## Appendix A: Additional Figures and Tables

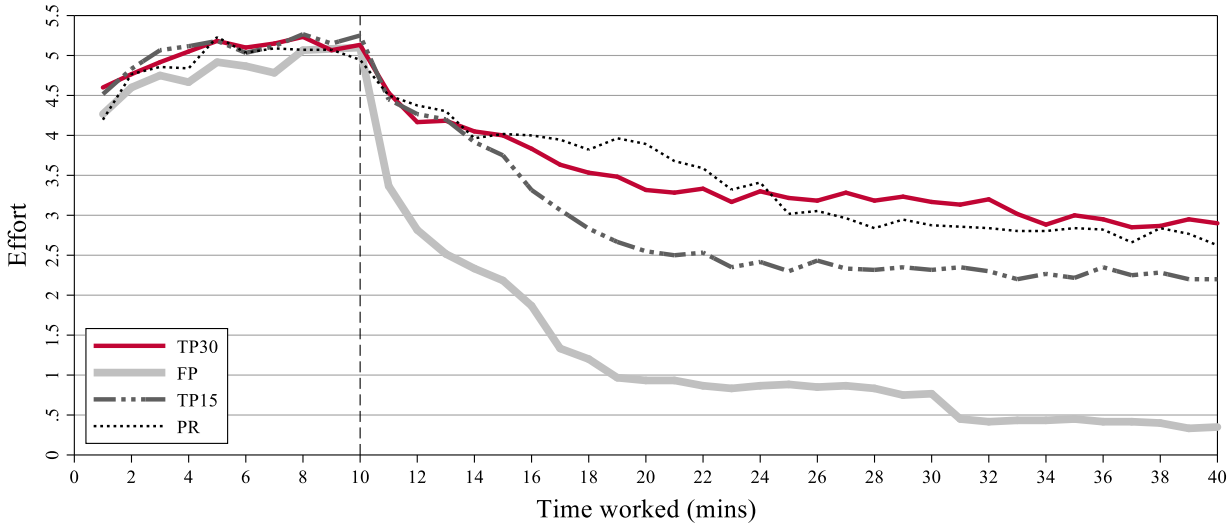



**Figure A1: Total Room Effort per Minute – All Treatments**

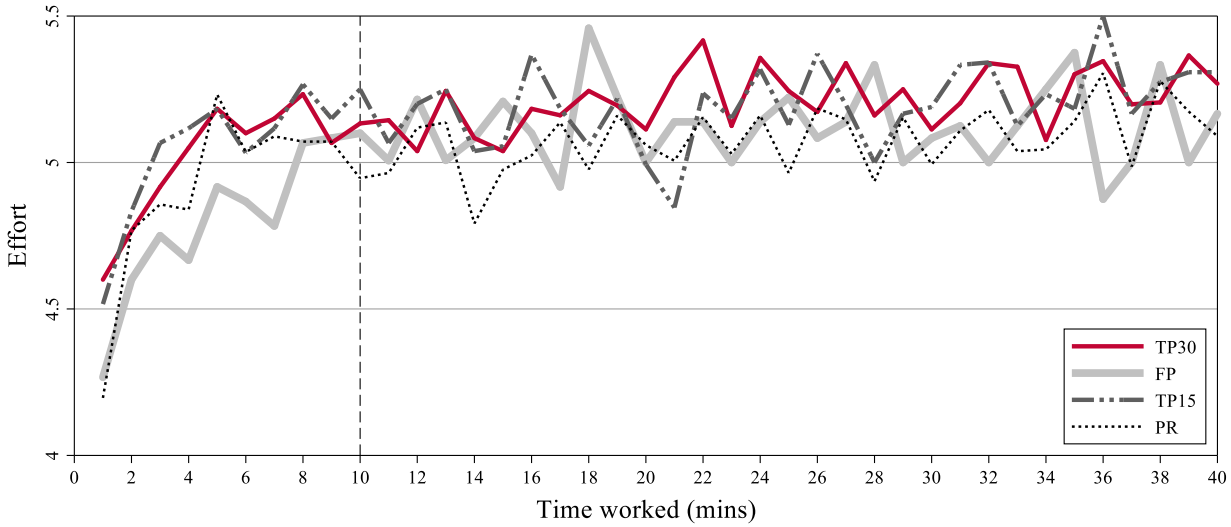



**Figure A2: Mean Room Effort Conditional on Working per Minute – All Treatments**

**Appendix Table A1: Balance Tests**

| | (1) | (2) | (3) |
|---|---|---|---|
| Sample | All | Male | Female |
| GPA | 0.166 | 0.518 | 0.245 |
| GPA-Squared | 0.172 | 0.508 | 0.254 |
| Economics or Commerce Major | 0.503 | 0.298 | 0.988 |
| Engineering or Science Major | 0.931 | 0.703 | 0.992 |
| Other Major | 0.722 | 0.820 | 0.845 |
| White | 0.940 | 0.789 | 0.612 |
| Asian | 0.948 | 0.828 | 0.601 |
| Other Race | 0.591 | 0.882 | 0.693 |
| First Year | 0.109 | 0.173 | 0.0208 |
| Second Year | 0.179 | 0.412 | 0.279 |
| Third Year | 0.00539 | 0.0514 | 0.215 |
| Fourth Year | 0.662 | 0.403 | 0.473 |
| Fraction of Screens with a Gold Square | 0.630 | 0.374 | 0.384 |
| Fraction with Gold Square Squared | 0.769 | 0.460 | 0.343 |

*Notes*: Table reports *p*-values from F-tests on the joint significance of the treatment indicator variables (FP, TP15 and PR) following separate regressions on each of the outcomes listed in the rows. Column 1 uses all workers; column 2 is limited to men; and column 3 is limited to women.

**Appendix Table A2: Intensive Margin Effort Measures by Gender**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | Male | | Female | | All | |
| | Mean Effort | Effort in the First 10 Minutes | Mean Effort | Effort in the First 10 Minutes | Mean Effort | Effort in the First 10 Minutes |
| FP | -0.227 | -0.310* | -0.168** | -0.110 | -0.227 | -0.310* |
| | [0.138] | [0.180] | [0.081] | [0.133] | [0.139] | [0.181] |
| TP15 | -0.021 | 0.010 | -0.004 | 0.057 | -0.021 | 0.010 |
| | [0.104] | [0.110] | [0.085] | [0.144] | [0.104] | [0.110] |
| PR | -0.063 | -0.228 | -0.099* | 0.009 | -0.063 | -0.228 |
| | [0.115] | [0.187] | [0.059] | [0.112] | [0.116] | [0.188] |
| Female x TP30 | | | | | 0.020 | -0.080 |
| | | | | | [0.074] | [0.099] |
| Female x FP | | | | | 0.080 | 0.120 |
| | | | | | [0.126] | [0.184] |
| Female x TP15 | | | | | 0.038 | -0.033 |
| | | | | | [0.075] | [0.093] |
| Female x PR | | | | | -0.017 | 0.157 |
| | | | | | [0.079] | [0.154] |
| Constant | 5.093*** | 5.060*** | 5.113*** | 4.980*** | 5.093*** | 5.060*** |
| | [0.090] | [0.099] | [0.049] | [0.103] | [0.090] | [0.100] |
| Prob>F: TP15 = FP | 0.085 | 0.047 | 0.088 | 0.207 | 0.085 | 0.047 |
| Prob>F: PR = FP | 0.201 | 0.709 | 0.344 | 0.211 | 0.202 | 0.709 |
| Prob>F: TP15 = PR | 0.644 | 0.156 | 0.217 | 0.665 | 0.645 | 0.157 |
| Observations | 118 | 118 | 118 | 118 | 236 | 236 |
| $R^2$ | 0.046 | 0.055 | 0.050 | 0.017 | 0.049 | 0.042 |

*Notes*: Observations at the worker level. Mean effort is total effort in the session divided by minutes worked. Robust standard errors clustered at the room level in brackets. *** p < 0.01, ** p < 0.05, * p < 0.1

## Appendix B: Field Experiment Materials

### B.1 Recruitment email

Subject: Hiring for Paid Short-Term Research Assistance Positions in the Upcoming Weeks

Dear students,

I am looking to hire a large number of students ASAP to help me test and benchmark a computer program that will be used in research to measure economic preferences. I would like to learn about user experience and performance under different parameters of the program.

Students invited to work on the project will be paid $25 for testing the program and responding to a questionnaire about the work. All work will be done using tablet computers that I will provide during the work session. Plan on being available for an hour. Payments will be made via PayPal within 2 days after the work session.

Work on the project will begin on Monday, February 15, 2016 and will take place over the next couple of weeks.

If you are interested, please provide information about yourself and your availability as soon as possible through this online form: <survey link>

The deadline to complete the online form is Wednesday, February 10, 2016 @ 12

Thanks!
Professor X

## B.2 Sign-up Survey

### Sign-up Page for Professor X's RA Positions

Thank you for your interest in Professor X's program testing and benchmarking short-term research assistant positions. To apply for a position, complete the following form to provide information about yourself and your availability.

Students invited to work on the project will be paid $25 for testing the program and responding to a questionnaire about the work. All work will be done using tablet computers that will be provided during the work session. Plan on being available for an hour. Payments will be made via PayPal within 2 days after the work session.

If you have any questions, please contact Professor X by email at professorx@abc.edu with the subject line "Program testing RA Position Question."

Personal Information

1. First name
2. Last name
3. Paypal email (We will pay assistants via PayPal. Please enter the email you use for PayPal below.)
4. Contact email (Please enter the email you want us to use to contact you here.)
5. Phone number
6. Sex (male/female)
7. Birth date (mm/dd/yyyy)

Academic Information

1. University ID number
2. Year in school
3. Major
4. GPA at university
5. Is there any other information you want to provide about yourself? If so, use the space below
6. Availability – please indicate all time slots when you could be available to work (M-F, 9:30-10:30am, 11am-12pm, 12:30-1:30pm, 2-3pm, 3:30-4:30pm)
7. If you are selected for this task, the data that you enter, as well as de-identified data about your work performance and compensation, may be used for research by the professor and coauthors. Check this box to indicate that you consent to have your data used in these ways. <click box>

**B.3 Invitation to work session email**

Subject: Program Testing RA Position

Dear <Name>,

You have been selected by Professor X as one of her RAs. Please come to Library Room <number> on <date> at <time> to perform the work**.**

You will be working in groups, so please make sure that you arrive a few minutes earlier so that we can start on time and you do not delay your colleagues.

Please make sure that your Paypal account is linked to this e-mail address. Otherwise, I will not be able to pay you.

**You were assigned this day and time based on your availability.** Please fill out the following form by 5pm on Thursday, February 11 to confirm you will be able to work at the assigned date and time.

<survey link>

If we do not hear from you by then, we will assume you cannot work and will find a replacement.

Thanks,
Research assistant

**B.4 Work session instructions (read out loud by research assistant)**

Script for Program Testing and Benchmarking RA Position Intro

<As students arrive>
Thanks for coming. Please sit at a tablet computer that you will use for the job.

<When everyone is seated>
I am XX, a PhD student in the department, working with Professor X on this project. I have some information and instructions that Professor X wants me to provide you about the position and the task before you get started.

The main task that you are going to complete is to play a simple game on a tablet computer. In the game, your job is to earn as many points as you can by clicking squares that appear on the screen. Once you log in to the game, it will alternate between "active screens" that have squares and "rest screens" with no squares. For this group, the active screen will last for at most 10 seconds. If you don't click a square during that time, the program will go to a rest screen. If you click on a red square before the 10 seconds are up, you earn 1 point and you are given the option to "fast forward" to the start of the next rest screen. Also, with a 10% probability (so on average in 1 out of every 10 screens), you will also see a gold square. You get 5 points for clicking the gold square. You can also click the red square in the same active screen.

As was explained in the recruitment e-mail, Professor X wants to use this program in her research. For that she needs to get a benchmark of the distribution of response times and how people score under different parameter conditions and time lengths. **So, please try your best.**

She also wants to learn about how the program functions and the user experience. Therefore, once you are done with the task, the program will ask you to complete a short questionnaire about your work experience today.

While Professor X asked you all to be available for a full hour, she thought that it might be too taxing to do this task for so long. This is another thing that she is trying to figure out. Therefore, while she would like you to stay for as long as you can, in order to get paid you only need to perform the task for at least 10 minutes and answer the questionnaire about the task. Once you have finished the questionnaire, you will be free to leave. This is how it will work. After 10 minutes of testing, a button will appear in the upper right-hand corner of the screen. You can click that button to end the program testing and move to the questionnaire. If you choose not to move to the questionnaire, after 40 minutes, the program will automatically end the program testing and direct you to the questionnaire. So, you will need to work for at least 10 minutes and are free to stay and work up to 40 minutes. **Please stay as long as you can.**

**<For Tournament treatment:>** To provide you with additional incentives to try your best, Professor X decided to pay a bonus in your group. The person that gets the most points, will receive [$30, $15] in addition to the promised $25, for a total of [$55, $40]. In case of a tie, the winner will be chosen randomly from among those who have the most points. This will be done with a roll of a die.

**<For Piece Rate treatment:>** To provide you with additional incentives to try your best, Professor X decided to pay a bonus in your group. You will be paid additional money for every point that you earn, with an exchange rate of 10 cents for every 3 points. You will be paid a bonus of 3 and one third cents for every single point you earn. This bonus payment is in addition to the promised $25, so your total payment will be $25 plus whatever bonus you earn.

**ANY QUESTIONS?**
<If there are no questions move to program>

Please login to the program. **Please use the same PayPal e-mail address you gave Professor X for the registration form.** She will not be able to pay you if you use a different email address.

Once you login you will see the parameters of the program for this group. Please wait on this page while I complete the instructions. Does everyone see the information about parameters? I will go around to make sure that there are no problems.

In case you encounter problems with the program please text Research assistant at the number written on this sheet of paper <Indicate paper>.

Once you start the program, you must not exit or minimize the program or turn off the tablet for any reason. If you do that, you may not be able to complete the task and you will not get paid.

Once you are done with the questionnaire, and submit your answers, you are free to leave. Please leave your tablet on the table. If you are the last person in the room to leave, please text Research assistant so he can come collect the tablets.

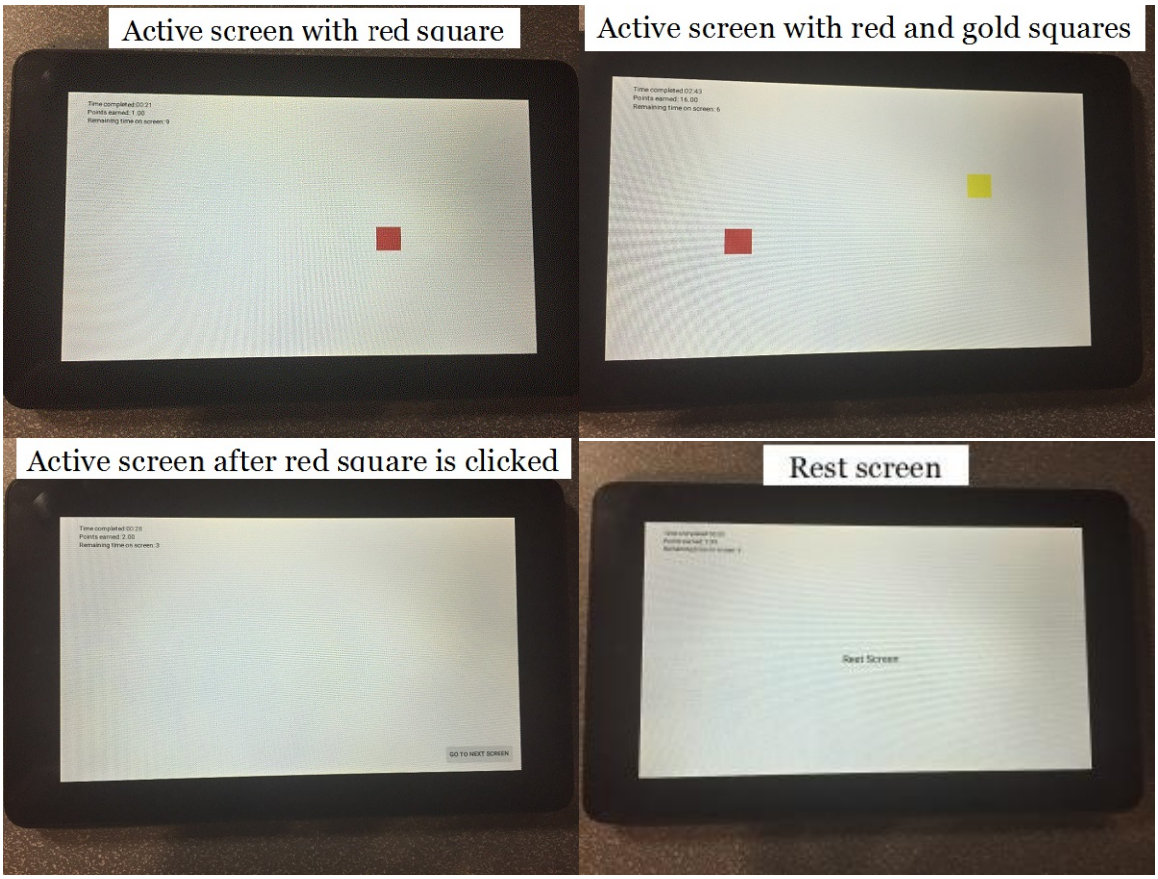
You will receive your payment on your Paypal account. If there are any problems with payments, please contact Professor X.

**ANY QUESTIONS?**

<RA goes around to check that everyone had the parameter screen up and is ready to start the task>

Please press the "**Start Task"** button to begin.

## B.5 Red Square program – screen shots



**B.6 Post work session questionnaire**

1. How much did you enjoy the game? (0-not at all, ..., 10-very much)
2. How exciting did you find the game? (0-not at all, ..., 10-very much)
3. Did you have enough time for the active screen? (0-not at all, ..., 10-very much)
4. Did you have enough time for the rest screen? (0-not at all, ..., 10-very much)
5. How tiring did you find the game? (0-not at all, ..., 10-very much)
6. Do you have any suggestions for improvements in the game design? (open ended)
7. Why did you decide to leave when you did? (open ended)
8. Would you be willing to come back for another trial under the same conditions? (yes, no)
9. How was the work environment? Mark all that apply. (Friendly, Quiet, Relaxing, Meditative, Boring, Stressful, Painful, Exciting)
10. Do you have any suggestions for improvement in terms of the work environment? (open ended)
11. Did you experience any technical problems (tablet/stands/program)? (open ended)
12. Do you have any suggestions for improvement in terms of technical problems? (open ended)
13. Are you happy with the payment scheme? (open ended)

## Appendix C. Detailed Comparison of Main and Auxiliary Treatments

The results in Section 5 show very similar overall effects on labor supply and costs from the alternative bonus schemes. This appendix compares the different treatments in greater detail.

### C1. Comparison between TP30 and TP15

It is natural to expect effort to be weakly lower in TP15 relative to TP30. The increase in prize value between TP15 and TP30 suggests an increase in labor supply. However, as we discussed in the main text, the number of competitors is not fixed. Thus, if multiple workers increase their effort in response to the increase in prize, then the probability of winning for each of them is reduced, and that would have a mitigating effect.[1] Furthermore, if a significant fraction of workers is at a corner (i.e., given the number of actual competitors in their room, they were willing to work even longer than 40 minutes for the $30 bonus), then this would also mitigate the effect of the prize on labor supply.

We do find lower effort (total and time worked) supplied in TP15 relative to TP30, but the estimates are not significant (Table 1, columns 1 and 2). In non-parametric tests, we also find fewer workers worked the full time in TP15 than in TP30 (41.67% versus 55%) and more rooms in which all workers left before 40 minutes (4 versus 2), but again, these differences are statistically insignificant.

In a supplemental analysis reported in Appendix Table C1 below, we find statistically significant differences in work time between TP15 and TP30 when we control for random variation in "luck" across workers in the experiment. The probability of being offered a gold square was 10% for each screen, but the realized share of gold squares offered varied across workers. This variation was not systematic across treatments (see Appendix Table A1), but even random noise would reduce the precision of our main estimates if luck affects labor supply. For

[1] It may be worth noting that the spillover effect from competitors' additional effort in TP30 can also *increase* the returns to effort for some workers and cause them to increase effort more between TP15 and TP30 than they otherwise would have. An example can illustrate this. Consider a room with 2 workers with high effort costs who quit immediately in both TP15 and TP30, a 3rd worker with moderately high effort costs who quits immediately in TP15 but is willing to stay 40 minutes for a ½ chance of winning TP30, and a 4th worker with low effort costs who is willing to stay 40 minutes for a ½ chance of winning TP15 or TP30. Equilibrium in TP15 can have all the workers leaving early, including the one willing to stay 40 minutes (because she wins soon after the 3rd worker leaves). A switch to TP30 could produce an equilibrium in which effort is still low for the 2 high cost workers, but both the 3rd and 4th workers stay the full 40 minutes. In this case, both the 3rd and 4th workers increase effort *more* in TP30 when the equilibrium changes than they would have if competitor effort had stayed the same.

tournaments, it is plausible that luck will indeed affect labor supply because it affects the chances of winning the prize, and therefore, the returns to effort. Unlucky workers who have very low chances of winning may decide to leave earlier than luckier workers. But very lucky workers might expect to win with certainty, which might induce them to stop working earlier. We therefore expect a concave relationship between workers' luck (i.e., the fraction of screens displayed that included a gold square) and labor supply.[2] The estimates in columns 1 and 2 of Appendix Table C1 confirm this prediction using an expanded version of the regressions of Table 1 (columns 1 and 2) with controls for the average of the fraction of gold squares (and square of that fraction) offered to workers in the room. We also confirm this at the worker level in the last two columns of Appendix Table C1, clustering standard errors at the room level. Across all of these models, controlling for luck has the dual effect of increasing the point estimate for TP15 and decreasing its variance, rendering the estimated lower labor supply in TP15 (compared to TP30) statistically significant.[3]

These results, together with the imprecise estimates from the main model support the interpretation that effort is increased by higher tournament prizes. Theory is less clear about the effects of prize amount on labor costs. Even if TP15 induced slightly lower effort than TP30, it was still possible for labor costs to be lower because of the smaller prize amount. That was not the case. All the coefficients on TP15 in Table 2 are positive and insignificant, indicating insignificantly higher costs relative to TP30.

**C.2 Comparison between TP30 and PR**

We set the PR amount to match average employer costs per point in TP30, but this is not the same as matching the marginal benefit of effort for workers in the two treatments. In particular, workers' payoffs in TP30, but not in PR, depend on the behavior of their competitors. For workers in TP30 who expect to face 3 other competitors for the full 40 minutes, the expected bonus payment for working the full time is close to the expected value of the incremental bonus

[2] If workers decided to leave after reaching a certain number of points (10.7% of workers in PR and 20% of workers in FP reported setting a point goal, in the tournament rooms, one worker, in TP30, set such a goal) then we would expect to find negative (or concave) relationships between work time and luck. In PR, for lucky workers the income effect may become larger than the substitution effect. In that case, they will decide to leave, and we expect concave relationships.

[3] The results remain intact if we restrict attention to the tournament treatments, run OLS regressions for worktime, or add demographic background characteristics to the regressions.

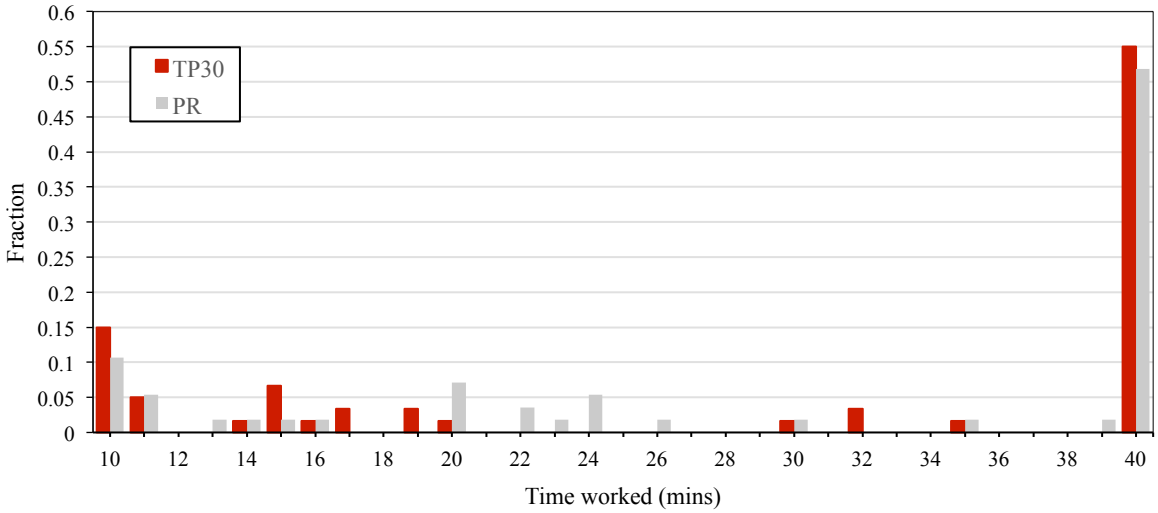
earned in in PR for staying an extra 30 minutes. The TP30 worker who works 40 minutes along with 3 others in the room should anticipate a 25% probability of winning the tournament, which implies an expected bonus of $7.50. Assuming the average cycle of active and rest screens lasts 11.5 seconds, workers should expect to see 156 red squares and 15 gold squares over the course of 30 minutes of (optional) work time. At the PR price, this yields an expected (incremental) bonus of $7.80. This similarity suggests that comparable shares of (risk neutral) workers might be willing to work the full time in PR and TP30. Indeed, as shown in Table 1, we received similar average effort from workers in the two treatments.

We also examined the prediction from the war of attrition setup that workers in TP30 should be more likely to leave the room after the other 3 competitors have left. In a pure war of attrition, the last TP30 worker in a room has no incentive to work more. Our setting has luck, and coworker output is unobservable, so it could still be worthwhile to supply some additional effort, but that is unlikely to mean working for much longer. Without a fellow competitor at work, we do not expect TP30 workers to stay the full 40 minutes, so rooms with only one worker staying 40 minutes should be relatively rare in TP30. There is no such expectation in PR, where compensation is set by individual performance. Consistent with this prediction, we find at least two workers at the end of the session in all 13 of the TP30 rooms with at least one worker persisting for the full 40 minutes. In contrast, 4 out of the 12 PR rooms in which at least one worker worked the full 40 minutes had only one person working the full time. This difference across treatments is statistically significant (Fisher exact test yields $p = 0.039$). We also formalized this with a regression model in which the outcome is the difference in quit times between the last 2 workers in the room. We find that this value is significantly smaller in TP30 as compared to PR.

Finally, we considered the stronger prediction from the finite horizon war of attrition setting that individuals will decide from the outset to either compete for the prize and work to the end or to leave immediate and incur no additional effort costs. Although this is what we find for the large majority of our sample, as shown in Figure 1 in the main text and in Appendix Figure C1, it is not universal: about 30% worked between 11 and 39 minutes. These individuals might have been taking additional time to see what others in the room were doing; or they might have decided not to compete, but stayed longer than the minimum because of the behavioral motivations described in Section 2.1.2. Even without behavioral or peer considerations, the PR

setting lacks a stark prediction about workers being driven to the endpoints of the available time, though that would be implied if both the costs of effort and the utility of income were linear over the work period. In that case, low effort cost workers would stay until the end and high cost workers would leave at 10 minutes. In typical settings with continued employment, it is more natural to expect effort costs to be convex and increasing in work hours and for the marginal utility of income (or consumption goods) to be declining. These forces would lead to more dispersion in work time in PR as compared to tournament settings, because only PR workers would find it worthwhile to invest intermediate or moderate numbers of work hours. We see a hint of this in our sample, as depicted in Appendix Figure C1: 37.5% of workers in PR left between 11 and 39 minutes. This corresponds to a 25% increase over the rate in TP30, consistent with tournament structure polarizing work times, though the difference between the two treatments is not statistically significant.

## Figure C1: Histogram of Time Worked in TP30 and PR




## Appendix Table C1: Effects of Treatments and Luck on Labor Supply

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | Room Level | | Worker Level | |
| | Time Worked | Total Effort in the Session | Time Worked | Total Effort in the Session |
| | **Main Treatments** | | | |
| FP | -12.031*** | -63.227*** | -17.681*** | -64.268*** |
| | [2.437] | [12.567] | [3.343] | [11.857] |
| | **Auxiliary Treatments** | | | |
| TP15 | -5.243* | -25.920* | -7.486* | -23.834* |
| | [2.711] | [13.907] | [4.162] | [13.675] |
| PR | -0.294 | -5.700 | -1.217 | -3.851 |
| | [2.532] | [11.780] | [4.009] | [11.697] |
| | **Luck** | | | |
| Fraction Gold | 653.049*** | 3,337.135*** | 701.906*** | 2,623.025*** |
| | [195.432] | [1,067.813] | [144.688] | [501.627] |
| Fraction Gold Squared | -2,993.232** | -15,144.249** | -3,471.844*** | -12,916.004*** |
| Constant | -2.071 | -13.233 | 4.450 | 31.676 |
| | [7.789] | [43.945] | [6.032] | [21.841] |
| Cluster Level | None | None | Room | Room |
| Prob>F: TP15 = FP | 0.018 | 0.019 | 0.007 | 0.006 |
| Prob>F: PR = FP | < 0.001 | < 0.001 | 0.201 | 0.405 |
| Prob>F: TP15 = PR | 0.086 | 0.163 | 0.143 | 0.149 |
| Observations | 59 | 59 | 236 | 236 |
| $R^2$ | | 0.483 | | 0.269 |

*Notes*: Observations at the room level in Columns 1 and 2 and worker level in Columns 3 and 4. Tobit model is estimated for time worked and linear regressions for effort measures. Fraction gold is the fraction of active screens displayed that included a gold square. Robust standard errors in brackets. *** p < 0.01, ** p < 0.05, * p < 0.1