form of uncertainty has increasing density, increasing the number of contestants will increase effort. They first use a laboratory setting to study experimental markets in which on the shape of the common uncertainty changes. Under risk-neutrality, they find weak support for the theory. But, under risk aversion, the results strongly support their theory.

The second study uses a field experiment from a dutch recreational fishing outlet. They control inputs like number of fish stoked and find evidence supporting their theory. Thus, they show that finding any relationship between the number of contestants and effort is possible, but is really reflecting the shape of uncertainty in output.

Harbring and Irlenbusch (2002) also test this specific prediction from Gibbs by experimentally examining how both contest size and prize structure can affect effort levels. They implement rank-order contests with 2,3 and 6 homogeneous agents. They analyze the effect of prize structure by varying the number of winners prizes. They look at the fractions 1/2, 1/3 and 2/3 (number of participants divided by number of winners prizes.) The contests are designed so each agent in each round simultaneously choose effort levels and the corresponding convex cost function. They measure effort as output and after each round, agents are informed of the other agents' output levels. They find that the prize structure does in fact have an effect on effort levels; the higher proportion of prizes leads to higher average effort. They further investigate this effect by allowing a restart effect after 10 rounds. After the restart, they find that effort decreases dramatically, therefore supporting the notion that learning or strategic behavior is occurring over the first 2 rounds which drives effort to increase. They conclude that effort increases over both aspects of contests: prize structure and size, with the former being the more significant motivator. The authors recognize that this finding is similar to the body of literature that analyzes the effect of prize spread on effort (the "spread effect") and note that the two tests are similar in the incentives they offer to agents. From a theoretical point of view, the expected value of winning is the same in both schemes, but psychological factors like relative deprivation may lead to different efforts levels in an empirical context.

The Spread Effect

Non-experimental evidence mostly uses sports data to test this hypothesis. Ehrenberg and Boganno (1990) study the incentive effects of contests using data from the PGS tour. Their analytical framework uses Lazear and Rosen as the theoretical basis, and views the two person contest as one person vs. the rest of the field. Using this assumption, many of the theoretical conclusions drawn by Lazear and Rosen can be tested using the PGA data. The authors assume that effort is a choice variable and that each individual faces a "cost of effort-concentration" function and that the marginal cost of effort is positive and increases as effort increases. The data comes from 1988 PGA results and the player's scoring average on all rounds during a year as a measure of his ability. They estimated equations after pooling the data across individuals and contests. They estimate ability as the dependent variable and as independent variables: total prize amount awarded, a dummy variable to represent a major contest, a vector of variables to control for the difficulty of the course, a vector of variables to control for the quality of other players, a vector of variables to measure an individuals ability and a random error term. The authors analyze the marginal return to effort in the final round of play in all contests. They find that higher marginal return of effort increases performance. This is very similar to prize spread increasing effort and may be an alternative way to approach the question.

Becker and Huselid (1992) use panel data from auto racing to show that prize differentials have incentive effects on individual performance. The data used come from the 1990 NASCAR circuit and the 1990-91 IMSA circuit. The NASCAR data use 44 different drivers over 29 races and exclude drivers who did not compete in at least 5 races. One of the races was excluded because its prize was significantly larger than others. They measure performance by constructing a variable that reflects both the order of finish and average speed of the race. The races were normalized such that the fastest race equaled 1 and slower races have proportionally larger scores. The authors find a significant relationship between prize spread and racer performance. They also find that these incentive effects have a limit and begin to decrease after the spread increases past a certain threshold.

Frick and Humphreys (2011) also use data from NASCAR using race-level data. They find support for the n effect hypothesis by looking at the effect the prize spread has on average race speed. Their argument is that effort, in the NASCAR context, can be interpreted as increasing speed in a race. They find that the average speed does in fact increase with the prize spread.

Experimental evidence testing this hypothesis has found the existence of this effect both in the laboratory and the field. Falk et al. (2008) approaches this question in a slightly different way from others, they look at contest design from the principal's point of view. Their results are still relevant and operate off the same theoretical foundation (Lazear and Rosen (1981)). They use an experiment to test what factors affect principals' decisions in contest design. They find that principals do in fact have incentive to increase prize spread; effort is monotonically increasing in the spread. The firms profits are also increasing, thus providing evidence that the principal also has sufficient reason to maximize the prize spread (in this particular experimental setting.)

Delgaauw et al. (2012) use a natural field experiment to test this hypothesis. Their data comes forma large retail chain that had a random subset of 208 stores participate in a two-stage elimination contest, The distribution of contest prizes varied across rounds, and the authors were able to take advantage of this variation to identify the effects of spread on effort. They find that increased prize spread does in fact enhance performance (from the first round to the second round). They also find that workers with high variation in performance seem to respond less to contest incentives.

The Heterogeneity Effect

A major challenge to contest design in the real world is contestant heterogeneity. The competition triggered by closely matched competitors quickly dwindles as the gap between competitors ability is increased. This type of asymmetry is expected to produce sub-optimal effort levels. Riis (2010) suggests that heterogeneity tends to make contests less predictable and stimulates strategic behavior amongst contestants.

The choice of risk in tournaments and contests has been studied in a limited number of papers. Bronars (1986) is responsible for first discussing the idea that a front runner prefers to take less risky strategies to protect his lead while a trailing opponent will be more prone to adopt a riskier strategy to change his ranking. Since his initial contribution, there have been few theoretical studies to extend his ideas. Hvide (2002) uses a purely theoretical approach to explore questions related to this hypothesis. He expands Lazear and Rosen's original model of output to include risk as a choice variable. With endogenous risk and heterogeneous players, the author shows that the high ability player engages in less risk through two different mechanisms. The first is a positive effect, decreased risk increases a strong players probability of winning. The second, negative effect, is that his equilibrium effort level will increase. Hvide (2002) shows that with high degrees of heterogeneity, the first effect dominates the second, but when the asymmetry is more subtle, and the second effect dominates due to the fact that the probability of winning converges to 0.5 (for a two-player model) as heterogeneity decreases. Thus, the strong player winds up adopting a more risky strategy if contestants have similar but not equal abilities.

Kraekel and Slikwa (2004) also use a purely theoretical approach to show that (similarly to Hvide) there are two important effects of risk taking in contests. On the one hand, it effects effort levels (effort effect) and on the other, it affects winning probabilities (likelihood effect). They show that if agents' abilities are similar, their efforts will decrease in risk, but if agents are different in ability, the opposite will occur. They also find that the direction of the likelihood effect does not depend on agent heterogeneity. In this case, the high ability agent always prefers a low-risk strategy and the low-ability agent always prefers high-risk. They conclude that different equilibria are possible depending on the magnitude and interaction of the two effects.

Non-experimental Evidence is also limited for this hypothesis. One of the most prominent papers that tests this prediction is Knoeber and Thurman

(1994). The authors use data from broiler chicken producers and are one of the only studies that use organizational data to study the incentive effects of contest heterogeneity on effort by examining the strategies chosen by different "types" of participants. Theory predicts that in mixed contests (contests will heterogeneous players) higher ability players will choose less risky strategies while low ability players will do the opposite. Using data from broiler chicken growers' production, the authors develop a linear model that relates settlement cost (the inverse measure of performance) to two numeric variables, two sets of dummy variables and a set of seasonal variables. They find that there is in fact a negative relationship between grower quality and variance of performance. They interpret variance of performance to be indicative of a risky vs. non-risky strategy. Thus, they find that high-ability growers do in fact adopt lower-risk strategies than their low-ability counterparts.

Grund, Hocker and Zimmerman (2010) examine this question in the context of the NBA. The use 3-point shots in the fourth quarter as a measurement of risk and find that trailing teams increase the percentage of 3-point shots, suggesting that being in a disadvantaged position evokes riskier strategies than being the front-runner.

Experimental research on this hypothesis was first conducted by, Bull et al. (1987) using an experiment with chosen effort. The authors find that in asymmetric tournaments, disadvantaged subjects, on average, display effort levels above the levels predicted by theory, while the effort levels of the advantaged subjects converged to predicted levels. Their findings offered the initial evidence in favor of the hypothesis that strategies might differ with heterogeneity of ability. Van Dijk et al. (1999) found similar results, this time using a real effort experiment. They attribute the findings to either a strategic decision or a decision made because of "peer-pressure" from fans or a combination of the two.

Nieken and Silkwa (2007) develop a theoretical model that predicts that risk-level choice in strategies depends not only on ability, but also the correlation of the outcomes of the strategies. Their model suggests that when the correlation equals 0, then we can expect agents to behave as other papers have found (Hvide, Kraekel and Slikwa) whereas if the correlation approaches 1, then the results of the aforementioned papers do not hold. Intuitively, it seems correct that if outcomes are highly correlated, than strategy choice becomes less impactful. They run an experiment where outcomes are not correlated, perfectly correlated and correlated with a coefficient of .5. They find that previous results hold and high ability agents choose low risk strategies 98.9% of the time, while the results are less predictable in the case of correlation.

4 Theoretical Model

Lazear and Rosen (1981) provide the seminal theoretical analysis of the basic theory of contests and tournaments. Their model is structured as a two-person competition where both competitors are vying for prizes paid by a firm. Both competitors are risk neutral and identical, as is the firm. The analysis relies on the fact that competition takes effort, thus players will only participate if the expected prize is larger than the anticipated cost of effort required to win. Thus, the decision-making process is two-fold; on one hand, the players must decide how much effort to exert, while on the other, the player must then make a determination as to whether or not the expected winnings justify the corresponding amount of effort required to obtain the prize. The authors propose that, in equilibrium, the firm must set prizes that are efficient on both margins of the player's decision making. The theory laid out in their paper shows that such an equilibrium exists and compares the outcome with that of piece-rate pay. The implied argument throughout the paper is that under risk neutrality, tournaments can perfectly replicate the efficient results of piece-rates by producing the Pareto optimal allocation of resources.

I begin by considering the simplest, two-player tournament that follows Lazear and Rosen. The rules of the game specify a fixed prize W_1 to the winner and a fixed prize W_2 to the loser and $W_1 > W_2$. The game is considered rank-order because the margin of winning does not affect the outcome, whereas relative performance does. All essential aspects of this game can be generalized to multiple contestants, but introducing complications like contestant heterogeneity complicate the mathematics somewhat. These complications will be considered below, but first, it is helpful to begin with this simple case in order to see the fundamental theoretical implications and how they change as the model becomes more complicated. This process will help paint a clearer picture of how theoretical predictions can be tested in a rural development setting using PMR's data. Let output, or in this case performance for household i, be measured by q where:

$$q_i = u_i + e_i \tag{1}$$

Here, u is an action taken by the household and e is a random component that can be thought of as representing random shocks or luck to the individual. In this setting, household action can also be thought of as making effort towards improving agricultural technologies or living conditions. This action, however, is costly and therefore each household has a cost function, $C(u_i)$ associated with u_i where $C^{'} > 0$ and $C^{''} > 0$. The larger prize will always be awarded to the household with the higher relative performance.

The probability that household i wins, depends positively on its effort level, u_i and negatively on the effort level of other households, u_j and upon the distribution of the random component of performance, e_i. The expected payoff for household i is:

$$P\left[W_1 - C\left(u_i\right)\right] + (1 - P)\left[W_2 - C\left(u_i\right)\right]$$

$$= P\left(W_1 - W_2\right) + W_2 - C\left(u_i\right), \tag{2}$$

where P is the probability of winning. The probability that household i wins can be defined as

$$P = prob(q_i > q_j) = prob(u_i - u_j > e_i - e_k)$$

$$= prob(u_i - u_j > \xi) = F(u_i - u_j), \quad (3)$$

where F(.) is the CDF of ξ. Each household chooses to maximize equation (2) subject to chosen effort level u_i, this implies

$$\frac{\partial P}{\partial u_i}(W_1 - W_2) - C'(u_i) = 0. \quad (4)$$

Equation (4) shows the equivalence of the marginal value of effort and the marginal cost of effort. It should also be noted that $\frac{\partial P}{\partial u_i}$ is the pdf of $F(u_i - u_j)$ and can be expressed as $f(u_i - u_j)$. Thus, the above equation shows that two factors affect the cost of effort: the prize spread and the distribution of the uncertainty or error component. This distribution can also be thought of as a density of "luck." A reasonable assumption is that the error terms are normally distributed. Symmetry (homogeneity) implies that a Nash-Cournot equilibrium exists, implying that $u_i = u_j = u^*$ and $F(.) = 0$ and therefore $P = \frac{1}{2}$. In this case the outcome is purely random in equilibrium.

The "N-Effect" Hypothesis:

Increasing number of households from n to $n+1$ while holding the number of prizes fixed will decrease performance

It follows logically that by increasing the number of households while holding the number of prizes constant, the probability of winning a prize falls and therefore household i has less incentive to invest a large amount of effort. In fact, many empirical findings have supported this idea. But, if households cooperate, then increasing n makes the outcome less dependent on luck and more dependent on effort. Thus, increasing n can potentially increase performance. Following List et al. (2010), I relax the assumptions of a two households model and allow for multiple households, n. Consequently, when a household maximizes expected utility, equation (4) becomes

$$(W_1-W_2)\int_{-\infty}^{+\infty}(n-1)f(u_i-u^*+e_i)F^{n-2}(u_i-u^*+e_i)f(e_i)de_i-C^{'}(u_i)=0 \quad (5)$$

and therefore in symmetric equilibrium $u_i = u^*$ the above reduces to

$$(W_1-W_2)\int_{-\infty}^{+\infty}f^2(e_i)F^{n-2}(e_i)de_i-C^{'}(u_i)=0. \quad (6)$$

Using Integration by parts, equation (6) becomes

$$\int_{-\infty}^{+\infty}(n-1)f^2F^{n-2}de=f(+\infty)-\int_{-\infty}^{+\infty}F^{n-1}f^{'}de. \quad (7)$$

Since $f(+\infty)$ does not depend on n, taking a derivative with respect to n yields

$$d\int_{-\infty}^{+\infty}(n-1)f^2F^{n-2}de/dn=\int_{-\infty}^{+\infty}(-lnF)F^{n-1}f^{'}de, \quad (8)$$

which indicates that the relationship between the number of contestants and effort is dependent on the form of the density of uncertainty, $f^{'}$. Naturally, when a contestant chooses effort levels in a contest, he naturally compare the marginal benefits to the cost of effort. As contest size increases, the probability that another contestant will win increases, regardless of the form of the density function of uncertainty. For the relationship between contest size and effort to be positive, the density function must be increasing in n. This means, more contestants means a higher probability of receiving a good “draw”. Hence, as the group becomes larger, pure luck is less likely to determine the winner and effort becomes to most influential factor in determining the winner. This is a reasonable assumption in this setting for two reasons: (1) PMR encourages peer learning as a central pillar of their development model. Thus, within a given community, households are encouraged to share expertise, best practices, etc. In the context of a contest, this means that a certain level of cooperation between families within the same community is expected. (2) PMR also encourages cooperation by implementing an inter-community contest alongside the intra-community contests. This extra "layer" of competition provides additional incentive for households to work together. Thus, assuming that increasing contest size also decreases the likelihood that pure luck will determine the winner is a logical conclusion; since knowledge and specialization are shared, effort becomes the most important factor in determining a winner. As a result, it is possible to see a positive relationship between contest size and performance in this setting.

The "Spread Effect" Hypothesis:

Performance is increasing in the prize spread

This hypothesis is implied by equation (4). It follows that effort, u_i, and therefore performance, q_i are both positively dependent on the prize spread, $W_1 - W_2$. Note that the relationship does not depend on prizes levels. It also follows that performance and effort are both negatively related to the marginal cost of effort. Given the above implications, when all households in a contest are maximizing expected utility, then effort becomes a function of the prize spread and the shape of the random error component, implying:

$$u^* \equiv u^* \left(\left[W_1 - W_2 \right], \sigma \right). \tag{9}$$

This is one of the most commonly tested hypothesis in the empirical literature, which has consistently found this prediction to hold in a variety of business and sports settings. This will be the first time the hypothesis is tested in a rural development setting. Equation (4) shows the relationship between effort and prize spread for the simple, two player contest, but the PMR contests are much larger than two households, so an appropriate theoretical prediction is still needed for this particular setting. As in hypothesis 1, I refer to the theoretical model proposed by List et al. (2010). Equation (5) describes an agent maximizing expected utility in a contest with n households. As shown above, the identity in equation (9) still holds, regardless of the number of contestants. Thus, there is a clear theoretical prediction that prize spread and performance are positively related.

The "Heterogeneity Effect" Hypothesis:

High ability households will adopt low-risk strategies while low ability households with adopt high-risk strategies

This hypothesis is implied by Equation (9), which shows that effort is a function of both the prize spread and σ, the variance of $g(.)$.[12]Assuming a mixed contest, I will test the theory that the household with the lower marginal cost of effort (high-ability) will choose a lower variance strategy in order to maintain its position at the top of the rankings. This is because high ability households have little incentive to take risk (high-variance strategies), while households with higher marginal cost of effort (low-ability) have little to lose by taking on risk (high-variance strategy). By adopting risk strategies,

[12] Assuming a normal distribution is a simple way to show that equilibrium effort is decreasing in σ. Substituting for a normal density in equation (8)

$$\frac{\partial C}{\partial u_i} = \frac{\triangle W}{2\sqrt{\sigma^2 \pi}}$$

shows that effort is decreasing in σ. Hvide (2002) notes the intuition behind this, mainly that higher variance makes outcomes more "noisy" which can lower equilibrium effort levels.

it is meant that a household induces a mean-preserving spread of output, y_i, through increasing the variance of e_i. A simple way of thinking about the mechanism in which variance of e_i can affect output is reverting back to equation (1). In this simple additive approach, the only choice variable is effort, u_i, but the above hypothesis implicitly changes that assumption. In order to adopt low or high variance strategies, and have that choice affect output, risk must become endogenous. Hvide (2002) outlines the implications of this and shows that the variance of the error term, e_i, can be thought of as

$$\eta_i^2 = \sigma^2 + s_i^2 \tag{10}$$

where σ is the level of background noise and s_i is the degree of voluntary spread in the output distribution.[13] Now household i has two choice variables: u_i and s_i. With this modified variance, we can now expect the choice of risk in strategy to affect effort levels, which is characterized by equation (8).

By assuming a mixed contest, I am implicitly introducing complications into the above model in one of two ways: by augmenting equation (1) to include an individual measure of ability, α_i, or assuming heterogeneity in cost. The former implies a slight modification of equation (1) to become

$$y_i = u_i + \alpha_i + e_i \tag{11}$$

This implies a change to equation (4) by means of affecting $\frac{\partial P}{\partial u} = f(\cdot)$. The change suggests the density function is now also a function of ability, α. But, assuming cooperation within contests, this type of heterogeneity might be mitigated when households work together. Instead, it is more logical to think

[13] It is important to note this does not change the implication of the previous footnote. Instead of plugging in σ, you plug in η which is a function of σ and retains the relationship implied by the hypothesis.

of heterogeneity in terms of cost.

Thus, the second possible modification considers the heterogeneity to come from different levels of ability in that some agents are more "able" to cover the associated costs, i.e. each household has a different cost function. It the follows that more able households can produce the same output at a lower cost than low ability households. More formally, if households are assumed to be of two types, a or b, where $C_a^{'}(u_i) < C_b^{'}(u_i)$, suggesting that the marginal cost of investment is higher for type b. This can also be interpreted as type a households being of higher ability than type b households. I also assume that the proportion of $a's$ and $b's$ is β and (1-β), respectively, then equation (2) for type $k = a, b$ becomes

$$W_2 + [\beta P_a^k + (1-\beta) P_b^k](W_1 - W_2) - C_k(u_k) \tag{12}$$

where (W_1^M, W_2^M) are the prizes for the mixed tournament and P_t^k is the probability that a household of type k defeats a household of type t, where $t = a, b$ and $t \neq k$. The first order condition of (6) changes equation (3) to

$$\left[\beta \frac{\partial P_a^k}{\partial u_k} + (1-\beta) \frac{\partial P_b^k}{\partial u_k}\right] (W_1 - W_2) - C_k^{'}(u_k) = 0 \tag{13}$$

Thus it shows that in mixed tournaments, effort is still dependent on the wage spread. This also shows that mixed contests yield inefficient outcomes compared to homogenous contests. In order for a mixed contest to be efficient, then $C_a^{'}(u_a) = V = C_b^{'}(u_b)$ must hold. This equality only holds in the special case where $\beta = \frac{1}{2}$, otherwise the heterogeneity causes both types of households to under-invest.

Therefore, regardless of the way in which you quantify heterogeneity, the effect is only on the probability of winning. Thus, sorting households according to their abilities, k, can improve efficiency and reduce the effect of heterogeneity. Since there are efficiency costs to mixing households of unequal ability, both able and less able households have incentive not to exert effort because their final position in the contest is relatively dependent on their innate ability, or financial resources. Since ability can be thought of as having more financial resources in this context, it is probable that heterogeneity of ability exists and therefore theory predicts that sorting can mitigate the disincentive effects created by heterogeneity.

5 Data

The data come from three regions in southern Peru and track the progress of all participating households over two contests. The three regions were all a part of PMR's Ayuper (Aid for Peru) project, which focused on post-disaster relief for the provinces south of Lima hit hardest by the 2007 earthquake. The contests took place from July 2009-July 2010 for all three regions of the Ayuper project.

At the beginning of each contest, every community elects "jury members" to be in charge of visiting households, give feedback on their progress and give the final scores. The feedback is documented by the jury members as a binary score (check-in score) indicating whether the households has made progress on each activity listed in the appendix, where 1 indicates progress and 0 indicates no progress. The final scores are given at the end of each contest and grade to progress a household has made on each activity relative to the beginning of the contests on a scale of 1-10. After totaling the final scores of every participating household, the top 6-7 households receive cash prizes in a public ceremony.

PMR stored both the check-in and final scores in separate excel files for each community. The data compilation process consisted of appending each communities file, then merging the files across contests. The performance measures were created by summing or averaging scores for each individual in a given month and contest. To create variables indicating the prize spread, the prize amounts (provided by PMR) were entered and the spread were created by taking the difference between prize amounts. The contest size variable was created independently from PMR by adding the number of households in a given community during a given month. Additionally, various dummy variables were created to indicate community, the winner of a contest, and contest.

Each region uses the same list of activities that focused on repairing homes, improving agricultural practices and starting small businesses. The contest rules are also the same across regions, but variables like prize structure, number of contestants, and number of check-ins change across region and contest number. Tables 1 summarizes key variables for the entire dataset and a balanced panel of the data.

All of the prize variables (first prize-seventh prize) show the amount awarded to the household that finishes in each place and is measured in Peruvian Soles. The approximate conversion rate during the time of the contest ranged from \$.33-\$.40 USD. This implies that the average first prize ranged from about \$115 to \$215 USD across both contests. The magnitude of these prizes is very large for most households, representing approximately 2-3 months income.[14] Contest size counts the number of households in the contest at any given month. Since PMR households are free to enter or exit the contest at any time, this

[14]Since income measures were not available in the PMR data, monthly income was approximated using summary statistics from the "Extreme Poverty Graduation Program" baseline survey conducted by Innovations for Poverty Action (IPA) in the summer of 2011. The survey showed that each individual in a household earns about 60 Soles a month and each household has, on average, about 2 working adults. Thus, I approximate monthly household income to be around 120 Soles a month, or \$46 USD.

number can vary month to month. For consistency, all estimates use either the contest size from the final round or the check-in rounds exclusively. The spread variables all measure the corresponding prize differential. For example, “1st 2nd Spread” measures the differential between first and second prizes. The prize spread only changes across contests and typically increase during the second contest. This is because PMR generally offers a more lucrative prize structure after the first contest. The total spread variable measures the difference between the first and last prizes and was created in order to capture any effect the overall spread might have on performance. Unfortunately, none of the regions had much variation in the spread or total spread variables. This is because many communities had the exact same prize structure.

Raw final score is the main performance measure used for all estimations. It is simply the total number of points accumulated from the pre-determined list of activities for each household in a given contest. Raw check-in score is an additional performance measure used in hypothesis 3 and is the average of the binary scores accumulated by each household during a contest. The distinction between these two types of scoring methods is important for the analysis. Since the dataset is only comprised of two contests, the raw final score is limited to 2 observations per household at most, thus the raw check-in score offers the benefit of more observations per household, but sacrifices accuracy by simplifying performance to a binary score. For this reason, raw final score is used as a performance measure in the first two hypotheses while raw check-in score is preferred for the third.

The only demographic and geographic controls this dataset provides is children per household and regional location of each household. The general lack of controls implies the potential for omitted variable bias in my analysis. Crucial indicators like income, proximity to local markets and education would ideally be controlled for, but the limits of the data restrict my ability to control for other factors. All regressions use these controls in an attempt to minimize omitted variable problems.

6 Identification and Empirical Testing

This section describes the data and empirical strategy used for each hypothesis. For each hypothesis, summary statistics are given for variables relevant to the analysis followed by a detailed description of the estimation strategy.

The "N Effect" Hypothesis

This hypothesis predicts a negative relationship between contest size and performance, based on the findings of several empirical studies. In the context of repeated contests with sufficient incentive to cooperate, it is possible that increasing the number of contests will increase cooperation and sharing of expertise amongst households. This implies that the probability that luck determines the outcome is reduced, making effort (or performance) become the most important factor in determining a winner. Thus, increasing n can potentially increase performance. An additional factor that needs to be controlled for in this case is the size of the prize in each community. If prize size is ignored, there is a potential that the results will be confounded by the the incentive effects of the prize as opposed to the size of the contest. Theory predicts that larger prizes will induce higher effort levels, while the theoretical predictions for the effect of contest size on effort is less clear; as List et al (2010) demonstrates, effort can be increasing, decreasing or remain the same depending on the form of the density of uncertainty.

Table 2 shows the variation of contest size within each prize amount category that is used to identify the effect of increasing contest size on performance. The left column shows all the contest sizes, and the remaining columns reports the number of communities for each contest size. Since I am looking at how the size of the contest as a whole affects performance, the unit of observation for this hypothesis is community and therefore any relevant variation of contest size takes place on the community level. Therefore, Table 3 shows this variation by displaying how many observations (number of communities) are available for the analysis. The largest number of observations occurs when the prize amount equals 350 or 450. The other prize categories have fewer observations and consequently less variation. The first two prize amounts correspond to the first contest and the last three correspond to the second. Thus, in each contest, there is at least one prize amount that had a sufficient number of observations and variation.

Estimation Strategy

The empirical approach uses a simple OLS regression of performance on controls and contest size while holding the prize amount fixed:

$$P_{itc} = \beta_0 + \beta_1 X_{itc} + \beta_2 Contestsize_{tc} + \beta_3 Contestsize^2_{tc} + \beta_4 Prize_{tc} + \varepsilon_{itc} \quad (14)$$

Where P is average performance of household, i, during contest, t, in community, c. Performance is measured by the final number of points received. X_{itc} is a vector of controls that includes the number of children in each household and the specific region within Ayuper that the individual lives in. $Contestsize$ is the size of the contest, t, for a given community, c and $Contestsize^2$ is used to measure any non-linearities in the relationship. $Prize$ is a vector of dummy variables, all of which equal 1 for a certain prize amount. Since there are 5 total prize amounts, this vector consists of 4 different dummy variables. Any error is captured by standard errors, ε_{itc}, which are clustered at the community level. As outlined in section 3, theoretical predictions suggest that increasing the number of contestants can have either a negative, positive or neutral effect on performance. Following the model put forth by List et al. (2010), the result $\beta_2 < 0$ would indicate that either households are risk averse or the density of uncertainty is decreasing in n. If $\beta_2 > 0$, the implication is that increasing contest size has the predicted effect on performance. This implies the form of uncertainty has an increasing density, or that the incentive effects created by peer learning and the inter-community contests are strong enough to encourage cooperation. If $\beta_2 = 0$, then there is no effect of size on effort. This also suggests that the form of uncertainty has a uniform density.

The "Spread Effect" Hypothesis

This hypothesis predicts that increasing the spread between prizes will increase performance, on average. The spread effect is identified from household-level variation in the spread variables. Due to data restrictions, I only look at the spread between the 4th and 5th prize amounts. Ideally, the analysis would use all spread variables, but in this case of this dataset, the other spread variables do not have sufficient variation.[15] Table 3 shows the variation of the spread between 4th and 5th prizes across contests. The upper half of the table shows the data for the spread variable using all observations. This includes households who participated in only one of the two contests. In order to develop a clearer picture of the size of both the treatment and control

[15]The variation needed to test this hypothesis is in prize spread across contests. Table 1 shows that for many communities, the prize spread remains the same in both contest 1 and contest 2 and therefore can not be used for this analysis.

groups, the lower half shows only the households who participate in both contests. The table shows that as households move form the first contest to the second, there is a group of households that experience an increase in the spread between 4th and 5th prizes, while the rest of the households do not. These two groups represent the treatment and control, respectively and, assuming that all communities can be considered more or less the same, any difference in effort levels seen between the two groups can be interpreted as the effect of increasing the prize spread.

Table 3: Prize Spread and Number of Households For Each Contest

All Contestants		
Difference between 4th and 5th prizes	Contest 1	Contest 2
0	709	987
50	101	318
100	0	204
Total	810	1509
Balanced Panel		
Difference between 4th and 5th prizes	Contest 1	Contest 2
0	364	292
50	48	102
100	0	18
Total	412	412

Note: The top table uses the entire sample, which includes individuals who participate in the first contest, but not the second. The bottom table restricts the data to be only individuals who participate in both contests. This helps clarify the number of observations actually used to identify an effect of prize differentials on performance using a difference in differences approach.

The lower half of Table 3 shows that 364 household have no spread during contest 1. The during contest 2, 102 of these households experience and increase in spread from 0 to 50 Soles while 48 of these households do not.

Estimation Strategy

According to theoretical predictions, I expect to find an increase in performance levels as the spread increases. To identify this effect, I use a difference in differences (DID) approach. The outcome variable is the same performance measures used to test the "N Effect" and the treatment in this case is an increase in the prize spread. The control group are the households who progress from Contest 1 to Contest 2 with no change in prize spread. The treatment group are those who experience an increase in prize spread from Contest 1 to Contest 2. The treatment effect is estimated by

$$P_{it} = \beta_0 + \beta_1 X_{it} + \beta_2 Contest_t + \beta_3 Spread_t + \delta_i (Contest_t \times Spread_t) + \varepsilon_{it}, \quad (15)$$

where X_{it} is a vector of geographic and community characteristics, *Contest* is a dummy variable indicating either the first or second contest. Prize amount would, in theory, need to be controlled for, but is collinear with the contest dummy and contest size variables. Thus, it was excluded from the specification. *Spread* is a dummy variable that indicates a change in the prize spread of 50 Peruvian Soles and the interaction of the spread and contest dummies identifies the individual treatment effect:

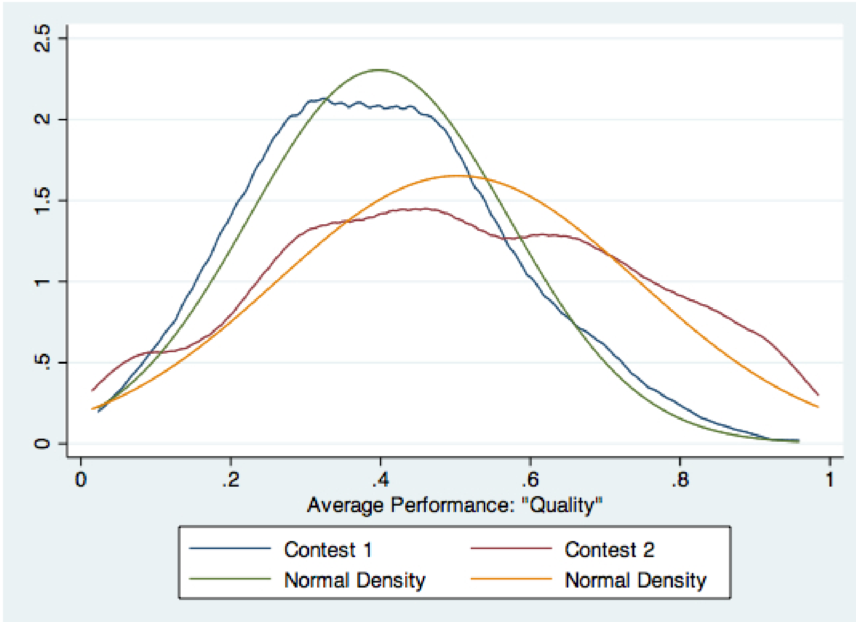
$$(\overline{P}^{Post}_{treat} - \overline{P}^{Pre}_{treat}) - (\overline{P}^{Post}_{control} - \overline{P}^{Pre}_{control}) = \delta_i.$$

Theory predicts that this treatment effect is positive, meaning that increasing the prize spread will have a positive effect on performance. For households on the margins of each prize category, having a larger spread equates to a higher benefit of moving from a lower position to a higher position. Thus, the marginal benefit of moving up one position increases while the marginal cost of effort remains the same.

The "Heterogeneity Effect" Hypothesis

This hypothesis predicts that, if effort levels are endogenous, then there should be a negative relationship between household ability and variability of performance. One of the main aspects that distinguishes the analysis of this hypothesis from the other two is the performance measure used. Since the data only cover two contests, using the final scores does not provide enough observations per household to measure variation of performance. To circumvent this problem, the scores from the monthly check-ins are used to measure performance variability. These scores are the average of the check-in scores received by each household during a given month. The check-in scores are simply a binary variable that indicates whether or not a household has made progress on a given activity. Therefore, a higher score indicates more effort put forth, which should imply better performance during the contest. Since these check-ins were conducted on a monthly basis, each household in the dataset has, on average, twice the number of observations for check-in scores than they do final scores, but this number can vary across households.

Figure 2: Distribution of Performance Measure for Contests 1 and 2



Since the performance measure in this case is an average, it is important to check the distributions of the measure in each contest to ensure that the average value is an appropriate estimate of average performance. Figure 2 shows that the performance measure is almost normally distributed for both contests in Ayuper, thus verifying the validity of its use in this context. This is because having normally distributed scores indicates that using the mean of scores are a proxy for performance, or effort, is representative of the average household.

An important assumption behind this hypothesis is that households are heterogeneous in ability. This results in an inefficient effort allocation by households; with heterogeneous contests, both low and high ability households tend to under invest effort. If this inefficiency is present in PMR contests, then incentive effects of the prizes are distorted and effort levels are sub-optimal. Therefore, the effort levels in contests can be increased by reducing heterogeneity. Theory posits that identifying and sorting households according to their ability can reduce the inefficiencies created by the heterogeneity of ability.[16] The PMR data has a unique structure that allows the study of both the heterogeneity effect and the implied sorting effect. During the first contest for any participating community, there is a certain degree of assumed heterogeneity due to the fact that everyone in a community is eligible to participate.[17] In the second contest, however, PMR automatically sorts the winners from the first contest into a separate "champions" league. Typically, only the first place finisher is sorted out of contest 2. It is important to note that sorting out the winners changes the contest on two levels. The first and most obvious level is that they are no longer competing against their neighbors.[18]

This leads to the second, less obvious effect on the mentality of the remaining households. In theory, since each household knows that the winner is no longer competing, there is incentive for some to change their strategy, seeing as the competition has become slightly more homogenous. The implication of this effect for the analysis is that sorting should change the strategies observed in the first contest due to the decrease in heterogeneity. However, this implication comes with caveats; it assumes that there was heterogeneity to begin with and it assumes that sorting out one household of higher ability reduces the amount of heterogeneity enough to spur a change in strategy for the remaining households.

[16] Backes-Gellner and Pool (2008) show this effect using organizational data from travel sales contests. The theoretical inefficiencies of heterogenous contests was first discussed by Lazear and Rosen (1980).

[17] From my observations in the field, assuming heterogeneity in this context is a reasonable assumption. Some might assume rural communities to be rather homogeneous due to the fact that everyone is "poor", but in this context, even in the poorest communities, heterogeneity can exist on several levels. Ability and knowledge are important factors that can vary within communities. Even small differences in income can create significant heterogeneity of wealth due to the fact that small increases in income can result in large economic differences.

[18] The champions league is basically a contest between the winners from each community, so they never actually see each other's progress and the competition takes place from a distance, as opposed to the normal contests where neighbors' actions can be affected by seeing each other's progress and potentially working together to complete activities.

It also assumes only one type of sorting; however, there may also be a sorting-in effect of households entering the second contests with higher ability levels. Their ability could be higher due to free-riding during the first contest, where non-participating households observe the strategies or accumulate resources to perform better during the second contest. This second type of sorting effect has the potential to negate any effect of sorting out the winners on heterogeneity. Table 4 describes the different sorting effects happening as communities transition from contest 1 to contest 2. In both regions, the group of households being sorted in after the first contest is considerably larger than the group being sorted out and has a higher mean ability. This implies that the composition of the contests is becoming more competitive and possibly more heterogenous despite the sorting-out of contest 1 winners. The contest sizes are smaller for each region than reported in Table 1. This is because the data are limited to households who have more than 1 observation of check-in scores, which is needed to generate the variability of performance measure.

Table 4: Hypothesis 3 Summary Statistics

Variable	All Households				Sorted Households			
	Contest 1		Contest 2		Sorted out		Sorted in	
	mean	*sd*	*mean*	*sd*	*mean*	*sd*	*mean*	*sd*
Performance Score	0.41	0.17	0.52	0.23	0.4	0.19	0.46	0.22
Variation of Performance	0.14	0.08	0.13	0.1	0.15	0.07	0.13	0.09
Average Effort (Quality)	26.14	9.04	30.03	10.54	18.38	8.52	29.37	11.7
Number of Households	521		889		55		423	
Number of Communities	46		69		16		59	
Standard Deviation Test (p-value)	0***		0***		0.0274**			
T Test (p-value)	0***		0***		.0019***			

Note: The above reports statistics for all households and for the two sorted groups of households. The third and fourth columns of the table shows that the sorted in group is not only much larger in size, but also has a higher average performance score. The standard deviation test is used to test the null hypothesis that the ratio of standard deviations of the performance score for each sorted group is equal to 1. The t-test uses the null hypothesis that the difference of the mean of performance between both groups is equal to zero. *,**,*** indicates significance at the .10, .05 and .01 levels, respectively. An additional test of standard deviation and mean was conducted on the variation of performance across contests, the resulting p-values were 0 and .0049 respectively, indicating that the distributions and means are statistically different.

Since heterogeneity can not be directly observed in this case, the results form the empirical specification will shed light on the relationship between household ability and their chosen strategies. It will also show what effect, if any, sorting-out has on these strategies.

Estimation Strategy

The empirical strategy used loosely follows that of Knoeber and Thurman (1994). I carry out a simplified version of their analysis, which is a two stage process. The first stage regresses the performance measure on various controls for each contest in order to obtain the residuals for performance.

$$P_{it} = \alpha + \beta_1 X_i + \beta_2 Contestsize_{it} + \beta_3 Month_i + \beta_4 Community_t + \varepsilon_{it} \quad (16)$$

Using the estimated residuals, from equation (20), I quantify variability of performance by taking the standard deviation of the residuals, $\hat{\sigma}_i$, for household i. This measure represents the amount of variability in performance for each household in a given contest.

To visualize the relationship between this variability and household "quality", I first create a variable that averages check-in scores for each household. This measure is average performance of each household for the whole contest and is meant to act as a proxy for household ability, or quality. I then plot the variability measure against the quality measure for each contest in order to see the relationship and test the significance of this relationship using

$$\hat{\sigma}_{it} = \beta_0 + \beta_1 \overline{P}_{it} + \varepsilon_{it} \quad (17)$$

Theory predicts that quality and variance of performance are negatively related, meaning that high quality households adopt low variance, or low risk, strategies (and low quality households do the opposite). This implies an inefficient allocation of effort levels and in order to increase effort, theory suggests that sorting out the heterogeneity will achieve more efficiency. Thus, in order to test both effects simultaneously, I separate the analysis by contest and plot variance of performance on quality of household. Additionally, I test the effect of sorting in new households after the first contest. To do this, I Identify the sorted-in and sorted-out households in the data and conduct the same analysis on each group to see both the strategies adopted and size of each group. The figures from contest 1 in each region will provide evidence for the heterogeneity effect, while the figures of the second contests will show the effects of sorting in new households and sorting out winning households.

7 Results

The following section outlines the results for each hypothesis. The results from hypothesis 1are ambiguous with respect to theoretical predictions, and depend on the assumption of the density of uncertainty. Hypotheses 2 and 3 both conform to theoretical predictions and lend support to the idea that contest can be an effective tool for development.

The "N Effect"

The results for this hypothesis are reported in Table 5. Column 1 shows the results of contest size on average performance during the first contest. The coefficient on contest size is positive, contrary to my initial predictions, but it is very small in magnitude and not statistically significant. The lack of clear results for contest 1 could be attributable to several factors, but one in particular is that households' distrust of PMR or their lack of understanding of the rules, etc. might be playing a role in the result I find.

Table 5: Hypothesis 1 Estimation Results

	(1) Contest 1	(2) Contest 2
Contest size	6.65	7.79*
	(7.24)	(3.56)
Contest size^2	-0.14	-0.04
	(0.18)	(0.07)
Contest size (no prize controls)	8.49	5.55
	(6.94)	(3.71)
Number of Communities	44	68
Number of Housholds	810	1509

Note: All regressions use the final performance score as the dependent variable and control for prize amount along with geographic region. The contest size variable on both specifications reports the general effect of contest size on performance while controlling for prize amount. The Contest size ^2 variable tests for any non-linearity in the relationship and Contest size is a separate regression with no prize controls. All prizes are in Peruvian Soles. Standard errors are clustered at the community level. *,**,*** represent significance at the .10, .05 and .01 levels, respectively.

Column 2 shows the results for the second contest and reports a positive and significant effect of contest size on performance. More specifically, it shows that for the second contest, increasing the contest size by 1 person increases performance by about 5 points, which represents an increase of about .05 of a standard deviation. This effect is relatively small in magnitude, but indicates a general increase in performance as the contest size increases.

This result conforms to theoretical predictions assuming the form of uncertainty to have an increasing density, but are contrary to the general expectations of behavior found in the literature so far. However, the result is logical considering the social context of the contests. Households are encouraged by PMR to learn from their peers and the inter-community contests provide another incentive to cooperate. Additionally, the contests, by design, allow for neighbors to benefit from the success of one another; if one household figures out an effective way to complete an activity, then the practice is likely to spread throughout the entire community. This implies that specialized knowledge is shared, and a larger team results in more productivity. Therefore, having more households contributing to the pool of common knowledge, results in higher performance levels.

The result, however, is not highly significant and only shows up in the second contest. This means that other factors might be mitigating the effect. It is possible that households have incorrect beliefs that more people participating means a higher prize amount (this is true, but only for brackets of contest sizes, some households may belief it increases with every additional household.) This is assumed to be more of an issue during the first contest due to the inherent learning that must take place for households to understand the rules of the contest. A similar effect could be taking place, albeit to a lesser degree, during the second contest. The squared term of contest size reports negative coefficients for both contests. This is the expected result, signifying that the relationship between contest size and performance is not linear, and eventually becomes decreasing as n increases.

Another explanation behind the lack of a highly statistically significant result is the potential of omitted variable bias. The dataset used for this analysis is limited by its lack of demographic information for households. In addition to demographic information, data on community size or proximity to local markets might be needed to control for their respective effects. Larger communities could be more wealthy and closer to roads, which could have an effect on performance.

The "Spread Effect"

Table 6 reports the results from the difference in differences estimation of the effect that increasing the prize spread has on performance. For this analysis, the control group is the group of households who did not experience an increase in spread across contests and the treatment group are all those who did. The Spread*Contest variable in Table 6 reports the main result, which indicates that increasing the spread between 4th and 5th prizes by 50 Peruvian Soles

(~$20) has a positive and significant effect on average performance.

Table 6: Hypothesis 2 Estimation results

	(1) Performance
1(Spread Increase)	-149.58***
	(22.37)
1(Contest 2)	22.37
	(8.93)
*1(Spread*Contest)*	111.31***
	(22.40)
Intercept	101.72***
	(19.96)
N	806

Note: Results are from a DID strategy that uses raw performance score as the dependent variable. A vector of geographic variables and contest size were also included in the regression.Prize amount is not specifically controlled for for due to collinearity with the contest dummy and contest size. *,**,*** represent .10,.05 and .01 significance levels, respectively.

More specifically, the 50 Soles can be attributed to an increase of 102 points to a households final score, on average. Considering the distribution of final scores over both contests, the effect is almost a full standard deviation increase in final score. The implication is that increasing the marginal benefit to exert effort (prize spread) leads to a significant increase in effort.[19] Table 7 reports the predictive margins from the above regression.

Table 7: Hypothesis 2 Predictive Margins of OLS results

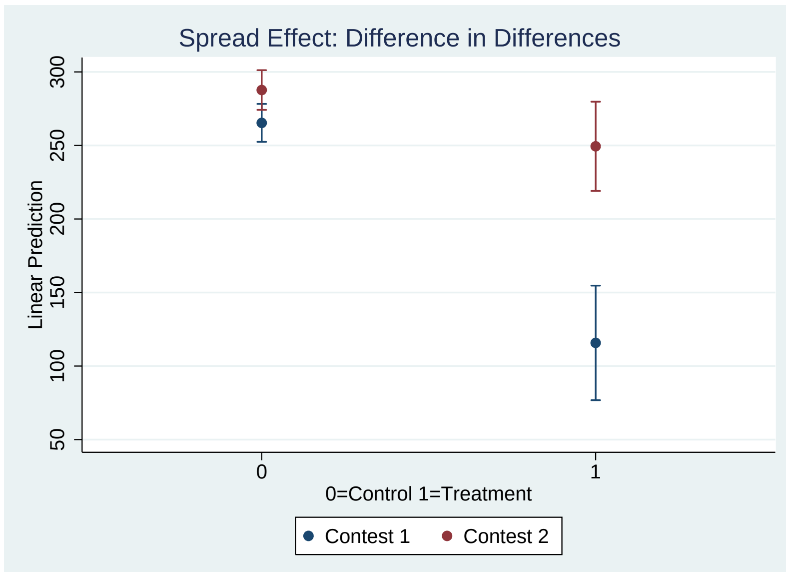
Spread Increase	Contest	Margin	Std. Err.	P-value
No	1	265.34	6.58	0
No	2	287.69	6.88	0
Yes	1	115.75	19.87	0
Yes	2	249.41	15.48	0

Note: The marginal effects reported in this table come from the previous DID regression.

[19]This result must be understood in the context that this analysis is conducted narrowly, using only the 4th-5th prize spread. It is unlikely that this one particular spread has such a large effect, but the data do not permit to study the effects of the other spreads due to lack of variation.

Figure 3 gives a visual depiction of the effect that increasing the prize spread has on performance. The graph give a visual explanation of the difference in difference approach. The treatment effect can be visualized by taking the differences of the points for each contest and subtracting the difference of the contest 1 points from the contest 2 points.

Figure 3: Difference-in-Differences Results



These results are in line with theoretical predictions and lend support to the idea that contests can create incentive for rural households to direct effort towards activities that could potentially be productive long run investments.

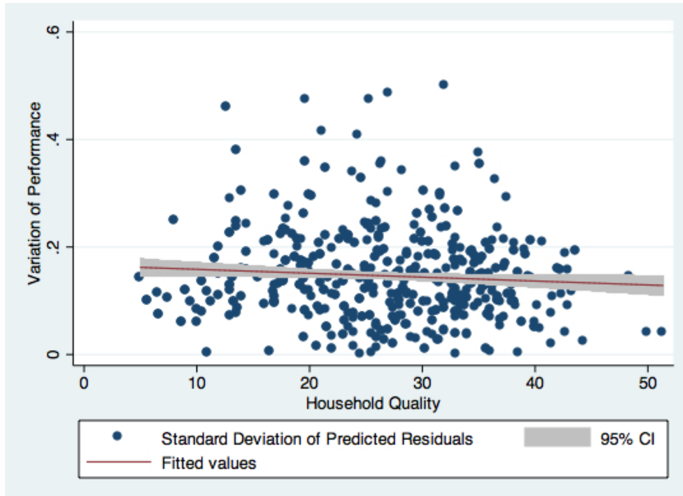
The "Heterogeneity Effect"

Figures 4-7 show the main results for this hypothesis. The second contest for each region excludes the first place winners from the first contest and includes new households who have not participated in contest 1. Theory predicts that Figure 4a should be negative, implying that the heterogeneity effect induces high ability households to adopt low-risk performance strategies and low ability households to do the opposite. Figure 4a shows a negative relationship between ability and variation of performance, validating theoretical predictions. As reported in Table 8, the pre-sorting coefficient shows that as quality is increased by 1 point (.11 standard deviation), the variance of performance will decrease by .0007 standard deviations. This means a standard deviation increase in quality results in a .006 standard deviation decrease in variation of performance.

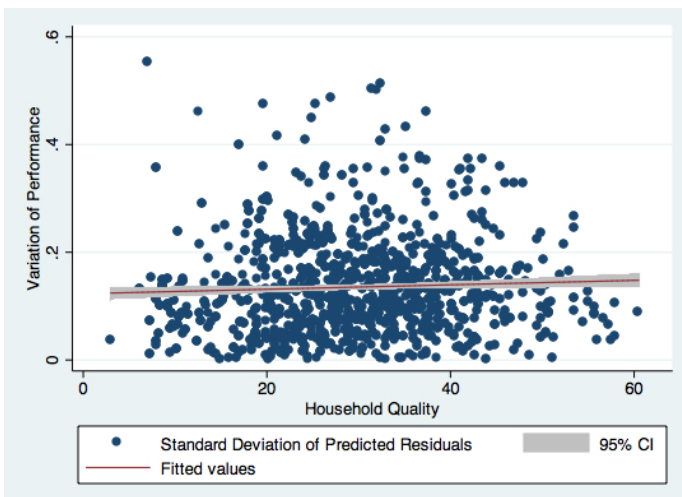
Figure 4b shows the same relationship during the second contest. If sorting winners into a separate league has any effect of heterogeneity, I would expect the slope of the relationship to change sign, or become flatter, indicating that sorting out the highest ability households reduces the heterogeneity effect of different risk strategies for different abilities. Figure 4b suggest that households in contest 2 are less likely to adopt a more or less risky strategy based on their ability. One potential problem with this observation is that while Figure 4b accounts for the sorting-out of winners, it does not depict the sorting-in of new households. As households move from Figure 4a to Figure 4b, there are 23 winners and 32 households who decide to not participate. This group of households constitutes the sorting-out group. The sorting-in group consists of 423 households who decide to join after the first contest.

Figure 4: Risk and Household Quality

(a) Contest 1 (Pre-Sorting)

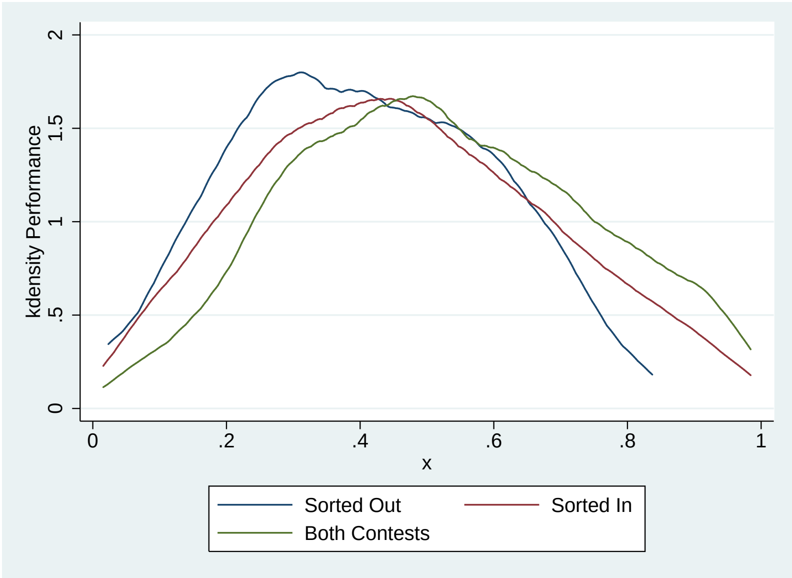


(b) Contest 2 (Post-Sorting)



With all the new participating households, the effect of sorting out the winners depends largely on the composition of the new group of households. If they are mostly of high ability, then heterogeneity can increase and mitigate any sorting-out effect. Thus, it is important to separate these two competing effects in order to identify the effectiveness of sorting out winners. Figure 5 illustrates the distribution of performance for the households in both the sorted-in, sorted-out and not-sorted groups.

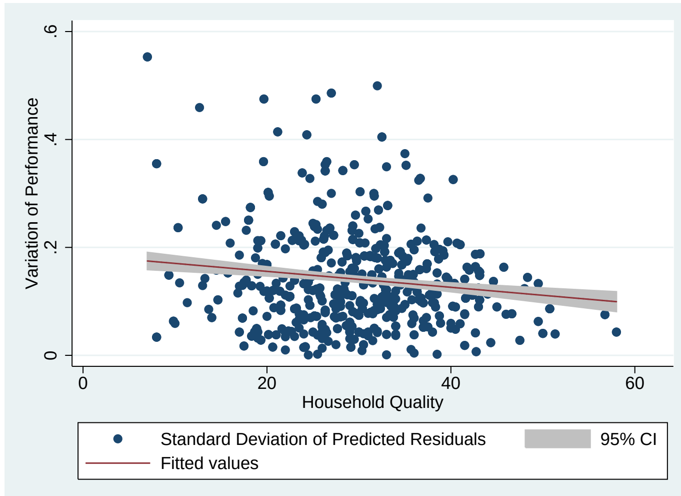
Figure 5: Distribution of Sorted In and Sorted Out Households



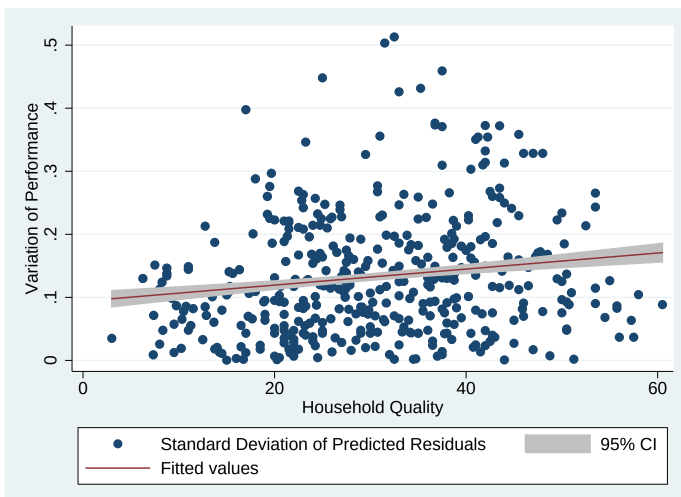
The figure suggests that the sorted-in households have a higher average performance than the sorted-out households. Simultaneous sorting in and out suggests that PMR's current policy of sorting out only the first place households does not "even the playing field" enough to compensate for incoming households of high ability. To see this more clearly, figures 6a and 6b decompose Figure 4b into two groups: households who participate in both contests and households who only participate in contest 2.

Figure 6: Decomposition of Figure 4b

(a) Households Participating in Contest 1 and Contest 2



(b) Households Sorted In After Contest 1



Separating the two effects shows the influence the incoming households have on the result seen in Figure 4b. Figure 6a shows that due to their higher average ability, the incoming households negate any effect sorting out the winners form contest 1 has on the strategies chosen by high and low ability households. Figure 6b shows the relationship for households who have previously participated in contest 1 remains negative, and the relationship becomes more pronounced. This suggests that the households in Figure 6b are driving the results seen in Figure 4b and the heterogeneity effect becomes more pronounced for households who participate in both contest 1 and contest 2 (Figure 6b).

Table 8: Hypothesis 3 Significance of Fitted Lines

	Fitted Lines for Figures 4 and 6			
	Pre-Sorting (Figure 4a)	Post-Sorting (Figure 4b)	Balanced Panel (Figure 6a)	Sorted In (Figure 6b)
Quality	-0.0007*	0.0004*	-0.0015***	0.0013***
	(0.00)	(0.00)	(-4.56)	(5.63)
N	521	889	466	423

Note: The dependent variable in all specifications is the standard deviation of the estimated residuals from equation (3). This is used as a proxy for variation of performance, or "risky strategy" and is regressed using OLS on average performance, which is a proxy for contestant quality. N shows the number of households for each column. *,**,*** represent .10, .05, and .01 significance levels, respectively.

8 Discussion of Results

The main goal of this research is to identify the behavior of households participating in PMR's contests and analyze whether or not their behavior is consistent with theoretical predictions. My results provide evidence that behavior is predicted well by hypotheses 2 and 3, but further investigation is need to clarify the ambiguous results of hypothesis 1.

The existing literature has mixed results regarding the first hypothesis; some have found a positive relationship, while others have found the opposite. My results help to clarify this ambiguity by providing evidence that in rural Peru, increasing the number of contestants tends to have a positive effect on performance. As noted by previous studies, this particular hypothesis is largely dependent on the context in which it is tested, as many factors can contribute to performance in increasing larger contests. The practical implications of the results suggest that PMR should aim to encourage larger contests and continue its philosophy of promoting cooperative behaviors on the community level.

My findings suggest that a considerable amount of cooperation is happening at the community level and this could be due to the incentive effects created by the inter-community contests or it could stem from PMR's philosophy of encouraging cooperation through peer learning. The benefits of increasing contest size, however, are not unlimited; understanding the optimal contest size in this context is an avenue for further research.

The results from the second hypothesis provide evidence that households are behaving as theory predicts, especially during the second contest. Having more predictable behaviors in the second contest is a sensible result seeing that households generally require the first contest to learn the rules and gain trust in PMR. As was noted in my field visits to these communities, many community members admitted a general distrust for non-profit organizations due to negative past experiences. The results show that after households learned the rules and were convinced that the prizes were legitimate, increasing the prize spread by 50 Peruvian Soles had a large effect on performance scores, increasing them by almost a full standard deviation on average. This implies that PMR should design prize structures to maximize effort by increasing the prize spread from contest to contest. Much like the first hypothesis, understanding the optimal prize spread is a topic for further research.

Lastly, the results from testing the third hypothesis suggest that households adopt different strategies based on their ability. High ability households tend to have lower variation of performance, indicating that they adopt a low-risk strategy and low ability households tend to have higher variation of performance, indicating they adopt high risk strategies. Theory predicts that the strategy choices resulting from the heterogeneity effect result in an overall under-investment of effort; high ability households invest less effort knowing that they already have a good chance of winning and low ability households under-invest effort thinking that since they are at an inherent disadvantage, investing valuable resources into a contest they can not win is futile. PMR attempts to reduce the disincentive created by the heterogeneity effect and sorts the winners into separate leagues, but the results suggest that their efforts are nullified by the sorting-in of new households after the first contest. In fact, the heterogeneity effect is more pronounced after the sorting out of winners due to the increased heterogeneity of the new households. To rectify the sorting-in effect, PMR should change the rules of their contests to separate a larger number of winners after each contest or identify a new mechanism that effectively separates households based on their ability.

Although the results show the presence of the heterogeneity effect, the relationship is noisy, suggesting that the general distribution of ability within communities is relatively uniform. Given the result from hypothesis 1, cooperation could be affecting the strategies adopted by some households. If communities tend to works as team, than the distribution of ability becomes more uniform as cooperation increases. Given the evidence from hypothesis 1 and the fact that PMR encourages community teamwork, this is a plausible explanation for the noisy relationship between quality and variation of performance and offers another potential solution to the inefficiency created by heterogeneity.

All three hypothesis lend support to the conclusion that contests have the potential to be effective tools for development in rural communities, but highlight the need for further research in order to identify the exact mechanisms through which the most effort can be incentivized. The results, however, must be understood as coming from an incomplete data set. A fundamental problem with this data set is the lack of controls and therefore, estimating the effect of contest design on behavior could potentially be affected by covariates that are not being controlled for. Additionally, the dataset suffers from a lack of variation in key variables. Hypothesis 1, for example, would ideally have many more communities and a larger amount of variation of contest size within each prize category. This would result in a clearer picture and help to clarify the ambiguous results. Hypothesis 2 is also likely to suffer from the limitations of the data. The estimated effect of increasing the prize spread is likely to be overstated due to the lack of controls and variation. Ideally, I would be able to test every prize spread, but in this setting was not able to. Hypothesis 3 was not as constrained by the dataset, but still suffers from a lack of controls when estimating the first stage predicted residuals. In general, some controls needed for future research would be education, income, proximity to local markets or roads, and knowledge of the social composition of communities. Knowing how many family members a household has in the community, as well as other demographic information would help clarify the results in this setting. Thus, the results from this research must be understood by taking these limitations into account.

9 Tables

Table 2: Number of Communities For Each Contest Size and Prize Amount

Contest Size	Number of Communities in Each Prize Category				
	350	400	450	500	600
8	2	0	2	0	0
9	0	0	1	0	0
10	2	0	0	0	0
11	0	0	1	0	0
12	2	0	4	0	0
13	1	0	2	0	0
14	3	0	4	0	0
15	7	0	6	0	0
16	6	0	5	0	0
17	2	0	3	0	0
18	1	0	4	0	0
19	0	0	3	0	0
20	3	0	5	0	0
21	4	0	1	0	0
22	2	0	2	0	0
23	1	0	4	0	0
25	0	0	3	0	0
26	5	0	1	0	0
27	0	0	2	0	0
29	0	0	1	0	0
30	0	0	1	0	0
31	0	0	0	2	0
32	0	2	0	0	0
33	0	0	0	1	0
35	0	0	0	1	0
37	0	1	0	2	0
38	0	0	0	3	0
46	0	0	0	0	1
48	0	0	0	0	1
53	0	0	0	0	1
57	0	0	0	0	1
Number of Communities	41	3	55	9	4
Number of Contestants	709	101	987	318	204

Note: Each column lists the number of communities for each prize amount. Each row represents the contest size reported in the left column. Thus the table shows the number communities and their contest size for each prize category. The bottom two rows report the total number of communities for a given prize category and the implied number of contests for each prize category.

Table 1: Summary Statistics

Variable	AYUPER											
	Contest 1				Contest 2				Contests 1 and 2			
Number of households	1410				2280				881			
Number of Communities	54				72				48			
Number of Regions	3				3				3			
	Mean	SD	Min	Max	Mean	SD	Min	Max	Mean	SD	Min	Max
Children per household	3.69	2.16	0	13	3.63	2.21	0	18	3.9	2.2	1	12
First prize	445.8	70.4	350	500	567.5	59.4	450	600	412.4	62.9	350	600
Second prize	345.8	70.4	250	400	463.8	60.4	350	500	310.2	57.9	250	400
Third prize	264.6	46.6	200	300	277.5	38.9	200	300	210.2	20.2	200	250
Fourth prize	195.8	70.4	100	250	215.7	59.2	100	250	111.3	23.4	100	200
Fifth prize	131.1	24.2	100	150	136.4	22.3	100	150	100	0	100	100
Sixth prize	81.1	24.2	50	100	86.4	22.3	50	100	50	0	50	50
Seventh prize	50	0	50	50	50	0	50	50	50	0	50	50
Contest size	23.3	8.7	4	48	26.0	11.4	1	57	22.6	8.4	8	48
1^{st} 2^{nd} Spread	100	0	100	100	103.7	18.8	100	200	102.2	14.6	100	200
2^{nd} 3^{rd} Spread	81.1	24.2	50	100	186.4	22.3	150	200	100	50	50	150
3^{rd} 4^{th} Spread	68.9	24.2	50	100	61.8	21.2	50	100	98.9	7.3	50	100
4^{th} 5^{th} Spread	64.6	46.6	0	100	79.3	38.7	0	100	11.3	23.4	0	100
5^{th} 6^{th} Spread	50	0	50	50	50	0	50	50	50	0	50	50
6^{th} 7^{th} Spread	50	0	50	50	47.6	10.7	0	50	0	0	0	0
Total Spread	395.8	70.4	300	450	517.5	59.4	400	550	362.4	62.9	300	550
Raw Check-in Score	0.4	0.2	0.02	0.96	0.5	0.2	0.02	0.98	0.52	0.22	0.02	0.98
Raw Final Score	227	109.7	2	618	270.6	126.9	9	645	263.3	121.22	2	645

Note: All prize variables represent the corresponding prize amount and are measured in Peruvian Soles (1 Sol=$.40USD.) Contest size measures the number of individuals participating in a contest at a given time. The spread variables are constructed as the difference between each of the above prize variables and the total spread is the difference between the first and last prize amounts.Raw check-in score is only used as a performance measure in hypothesis 3 and is the summation of all 1's received during a monthly check-in. Raw final score is the main performance measure used and is a summation of all the points earned by an individual during the final scoring round in a given contest. The third column shows the summary statistics for households who participated in both contests. This group of households will be used for the analysis of Hypothesis 3. Raw Check-in score is not reported due to the fact that the statistics are combined from both contests.

References

Backes-Gellner, U. and K. Paul (2008, January). Tournament incentives and contestant heterogeneity: Empirical evidence from the organizational practice. Tubingen Working Paper.

Bank, W. (2000, September). *World Bank Development Report: 2000/2001*. Oxford University Press.

Becker, B. E. and M. A. Huselid (1992). The incentive effects of tournament compensation systems. *Administrative Science Quarterly 37*, 336–350.

Bertrand, M., S. Mullainathan, and E. Shafir (2006, Spring). Behavioral economics and marketing in aid of decision making among the poor. *American Marketing Association 25*(1), 8–23.

Boudreau, K. J., C. E. Helfat, K. R. Lakhani, and M. Menietti (2012, August). Field evidence on individual behavior and performance in rank-order tournaments. Harvard Business School Working Paper.

Bull, C., A. Schotter, and K. Weigelt (1987, February). Tournaments and piece rates: An experimental study. *Journal of Political Economy 95*(1), 1–33.

Casas-Acre, P. and F. A. Martinez-Jerez (2009, August). Relative performance compensation, contests, and dynamic incentives. *Management Science 55*(8), 1306–1320.

Datta, S. and S. Mullainathan (2012, November). Behavioral design: A new approach to development policy. CDG Policy Paper 016.

Dechanaux, E., D. Kovenock, and R. M. Sheremeta (2008, September). A survey of experimental research on contests, all-pay auctions and tournaments. Working Paper.

Delfgaauw, J., R. Dur, A. Non, and W. Verbeke (2012, July). The effects of prize spread and noise in elimination tournaments: A natural field experiment. Tinbergen Institute Discussion Paper.

Ehrenberg, R. G. and M. L. Boganno (1990, February). The incentive effects of tournaments revisited: Evidence from the european pga tour. *Industrial and Labor Relations Review 43*(3), 74s–88s.

Eriksson, T. (1999, April). Executive compensation and tournament theory: Empirical tests on danish data. *Journal of Labor Economics 17*(2), 262–280.

Falk, A., E. Fehr, and D. Huffman (2008, March). The power and limits of tournament theory. Working Paper.

Frick, B. and B. R. Humphreys (2011, August). Prize structure and performance: Evidence from nascar. University of Alberta Working Paper.

Gibbs, M. (1990). Promotions and incentives. *Boston: Division of research Harvard Business School. Working Paper*(90-064).

Green, J. R. and N. L. Stokey (1983). A comparison of tournaments and contracts. *Journal of Political Economy 91*(3), 349–364.

Grund, C., J. Hocker, and S. Zimmerman (2010, March). Risk taking behavior in tournaments: Evidence from the nba. IZA Discussion Papper No.4812.

Harbring, C. and B. Irlenbusch (2003). An experimental study on tournament design. *Labour Economics* (10), 443–464.

Hvide, H. K. (2002, October). Tournament rewards and risk taking. *Journal of Labor Economics 20*(4), 877–898.

Knoeber, C. R. and W. N. Thurman (1994, April). Testing the theory of tournaments: An empirical analysis of broiler production. *Journal of Labor Economics 12*(2), 155–179.

Kosfeld, M. and S. Neckerman (2011, August). Getting more work for nothing? symbolic awards and worker performance. *American Economic Journal: Microeconomics 3*, 86–99.

Krakel, M. and D. Sliwka (2002, August). Risk-taking in asymmetric tournaments. University of Bonn Working Paper.

Lazear, E. P. and P. Oyer (2007, October). Personnel economics. NBER Working Paper 13480.

Lazear, E. P. and S. Rosen (1981, October). Rank-order tournaments as optimum labor contracts. *The Journal of Political Economy 89*(5), 841–864.

List, J., D. van Soest, J. Stoop, and H. Zhou (2010). On the role of group size in tournaments: Theory and evidence from lab and field experiments. Working Paper.

Nalebuff, B. J. and J. E. Stiglitz (1983, Spring). Prizes and incentives: Towards a general theory of compensation and competition. *The Bell Journal of Economics 14*(1), 21–43.

Nieken, P. and D. Sliwka (2010). Risk-taking tournaments-theory and experimental evidence. *Journal of Economic Psychology 31*, 254–268.

Orrison, A., A. Schotter, and K. Weigelt (2004, February). Multiperson tournaments: An experimental examination. *Management Science 50*(2), 268–279.

Riis, C. (2010, Fall). Efficient contests. *Journal of Economics and Managment Strategy 19*(3), 643–665.

van Dijkand Joep Sonnenmns, F. and F. van Winden (2001). Incentive systems in a real effort experiment. *European Economic Review 45*, 187–214.

10 Appendix

Ayuper Prize Structure: Contest 1

Number of Participating Families	First Prize	Second Prize	Third Prize	Fourth Prize	Fifth Prize	Sixth Prize	Seventh Prize
15 to 30	S/. 350	S/. 250	S/. 200	S/. 100	S/. 100	S/. 50	-
31 to 45	S/. 400	S/. 300	S/. 250	S/. 150	S/. 100	S/. 50	-
46 to 60	S/. 500	S/. 300	S/. 250	S/. 200	S/. 100	S/. 50	S/. 50
61 and up	S/. 500	S/. 400	S/. 300	S/. 250	S/. 150	S/. 100	S/. 50

Ayuper Prize Structure: Contest 2

Number of Participating Families	First Prize	Second Prize	Third Prize	Fourth Prize	Fifth Prize	Sixth Prize	Seventh Prize
15 to 30	S/. 450	S/. 350	S/. 200	S/. 100	S/. 100	S/. 50	-
31 to 45	S/. 500	S/. 400	S/. 250	S/. 150	S/. 100	S/. 50	-
46 to 60	S/. 600	S/. 400	S/. 250	S/. 200	S/. 100	S/. 50	S/. 50
61 and up	S/. 600	S/. 500	S/. 300	S/. 250	S/. 150	S/. 100	S/. 50

Ayuper Contest Activities

operation and maintenance of irrigation system	stove and chimney (no smoke in the house)
organic planters for fruit and vegetables	state and maintenance of washrooms
worm fertilizer for produce and trout raising	construction and maintenance of a latrine or bathroom
construction and maintenance of terraces and planters	general order and cleanliness of house
conservation and selection of seeds	organization of production
organization plan for pasture	organization of sales
animal holdings: quality and organization	presentation of family (order, clothing, health and hygiene)
planting of grasses (native and exotic) in weeded areas	quality and quantity of stored food
straw and feed (enough to feed all animals)	vaccination cards for the whole family
construction, quality and cleanliness of stables/feeding areas/etc	no parasites in domestic animals
selection, separation and castration of animals	art in the house, paint, furniture, musics, clothes, etc.
Yogurt, cheese and butter production/quality	separate sleeping quarters for boys and girls
Storage and use of guano	quality of furniture and beds
elimination of internal and external parasites in livestock	fences to make sure animals don’t enter the house/kitchen
Quality and cleanliness of stables, and other areas for livestock	areas to receive visitors/tourists
Documentation and control of breeding in guinea pigs and birds	irrigation for plants
Drawing of forestation plan	plants free of damages/problems
family greenhouse (quantity of plants, quality of installation)	weeding land that wasn’t in use before and rotations crops
variety of plants in greenhouse	health and weight of animals
use of biological/organic pesticides	notebook with information tracking livestock
quantity and handling of grafted plants	control and quality of reproducers
storage of products (artisan, dairy, food, etc.)	sufficient reserves of food for animals
Quality of products and services	tools and equipment, including a first-aid kit
presentation of products and services	weight and health of birds and guinea pigs
administration of business	availability and preparation of food for birds/guinea pigs
participating in community association	notebook tracking birds/guinea pigs
preparation to receive tourists	quality of the reproducing males
availability and handling of potable water	handling and cleanliness of plantations
card tracking height and weight of children	number of new plants in field
card tracking pregnancy/family planning	advances in forestation
ridding children and adults of parasites	use of biocides and organic fertilizers
drawing of present and future	quality of floors and ceilings/roofs
plaster interior and exterior walls	