Tests with Lab Experiments of Hotelling's Rule about Prices of Non-Renewable Resources

Scott R. Templeton, Clemson University, stemple@clemson.edu

Daniel H. Wood, Federal Trade Commission, dhwood@gmail.com

Selected Paper prepared for presentation at the 2023 Agricultural & Applied Economics Association Annual Meeting, Washington DC; July 23-25, 2023

Experimental Tests of Hotelling's Rule about Non-Renewable Resource Prices

Scott R. Templeton
Clemson University

Daniel H. Wood
Federal Trade Commission

May 17, 2023*

Abstract

In resource economics, Hotelling's rule is a prediction that, if certain conditions hold, the net price of an intertemporally scarce, non-recyclable, non-renewable resource rises over time (Hotelling 1931). The interest rate, costs of extraction, and market structure determine the time path of the net price. If the resource is inter-temporally abundant, however, the net price does not rise, regardless of the interest rate. We conduct laboratory experiments to test whether the behavior of mathematically adept sellers of a resource in a dynamic oligopolistic market is consistent with Hotelling's rule. The resource is inter-temporally scarce in our treatments but inter-temporally abundant in our control. In both treatments and the control, average quantities sold over time and associated prices are consistent with qualitative predictions of a Hotelling-inspired model of four Cournot sellers and, to some extent, with quantitative predictions of the model. The rule it is not, however, an accurate predictor of a minority of individual behavior.

Keywords: Hotelling's rule, resource economics, laboratory experiments, dynamic Cournot oligopoly, inter-temporal scarcity

AAEA Subject Code: Resource Economics and Policy

JEL codes: Q31; C73; C91; L13

*The views expressed in this article are those of the authors and do not necessarily reflect those of the Federal Trade Commission or Clemson University. We thank audiences at the 2019 and 2016 Southern Economics Association's annual meetings, 2016 South Carolina Applied Microeconomics Conference, Clemson University, and Lehigh University. We have benefited from feedback from Peter Blair, Daniel Friedman, Jorge Garcia, Doug L. Miller, Mike Makowsky, Jean Paul Rabanal, Dave Schmidt, Zachary Templeton, Chuck Thomas, Patrick Warren, and Bart Wilson. Ali Kuru, Kyle Rechard, Yuvaluck Setboonsrung, Guangua Wang, and especially Jacob Walloga assisted our collection of data. Mia Lv, Chris Cheng, other support staff at Stata, and Zachary Templeton occasionally assisted with programming. Finally, we thank Clemson's University Research Grant Committee for funding these experiments.

1 Background

The theoretical foundation of non-renewable resource economics is built with the models and predictions of Hotelling (1931). Consider a miner who accurately foresees the availability of fixed reserves of a non-renewable resource, sells in a market in which demand for the resource inversely varies with the resource's price but is otherwise constant, and maximizes the present value of his future profits. The miner's profit per unit is "the net price [p] received after [the miner's] paying the cost of extraction and placing [a unit of the mineral] upon the market" (Hotelling 1931, p. 141). The cost of extracton can vary with the degree of physical accessibility (Hotelling 1931, p. 140-141). If the miner sells in a perfectly competitive market, the net price in equilibrium will rise at the rate of interest r. If the miner is a monopolist in equilibrium, the miner's marginal profit, which is less than the net price, will rise at the rate of interest r. The predictions are known as Hotelling's rule (Livernois 2009), or the r-percent rule (Miller and Upton 1985).

Hotelling's paper has been the progenitor of a large theoretical literature. His models of a price-taking or price-setting single firm have subsequently been modified to represent extraction by multiple firms in a competitive market (e.g., Gordon 1967), multiple firms under the coordination of a manager of a common-pool version of the resource (e.g., Cummings 1969), multiple firms with entry and exit in a competitive industry (e.g., Schulze 1974), a dominant firm with a competitive fringe (e.g., Salant 1976), and Cournot oligopolists (e.g., Lewis and Schmalensee 1980). Instead of being embodied in Hotelling's net price of a resource, cost per unit has subsequently been modeled as a constant marginal cost of extraction (e.g., Gordon 1967, p. 277; Stiglitz 1976) or an increasing marginal cost of extraction (e.g., Gordon 1967, pp. 279-280; Schulze 1974). Instead of Hotelling's net price decreasing with past production to reflect increases in the cost of extraction as the mine goes deeper (Hotelling 1931, pp. 152-

153), cost has also subsequently been explicitly modeled to increase with past production (e.g., Gordon 1967, pp. 278-279; Cummings 1969, pp. 207-209), increase with cumulative extraction in the presence of a backstop technology (e.g., Heal 1976), or decrease with current reserves (e.g., Livernois and Martin 2001). In these numerous extensions, except for a few of the models with costs of extraction that increase with cumulative extraction (e.g., Hanson 1980), a version of the r-percent rule still characterizes marginal profit over time. In all of these models the price of the resource increases as the quantity produced decreases over time.

In spite of the canonical importance of Hotelling's models and the numerous extensions to them, the empirical evidence for his prediction about the trends in the marginal profit or price of a non-renewable resource is limited (Livernois 2009). In theoretical models (e.g., Pindyck 1978; Stiglitz 1976) and in actual economies (e.g., Baumeister and Kilian 2016; Fitzgerald 2013), discoveries of new reserves from exploration and cost-lowering innovations in methods of extraction are two important reasons for the frequently observed lack of rise in market prices of oil and other non-renewable resources. OPEC's occasional inability to agree on and maintain production quotas and unexpected recessions are other reasons for decreases in market prices of oil. Unexpected disruptions in production and growth in demand are reasons for price increases that Hotelling's models also do not predict. Discoveries, technical innovations, changes in market structure, disruptions in supply, and business cycle fluctuations are all factors that confound tests of the accuracy of Hotelling's prediction.

We experimentally test Hotelling's rule by inducing and then comparing inter-temporal prices and sales of a non-renewable resource in a laboratory. The simulated market's structure is a Cournot oligopoly with four sellers who have common knowledge about the interest rate, current availability of the resource, and inverse demand for it. The computer purchases the resource at a price that is determined by the total quantity offered for sale in each of four periods. Each seller receives the same endowment of the resource at the beginning of each four-period market. Each seller has the opportunity in each period to sell some, none, or all of the resource that he or she currently owns. A virtual bank pays interest per period on balances that comprise previously received interest and earned revenues. No marginal cost of resource extraction exists, by design.

We test three difference conditions of scarcity and interest rate. In our control, the resource is inter-temporally abundant – there is enough of the resource that it would not run out in equilibrium if the sellers sold as much as was optimal – and theoretically, Hotelling's prediction becomes the prediction of the static theory of four-firm Cournot oligopoly. In our treatments, the resource is inter-temporally scarce, but in one there is a relatively higher interest rate than the other. Theoretically the prices should increase (and quantities sold should fall) more steeply in the high interest rate treatment.

In spite of the advantages of conducting an experiment in a laboratory to test Hotelling's rule, only one peer-reviewed publication on this topic exists (van Veldhuizen and Sonnemans 2018).[1] In the published research, duopolists receive a relatively large endowment of a non-renewable resource in one treatment and a relatively small endowment in the other. The resource is inter-temporally scarce in both treatments. The duopolists with the large endowment "pay significantly less attention to dynamic optimization, and shift extraction to the present, leading them to overproduce relative to the Hotelling rule" (Neumann and Erlei 2014, p. 481).

In addition to directly relating to the Hotelling-inspired literature our work indirectly relates to three strands of literature from laboratory experiments in economics. First, ex-

[1]One working paper also exists, namely Neumann and Erlei (2014).

periments about use of natural resources have been focused on the extent to which people efficiently use renewable, common-pool resources, such as an open-access fishing grounds or unenclosed pastures (e.g., Stoop, van Soest, and Vyrasktekova 2013; Casari and Plott 2003). Second, experiments about prices of storable commodities or assets that are traded over at least two periods of time have been focused on effects of learning, through replication, on efficiency of intertemporal prices (e.g., Miller, Plott, and Smith 1977; Smith, Suchanek, and Williams 1988) or accuracy of forecasted prices (e.g., Kelley and Friedman 2002).[2] Third, experiments with subjects who tackle problems of dynamic optimization have been focused on, among other things, whether their multi-period choices about consumption and saving from income are dynamically efficient (e.g., Hey and Dardanoni 1988; Meissner 2016).[3]

A Hotelling-inspired model of four-firm Cournot oligopoly generates remarkably accurate qualitative predictions about average behavior over time in our experiment. In particular, average individual sales fall and market prices rise over time, as predicted, when the resource is inter-temporally scarce. Moreover, the average price of the scarce resource is initially lower but subsequently higher when the interest rate is high than when the interest rate is low, as predicted. Also, sales and prices do not usually change over time, on average, when the resource is inter-temporally abundant. Individual sellers exhaust their reserves if the resource is intertemporally scarce but rarely exhaust them if the resource is abundant, as predicted. Finally, the null hypotheses that average sales and prices equal the sales and prices implied by Hotelling's rule cannot be rejected in the half of the periods when the resource is scarce with a low interest rate and the first three periods when the resource is abundant. Nonetheless, the Hotelling-inspired model often does not accurately predict quantities sold by individual

[2]See Sunder (1995, pp. 475-481) and Duffy (2015, pp. 14-20) for reviews of literature on experiments about learning to forecast, or formation of expectations.

[3]See Duffy (2015, pp. 4-12) for a review of literature on experiments about intertemporal consumption-saving choices.

sellers and prices in individual markets. There is much dispersion around average behavior and noisiness of individual sales drives dispersion around prices.

In the next section we describe our experimental design. We then develop in Section 3 a four-period game-theoretic model of four Cournot sellers of a non-renewable resource and use the model to generate Hotelling-like predictions of sales and prices. In Section 4 we describe characteristics of the subjects, the seven four-period markets in which they sold their resource during a session, and the sessions that constituted two treatments and one control of the experiment. We present deviations-from-mean models that we estimate with bootstrapped standard errors to generate p-values for hypothesis tests of whether means are consistent with theoretical predictions in Section 5. We also briefly discuss non-parametric tests of whether medians are consistent with the predictions in the same section. We then describe sales of the resource in each period, display the estimated models, and present results of hypothesis tests in Section 6. We concentrate initially on aggregate behavior and subsequently on individual behavior. In Section 7 we describe associated mean and median prices of the resource and present results of hypothesis tests about prices. We then discuss our experimental results in Section 8. We conclude with implications for future research, teaching, and understanding real-world markets for non-renewable resources in Section 9.

2 Experimental Design

We focused our recruitment on undergraduate students with majors that emphasize quantitative or financial methods. Recruitment of subjects occurred from the first week of February 2016 through the first week of April 2016 for sessions with treatments and from the first week of September through the last week of October 2017 for sessions with the control. At our request program coordinators or their secretaries sent a formal invitation to the list-serves of these undergraduate majors: accounting, construction science and management, civil engineering, economics, finance, industrial engineering, and mathematical sciences. Masters students in Applied Economics and Statistics and first-year doctoral students in economics were also invited for sessions with treatments.[4] After participating in their initial session, subjects who had previously indicated an interest in participating in future sessions received an email invitation to participate in a second session. These subsequent sessions, with one exception, were exclusively made up of subjects who had already participated and each subject knew that all other subjects had also already participated. Subsequent sessions are "experienced" sessions while initial sessions are "inexperienced" sessions. Aside from the subject pool, however, the design and administration of inexperienced and experienced sessions were identical.

At the beginning of a session each subject sat in front of a terminal in our computer laboratory and, after reading the informed consent to participate, signed it. Each subject was also given written instructions (Appendix A) and asked to follow along as one of us read them aloud. Subjects could ask a question at any time during our reading of instructions. The instructions include a compound interest table. We also paused and asked for questions after we read how they could earn revenue and receive interest income. Immediately after the instructions were read and questions answered, subjects were required to answer six questions about earning revenue from resource selling, determining the price of the resource in any period, and compounding interest (Appendix B). We checked answers and, if necessary, gave step-by-step explanations to subjects until each of them correctly answered all questions.

[4]We did not have formal invitations sent to undergraduate majors in construction science and management or civil engineering or graduate students in economics for the three sessions with controls in October 2017.

We or our colleagues also read or handed out scripted invitations to students in these undergraduate courses: intermediate microeconomics, introductory environmental economics, introduction to econometrics, game theory, engineering economic analysis, and theory of interest.

Subjects interacted using terminals running z-Tree (Fischbacher 2007). Subjects were randomly assigned into groups of four sellers. We chose a group size of four because experimental Cournot markets with at least four sellers are "never collusive" (Huck, Normann, and Oechssler 2004, p. 435), while five sellers per group would have made recruitment of subjects and formation of same-sized markets more difficult. Subjects were each paid $5 to show up for a session regardless of whether they could be assigned to a four-seller group and, thus, could actually participate in the session.

The randomly-grouped subjects were individual sellers in a simulated *market* that comprised, coincidentally, four selling *periods*. The initial 4-period market was for practice. Subsequent markets in a session were non-practice markets and could potentially affect a subject's earnings. Each of the four periods in the practice market lasted at most 90 seconds and each period in subsequent markets lasted at most 45 seconds.[5] Subjects were randomly re-assigned by z-Tree into new groups of four at the start of each successive market. Subjects were instructed not to inform anyone about the group number into which they were assigned and they interacted anonymously with other sellers through their terminals. Also, to reduce the possibility of end-game effects the exact number of successive markets in which subjects were to participate was not provided in our written instructions or in our answer to any question. Subjects were only told that they would participate in several non-practice markets during the session, although each session was stopped after seven non-practice markets.

Each subject received an endowment of x_0 units of a non-renewable resource at the start of every 4-period market. In each of the four periods each subject s simultaneously decided on q_s, that is, on how much, if any, of her endowment to sell for experimental

[5]If subjects made decision before the time limit, the next period began. Subjects were also allowed to ask a question at any time and, in answering questions, we temporarily suspended the clock. However, almost no one asked questions after the practice market.

credits. The marginal cost of bringing the resource to market was 0 credits per unit. In each period the inverse demand for the resource that the four subjects could offer for sale was $P(Q) = 300 - Q$ credits per unit of the resource for $0 \leq Q \leq 300$ and $P(Q) = 0$ for $Q > 300$, in which $Q = \sum_{i=s}^{4} q_s$ in any period. The amount of the resource that a subject still had available for sale and the sum of the amounts of the resource that the three other subjects still had available for sale were displayed to the subject on her computer screen while she chose how much, if any, to sell in the period. Subjects were informed that any unsold endowment at the end of the fourth period of a market could not be carried over to the next 4-period market to be sold.

In choosing four selling periods per experimental market, or $T = 4$, we wanted subjects to have a time horizon during which making inter-temporal decisions and calculating compounding interest were cognitively neither too simple nor too complex and to be able to participate in several experimental markets within a session. We also chose a time horizon of four periods so that a subject's selling a positive amount in each period was the theoretical prediction.

Given our parametrization of market demand, marginal cost, and the time horizon for trading, the resource is inter-temporally scarce if $x_0 < 240$ units for each of four Cournot oligopolists. We set $x_0 = 100$ units to create inter-temporal scarcity for our two treatments and $x_0 = 300$ to create inter-temporal abundance for our control. We also chose $x_0 = 100$ to enable subjects to think about their decisions to sell as decisions about proportions of their initial endowment to sell (Table 1).

The sales revenue of subject s in a period, $P(Q)q_s$, the total amount sold by the four sellers in the period, Q, and the market-clearing price, P, were reported through z-Tree to the subject at the end of each period. Subject s's sales revenue was deposited at the end of each period in a virtual bank that paid interest on the balance at a rate of r per period in future periods. Two interest rates, $r = 0.5$ and $r = 0.25$ per period, were used in the resource-scarce treatments with $x_0 = 100$ and one interest rate, $r = 0.5$, was used in the control with $x_0 = 300$. The rates were constant within a session. Large interest rates were chosen for salience. (Table 1).

Table 1: Parameter Values for Each Type of Experiment

Type	Endowment x_0	Interest Rate r	Fixed Cost f
Low r Treatment	100	0.25	12500
High r Treatment	100	0.50	25000
High r Control	300	0.50	15000

We explicitly asked subjects to maximize their final earnings for each 4-period market, which were

$$\pi_s = \sum_{t=1}^{4}(1+r)^{4-t}P_t(Q_t)q_{s,t} - f.$$

The parameter f is a transaction-processing fee and was deducted once per market, at the end of the last period in each market. In particular, a fixed cost of $f = 25,000$ credits if $r = 0.5$ and $x_0 = 100$, $f = 12,500$ credits if $r = 0.25$ and $x_0 = 100$, and $f = 15,000$ credits if $r = 0.50$ and $x_0 = 300$ was deducted from virtual bank balances of all subjects to determine final net balances. Final net balances, π_s, could not, by design, become negative; a final net balance that would have otherwise been negative was converted to zero (Table 1).

The main purpose of deducting the transaction-processing fee was to accentuate the payoffs from optimal choices relative to payoffs from non-optimal choices. For example, the payoff to playing the subgame-perfect equilibrium (SPE) strategy with $r = 0.5$ and $x_0 = 100$ when other subjects also play that strategy generates 38,555 credits in sales revenue and accumulated interest, while naively selling $q_{s,t} = 25$ each period when three opponents play the subgame-perfect strategy generates 36,799 credits. Expected earnings would have been only 5% lower with the naive strategy than with the SPE strategy if $f = 25,000$ had not been not deducted, while expected earnings were 15% lower with the naive strategy than with the SPE strategy because the fee was deducted.

In addition to accentuating payoffs, the chosen transaction-processing fees also roughly equalized expected earnings per session across our two treatments and one control. In particular, if each subject chooses quantities in the subgame-perfect strategy of four Cournot firms, the expected earnings of a subject with the respective transaction-processing fees subtracted per market are these: 1) 13,555 credits with $r = 0.5$ and $x_0 = 100$, 2) 15,844 credits with $r = 0.25$ and $x_0 = 100$, and 3) 14,250 with $r = 0.5$ and $x_0 = 300$.

Finally, to attenuate any wealth effects of early-market earnings on late-market behavior (Friedman and Sunder 1994, pg. 51), we informed subjects at the start of their session that they would be paid for two 4-period markets that z-Tree would randomly select and reveal to them at the end of the session. We paid for two rather than one period in order to reduce the variance in subject earnings. Final net balances from the two randomly selected markets were converted at a rate of one dollar per 2000 experimental credits.

3 Theoretical Model and Predictions

Lewis and Schmalensee (1980) first developed r-percent rules for optimal, continuous-time extraction of a non-renewable resource under Cournot competition. In this section, we use the theory of subgame perfect equilibrium to develop analogous discrete-time predictions about behavior in our experimental environment.

N firms sell a non-renewable resource over T periods $t = 1, \ldots, T$. Demand is stationary and the inverse market demand each period is $P = a - Q$ where $Q = \sum_s^N q_i$. Each firm s

received an identical endowment of $x_{s,0}$ units of the resource before $t = 1$ and had interim stocks $x_{s,t} = x_{s,0} - q_{s,1} - \ldots - q_{s,t-1}$ at the beginning of period t.

Let the $q^*_{s,t}$ denote the the ex ante optimal quantity choice at period t. In other words, $q^*_{s,t}$ is optimal if every firm makes the SPE quantity choices $q^*_{s,1}, \ldots, q^*_{s,t-1}, q^*_{s,t+1}, \ldots$. The SPE quantity choice is unique. Appendix C contains proofs and further definitions.

Behavior qualitatively depends on the aggregate resource endowment.

Definition 1. If $x_0 < \frac{TA}{N+1}$ then resources are *inter-temporally scarce.*

If resources are not scarce, then firms should play the static NE of the Cournot game; quantity supplied and prices are then constant. If resources are scarce, however, the static NE is not feasible because playing it for T periods would overexhaust available stocks. Instead firms sell until the optimal stopping period.

Definition 2. The *optimal stopping period* T^* is the last period in which $q^*_{i,t} > 0$.

In both our intertemporally scarce treatments, $T^* \geq T = 4$.[6] In equilibrium all firms sell equal quantities, which combined with linear demand leads to each firm's marginal revenue being $A - (N+1)q^*_{i,t}$. Optimally responding to the available interest rate implies that every firm's behavior in equilibrium will satisfy three r-percent rules, namely $(1+r)[A - (N+1)q^*_{s,t}] = A - (N+1)q^*_{s,t+1}$ for $t < T$. Additional calculation shows that

Proposition 1. *If the resource is inter-temporally scarce and $T^* \geq T$, then the SPE quantity*

[6] $T^* \geq T$ if $\frac{x_0}{T-(1+r)+(1+r)^{T-1}} \geq \frac{A}{N+1}$, in which case firms never optimally sell 0 units. For $r = 0.5$, $T^* = 4$, while for $r = 0.25$, if there were five or more periods, T^* would equal 5. In other words, four selling periods were, by design, close to the optimal number periods for four Cournot oligopolistic sellers to exhaust their endowments for both treatments.

supply path is, for all s,

$$q^*_{s,1} = \frac{A}{N+1} - \left(\frac{1+r}{\sum_{t=1}^{T}(1+r)^t} \right) \left(\frac{TA}{N+1} - x_0 \right)$$
$$q^*_{s,t+1} = q^*_{s,t} - r \left(\frac{A}{N+1} - q^*_{s,t} \right).$$

Qualitatively, the SPE quantity supply increases over time, and the increase is steeper for higher interest rates or resource scarceness (i.e., lower x_0). The top rows of Tables 2 and Tables 3 report the SPE quantity paths for our treatments, as well as implied prices and marginal revenue.[7]

The second set of rows report predicted $q^*_{s,t}$ for two alternative behavioral assumptions. The first is quantity sold if firms successfully collude – i.e., the quantity sold choices that maximize joint profit, while the second is quantity sold if firms act as price-takers – i.e., the quantity sold choices ignoring own effects on P.

4 Description of Subjects, Markets, and Sessions

One-hundred forty students participated as inexperienced sellers in one of three types of experimental session. There were 60 first-time subjects in treatment sessions when $r = 0.5$ and $x_0 = 100$ and 52 first-time subjects in treatment sessions when $r = 0.25$ and $x_0 = 100$. There were also 28 first-time subjects in control sessions when $r = 0.5$ and $x_0 = 300$ (Table 4).

Of the 140 individuals who participated in treatment or control sessions as first-time subjects, 63 of them also participated again in a session of entirely experienced subjects.

[7] Quantity sold for $r = 0.25$ does not sum to exactly 100 due to rounding.

Table 2: $r = 0.25$ Predictions

Period	1	2	3	4
SPE $q^*_{s,t}$	35.7	29.7	22.1	12.6
SPE price	157.1	181.4	211.8	249.7
SPE marginal revenue	121.4	151.8	189.7	237.1
Collusive $q^*_{s,t}$	28.8	26.7	24.0	20.6
Price-taking $q^*_{s,t}$	40.3	31.6	20.8	7.2

Table 3: $r = 0.5$ Predictions

Period	1	2	3	4
SPE $q^*_{s,t}$	42.8	34.2	21.2	1.8
SPE price	128.9	163.4	215.1	292.6
SPE marginal revenue	86.0	129.1	193.8	291.1
Collusive $q^*_{s,t}$	31.3	28.3	23.7	16.7
Price-taking $q^*_{s,t}$	48.7	35.6	15.8	0

Table 4: Parameter Values and Number of Subjects by Experience and Type of Session

session type	a	x_0	r	f	inexperienced	experienced	all subjects
high-r treatment	300	100	0.50	12500	60	36	96
low-r treatment	300	100	0.25	25000	52	24	76
high-r control	300	300	0.50	15000	28	12	40

Thirty one of the 63 were second-time subjects in a session with a treatment of $r = 0.5$. Five additional students, not among the 140, also participated with experienced subjects in a session with a treatment of $r = 0.5$ (Table 4). The five were considered experienced because they had participated in a pilot control session of twelve subjects during February 2016. Four of the 63 experienced students had also participated, by accident, with four inexperienced subjects in a session with a treament of $r = 0.25$. Twelve of the 63 were second-time subjects in a control session with $r = 0.5$ and $x_0 = 300$ (Table 4). In total, 145 students participated as 172 subjects who received one of two types of treatments or as 40 subjects who received the control. The 212 subjects included 72 subjects who participated as experienced sellers, 4 of whom did so twice (Table 4).

All subjects had majors with quantitative or financial emphases (Table 5). The five most common majors were economics, industrial engineering, mathematical sciences, accounting, and civil engineering. In addition to economics and accounting, three other undergraduate majors–marketing, business management, and professional golf management–constituted the "economics and business" majors. They represented the majority of majors in the high-r treatment, almost half in the high-r control, and more than one-third of majors in the low-r treatment (Table 5). In addition to industrial engineering, mathematical sciences, and civil engineering, three other majors– biosystems engineering, computer engineering, and mechanical engineering–constituted the "math and engineering" majors. Math and engineering majors represented more than half of all subjects. "Other preprofessional" majors were biological science, packaging science, pre-pharmacy, or political science. Seniors and juniors were the first and second most represented class ranks (Table 5). They accounted for almost two-thirds of all subjects. Almost two-thirds of the subjects were male (Table 5). We conducted 10 sessions with the high-interest-rate treatment from February 12 to April 1 of 2016, 7 sessions with the low-interest-rate treatment from March 25 to April 15 of 2016, and and 3 sessions with the control from Sept. 29 to Oct. 26 of 2017. Most sessions lasted 60-70 minutes. Eight, twelve, or sixteen subjects participated in a session.[8] Thus, two, three, or four separate markets, called "groups" in z-Tree, each with exactly four sellers, were in operation at any point in time during a session. After the initial practice market each subject participated in seven additional, successive 4-period markets during a session. However, data

[8]Ten treatment sessions had 8 subjects, five treatment sessions and two control sessions had 12 subjects, and two treatment sessions and one control session had 16 subjects.

Table 5: Proportions (Percentages) of Subjects by Gender, Class, and Major for Each Type of Session

		high-r treatment	low-r treatment	high-r control	All Types of Session
Gender	Male	60.4	78.9	50.0	65.1
	Female	39.6	21.1	50.0	34.9
Class	Freshman-Sophomore	19.8	28.9	10.0	21.2
	Junior	22.9	21.1	55.0	28.3
	Senior	37.5	38.2	32.5	36.8
	Graduate	19.8	11.8	2.50	13.7
Major	Business and Economics	52.1	36.8	47.5	45.8
	Math and Engineering	46.9	56.6	50.0	50.9
	Other Pre-Professional	1.0	6.6	2.5	3.3
	Number of Sellers (S):	96	76	40	212

about the decisions of eight subjects in their seventh market, or replication, in a session with the high-r treatment do not exist because z-Tree crashed. Thus, the 96 subjects with the high-r treatment made 664 (=96*7 − 8) sets of decisions about individual quantities of the resource to sell per period during 166 four-period markets. The 76 subjects with the low-r treatment made 532 (=76*7) sets of decisions about individual quantities to sell during 133 four-period markets. The 40 subjects in sessions with the high-r control made 280 (=40*7) sets of decisions about four individual quantities to sell during 70 four-period markets. Given four sellers per market, the observations about individual quantities sold also became 166, 133, and 70 observations about market prices per period with the two treatments and one control.

The average–mean and median–monetary earnings of a subject in two randomly chosen markets in an experimental session are close to the predicted earnings of a four-firm Cournot oligopolist that sells a non-renewable resource in two markets, whether the resource is intertemporally scarce or not and whether the interest rate is high or low if the resource is scarce (Table 6).

Table 6: Predicted Earnings and Distribution of Actual Earnings ($ per session) of a Subject in Two Markets by Type of Session

session type	S	predicted	mean	std. dev.	minimum	median	maximum
high-r treatment	96	13.55	13.13	4.08	0.00	12.89	21.73
low-r treatment	76	15.84	14.77	2.21	6.27	14.86	19.83
high-r control	40	14.25	14.07	8.48	0.00	14.55	34.73

Predicted and actual earnings exclude $5 for showing up.

No type of seller, except sellers who are at least seniors in the high-r treatment, has mean earnings that statistically differ from those of any other type of seller. Nonetheless, a few patterns in the sample means of earnings are noteworthy. In each type of session the mean earnings of female subjects always exceeded the mean earnings of male subjects. Students who were at least seniors always earned more than students who were at most juniors and the difference was statistically significant (p-value = 0.0132) for subjects in the high-r treatment. The sample-mean earnings of math science majors were higher than sample-mean earnings of the other majors in the two treatments. (See Table **??** in Appendix **??** for details.)

5 Hypotheses and Statistical Methods to Test Them

5.1 Average Sales

Our model makes predictions about quantities that individuals sell in each period and changes in sales over time. We test the predictions with non-parametric tests of medians and means in regression models. While we use bootstrap methods to avoid distributional assumptions about standard errors in regression models, the non-parametric tests of medians are also agnostic about the data-generating processes. We view the two methods as complementary.

In particular, we test whether the average–median or mean–quantity sold per seller decreases over time if the resource is inter-temporally scarce or does not change over time if the resource is inter-temporally abundant. The associated null hypothesis is $H_0\colon \mu_t^q \leq \mu_v^q$ if the resource is intertemporally scarce or $H_0\colon \mu_t^q = \mu_v^q$ if the resource is intertemporally abundant for $t = 1, \ldots, 3$ and $v = t + 1, \ldots, 4$. Most importantly, we test whether average sales per seller differ from the quantitative predictions of our Hotelling-inspired model of four Cournot sellers. The associated null hypothesis is $H_0\colon \mu_t^q = \mu_t^{q_H}$ for $t = 1, \ldots, 4$. We then also test for systematic divergences from mean sales and predictions of our Hotelling-inspired model. A systematic divergence from mean sales would exist if, for example, mean sales of female subjects were different from mean sales of male subjects. The associated null hypothesis would be $H_0\colon \mu_t^{q_F} = \mu_t^{q_M}$ for $t = 1, \ldots, 4$. A systematic divergence from Hotelling would exist if, for example, mean sales of female subjects differ from the quantitative prediction of our Hotelling-inspired model. The associated null hypothesis would be $H_0\colon \mu_t^{q_F} = \mu_t^{q_H}$ for $t = 1, \ldots, 4$.

To test alternative hypotheses that the median quantity sold in the first, second, or third period exceeds or differs from the median quantity sold in any subsequent period, we use the "signtest" command in Stata. A sign test is a non-parametric test of exceedance or differences in observations of matched, or paired, samples (e.g., Cochran and Cochran 1989, pgs. 138-140). The test statistic for the sign test has a binomial distribution. In our tables the probability associated with the statistic is p1(bi_0) for a one-sided test of exceedance when the resource is scarce and pr(bi_0) for a two-sided test of difference from zero when the resource is abundant. We also use Stata's "signtest" to determine whether median sales differ from the quantitative predictions of the Hotelling-inspired model of four Cournot sellers with an inter-temporally scarce or abundant resource. The probability associated with the test statistic is $\mathrm{pr}(bi_H)$ for a two-sided test of difference from Hotelling's prediction. No assumption about the shape of the distribution of the quantities sold is required or made for these tests (Cochran and Cochran 1989, pg. 139).

To test an alternative hypothesis that the mean quantity sold in the first, second, or third period exceeds or differs from the mean quantity sold in any subsequent period, we use the original sample and bootrapped samples to regress differences in sales in any two periods on a constant and calculate cluster-robust standard errors. For calculation of the standard errors and resampling we specify a "cluster" as a participant, the unique identity of a subject, to account for possible correlation among choices that the participant makes within the seven markets in a session and, if applicable, across sessions. To ensure that a difference in sales between two periods pertains to only one particular subject-market-session, we use the same seed number in STATA to start the bootstrapping. We then adapt the steps in (Cameron and Miller 2015, p. 343) and calculate values of the square root of a Wald statistic, or values of a "Wald t-statistic", from the original sample and each bootstrap sample. Under $H_0\colon \mu_t^q \leq \mu_v^q$ when the resource is intertemporally scarce or $H_0\colon \mu_t^q = \mu_v^q$ when the resource is not for $t = 1, \ldots, 3$ and $v = t+1, \ldots, 4$, the Wald t-statistic from the original sample is $w_{tv}^q \equiv \hat{\mu}_{tv}^q / se_{\hat{\mu}_{tv}^q}$. The terms $\hat{\mu}_{tv}^q \equiv (\hat{\mu}_t^q - \hat{\mu}_v^q)$ and $se_{\hat{\mu}_{tv}^q}$ are the estimated mean of the difference in sales in two time periods, t and v, and the cluster-robust standard error of the estimated mean difference in the original sample. The Wald t-statistic from the b-th bootstrap sample is $w_{tv}^q(b) \equiv [\hat{\mu}_{tv}^q(b) - \hat{\mu}_{tv}^q]/se_{\hat{\mu}_t^q(b)}$ for $b = 1, \ldots, B$. The terms $\hat{\mu}_{tv}^q(b) \equiv [\hat{\mu}_t^q(b) - \hat{\mu}_v^q(b)]$ and $se_{\hat{\mu}_{tv}^q}(b)$ are the estimated mean of the difference in sales in two time periods and the cluster-robust standard error of the estimated mean difference for $b = 1, \ldots, B$.

To test an alternative hypothesis that the mean quantity sold by a seller in period t differs from the quantitative prediction of our theoretical model, we again use the original sample and bootrapped samples to regress, in this case, sales in the period on a constant. Each participant is again a "cluster" for calculation of standard errors and resampling. Under $H_0\colon \mu_t^q = \mu_t^{q_H}$ for $t = 1, \ldots, 4$, the Wald t-statistic from the original sample is $w_t^q \equiv [\hat{\mu}_t^q - \mu_t^{q_H}]/se_{\hat{\mu}_t^q}$. The terms $\hat{\mu}_t^q$ and $se_{\hat{\mu}_t^q}$ are the estimated mean sales of a seller in period t and the cluster-robust standard error of the estimated mean in the original sample. The term $\mu_t^{q_H}$ is the quantitative prediction. The Wald t-statistic from the b-th bootstrap sample is $w_t^q(b) \equiv [\hat{\mu}_t^q(b) - \hat{\mu}_t^q]/se_{\hat{\mu}_t^q}(b)$, in which $\hat{\mu}_t^q(b)$ and $se_{\hat{\mu}_t^q}(b)$ are the estimated mean sales and cluster-robust standard error of the estimated mean for $b = 1, \ldots, B$.

Our final step to make inferences about mean sales of a seller is to calculate a "bootstrap p-value"(Wooldridge 2010, p. 440), also known as an "achieved significance level"(Efron and Tibshirani 1994, p. 203). In particular, the bootstrap p-value for a one-sided test of exceedance of mean sales in period t over mean sales in any subsequent period s when the resource is intertemporally scarce is

$$p1(\mathrm{b}t_0) \equiv \sum_{b=1}^{B} \frac{\mathbf{1} \cdot [w_{tv}^q(b) \geq w_{tv}^q]}{B}.$$

The bootstrap p-values for a two-sided test of difference between mean sales in period t and in any subsequent period s when the resource is not intertemporally scarce is

$$p(\mathrm{b}t_0) \equiv \sum_{b=1}^{B} \frac{\mathbf{1} \cdot [|w_{tv}^q(b)| \geq |w_{tv}^q|]}{B}.$$

The bootstrap p-value for a two-sided test of difference between mean sales and the quantitative prediction of our Hotelling-inspired model in period t is

$$p(\mathrm{bt}_H) \equiv \sum_{b=1}^{B} \frac{\mathbf{1} \cdot [|w_t^q(b)| \geq |w_t^q|]}{B}.$$

To test Hotelling's predictions and possible systematic divergences from them, we estimate models of population means with deviations from them as suggested by Suits (1984). For each of the three types of experimental session, let q_{smt} be the non-negative quantity of the resource sold by the s-th subject in the t-th period of the m-th market in a session. Two types of systematic divergences are considered: 1) those that might originate with observed and unobserved characteristics of the subjects and 2) those that might originate with characteristics of our experimental design.

Let q_{smt} be the quantity that subject s sells during market m in period t. Suppose that the mean sales of female subjects might differ from the mean sales of male subjects and, thereby, from the overall mean, μ_t^q. Let FEMALE_s and MALE_s indicate whether subject s reports a gender of female or male. Let S_F and S_M represent the number of observations, or non-negative sales, of female and male subjects and $S = S_F + S_M$ represent the total number of observations.[9] In formal terms,

$$q_{smt} = \mu_t^q + \alpha_t^F \text{FEMALE}_s + \alpha_t^M \text{MALE}_s + \epsilon_{smt}, \text{for } s = 1, \ldots, S.$$

The gender-deviation-from-mean model is estimated with data from the original sample and also repeatedly re-estimated with data from bootstrapped samples. Each participant, who can be a subject more than once, is again a "cluster" for calculation of standard errors and resampling. Replication in each bootstrap of the original number of participants does not necessarily replicate the original sample size or the number of observations that female

[9] In the control and low-r treatment S_F, S_M, S could be interpreted as the number of female, male, and all subjects because each subject generates seven observations. In one session with the high r treatment, however, subjects generated only six observations.

and male participants generate, however, because almost all participants who sell in two sessions generate twice as many observations as participants who sell in only one session.[10] As a consequence, the sampling weights S_F, S_M, and S must be re-calculated with each bootstrap. The constraint $\alpha_t^F(S_F/S)+\alpha_t^M(S_M/S)=0$ is imposed in all samples and implies that the estimates of α_t^F and α_t^M are the mean deviations of sales of females and males from mean sales in each sample.

To test whether mean sales differ by gender, we again adapt steps in Cameron and Miller (2015, p. 343) and calculate values of a "Wald t-statistic" from the original and each bootstrap sample. Let $\mu_t^{q_F} \equiv \mu_t^q + \alpha_t^F$ and $\mu_t^{q_M} \equiv \mu_t^q + \alpha_t^M$. Then $H_0\colon \mu_t^{q_F} = \mu_t^{q_M}$ is equivalent to $H_0\colon \alpha_t^F = \alpha_t^M$ for $t = 1, \ldots, 4$. The Wald t-statistic for the original sample is $w_t^{q_{FM}} \equiv (\hat{\mu}_t^{q_F} - \hat{\mu}_t^{q_M})/se_{\hat{\mu}_t^{q_{FM}}} = (\hat{\alpha}_t^F - \hat{\alpha}_t^M)/se_{\hat{\alpha}_t^{FM}}$, in which $se_{\hat{\alpha}_t^{FM}} \equiv \sqrt{\hat{\sigma}^2_{\hat{\alpha}_t^F} - 2\hat{\sigma}_{\hat{\alpha}_t^F\hat{\alpha}_t^M} + \hat{\sigma}^2_{\hat{\alpha}_t^M}}$. The terms $(\hat{\alpha}_t^F - \hat{\alpha}_t^M)$ and $se_{\hat{\alpha}_t^{FM}}$ are the estimated difference between mean sales of females and males in period t and the cluster-robust standard error of the estimated mean difference for the original sample. The Wald t-statistic from the b-th bootstrap is $w_t^{q_{FM}}(b) \equiv [\hat{\alpha}_t^{FM}(b) - \hat{\alpha}_t^{FM}]/se_{\hat{\alpha}_t^{FM}}(b)$, in which $\hat{\alpha}_t^{FM}(b) \equiv \hat{\alpha}_t^F(b) - \hat{\alpha}_t^M(b)$ and $se_{\hat{\alpha}_t^{FM}}(b) \equiv \sqrt{\hat{\sigma}^2_{\hat{\alpha}_t^F}(b) - 2\hat{\sigma}_{\hat{\alpha}_t^F\hat{\alpha}_t^M}(b) + \hat{\sigma}^2_{\hat{\alpha}_t^M}(b)}$. The terms $\hat{\alpha}_t^{FM}(b)$ and $se_{\hat{\alpha}_t^{FM}}(b)$ are the estimated difference between the mean sales of females and males in period t and the cluster-robust standard error of the estimated mean difference for $b = 1, \ldots, B$. The bootstrap p-value for a two-sided test of whether mean sales of females differs from mean sales of males in period t is

$$p(\mathrm{bt}_0) \equiv \sum_{b=1}^{B} \frac{\mathbf{1}\cdot[|w_t^{q_{FM}}(b)| \geq |w_t^{q_{FM}}|]}{B}.$$

To test whether mean sales of females or males in period t differ from the quantitative prediction of the Hotelling-inspired model of four Cournot sellers, we again use estimates

[10] Suppose, for example, that a participant who sells in two sessions is randomly chosen twice and a participant who sells in one session is not randomly chosen for a particular bootstrap. *Ceteris paribus*, the number of observations in the bootstrap increases, on net, by seven.

in the gender-deviation-from-mean model. The null hypothesis related to female sellers is $H_0\colon \mu_t^{q_F} \equiv \mu_t^q + \alpha_t^F = \mu_t^{q_H}$. Under H_0, the Wald t-statistic for the original sample is $w_t^{q_F} \equiv (\hat{\mu}_t^{q_F} - \mu_t^{q_H})/se_{\hat{\mu}_t^{q_F}}$, in which $se_{\hat{\mu}_t^{q_F}} \equiv \sqrt{\hat{\sigma}^2_{\hat{\mu}_t^q} + 2\hat{\sigma}_{\hat{\mu}_t^q \hat{\alpha}_t^F} + \hat{\sigma}^2_{\hat{\alpha}_t^F}}$. The terms $\hat{\mu}_t^{q_F} \equiv \hat{\mu}_t^q + \hat{\alpha}_t^F$ and $se_{\hat{\mu}_t^{q_F}}$ are the estimated mean sales of female sellers in period t and the standard error of the female's estimated mean. In words, the original sample t-value is calculated under the null hypothesis that the effect of FEMALE is zero and, thus, the sum of the constant and the coefficient on FEMALE does not differ from the prediction of the theoretical model of four Cournot firms with an intertemporally scarce or abundant resource. (Keep the previous sentence?) The Wald t-statistic from the b-th bootstrap is $w_t^{q_F}(b) \equiv [\hat{\mu}_t^{q_F}(b) - \hat{\mu}_t^{q_F}]/se_{\hat{\mu}_t^{q_F}}(b)$, in which $\hat{\mu}_t^{q_F}(b) \equiv \hat{\mu}_t^q(b) + \hat{\alpha}_t^F(b)$ and $se_{\hat{\mu}_t^{q_F}}(b) \equiv \sqrt{\hat{\sigma}^2_{\hat{\mu}_t^q}(b) + 2\hat{\sigma}_{\hat{\mu}_t^q \hat{\alpha}_t^F}(b) + \hat{\sigma}^2_{\hat{\alpha}_t^F}(b)}$ for $b = 1, \ldots, B$. The bootstrap p-value for a two-sided test of difference between mean sales of females and the Hotelling-inspired prediction in period t is

$$p(\mathrm{bt}_H) \equiv \sum_{b=1}^{B} \frac{\mathbf{1} \cdot [|w_t^{q_F}(b)| \geq |w_t^{q_F}|]}{B}.$$

Substitute M for F in the formula above to characterize the null hypothesis, Wald t-statistics, and bootstrap p-value for male sellers.

The deviation of the s-th seller's sales from the overall mean in the t-th period might depend on, in addition to gender, any number of other observable and unobservable characteristics. To represent all possible characteristics of a seller, let $\mathrm{ID1}_s$, $\mathrm{ID}j_s$, and IDJ_s indicate whether the s-th seller is the first, j-th, or last (J-th) individual participant in one of the three types of experimental session. Note that $J < S$ because a participant may sell in more than one session and, thus, may be a subject more than once. In formal terms, our model is

$$q_{smt} = \mu_t^q + \alpha_t^1 \mathrm{ID1}_s + \ldots + \alpha_t^j \mathrm{ID}j_s + \ldots + \alpha_t^J \mathrm{IDJ}_s + \epsilon_{smt}, \text{for } s = 1, \ldots, S, \text{ in which,}$$

as before, q_{smt} is the sales of subject s during market m in period t and μ_t^q is the population mean of sales in the t-th period during sessions with one of two treatments or one control. Let $\alpha_t^1(N_1/N)+\ldots,+\alpha_t^j(N_j/N)+\ldots,+\alpha_t^J(N_J/N)=0$, where N_1, N_j, N_J, and N represent the number of observations that the first, j-th, last, and all participants generated in period t during one or more sessions of the same type. The constraint implies that the least-squares estimate of the parameter α_t^j represents the mean deviation of the j-th participant's sales from the overall mean of the sample. The standard errors in the participant's-deviation-from-mean model are clustered on each unique four-seller group to account for possible correlation between sales of four sellers during a particular market and session.[11]

To test whether mean sales of participant j in period t differs from the prediction of the Hotelling-inspired model, we use estimates and cluster-robust standard errors from the participant-deviation-from-mean model. The null hypothesis is $H_0\colon \mu_t^{q_j} \equiv \mu_t^q + \alpha_t^j = \mu_t^{q_H}$. Under H_0, the Wald t-statistic for the original sample is $w_t^{q_j} \equiv (\hat{\mu}_t^{q_j} - \mu_t^{q_H})/se_{\hat{\mu}_t^{q_j}}$, in which $\hat{\mu}_t^{q_j} \equiv \hat{\mu}_t^q + \hat{\alpha}_t^j$ and $se_{\hat{\mu}_t^{q_j}} \equiv \sqrt{\hat{\sigma}^2_{\hat{\mu}_t^q} + 2\hat{\sigma}_{\hat{\mu}_t^q\hat{\alpha}_t^j} + \hat{\sigma}^2_{\hat{\alpha}_t^j}}$. The terms $\hat{\mu}_t^{q_j}$ and $se_{\hat{\mu}_t^{q_j}}$ are the estimated mean sales of the j-th participant in period t and the standard error of the participant's estimated mean. The Wald t-statistic from the b-th bootstrap is $w_t^{q_j}(b) \equiv (\hat{\mu}_t^{q_j}(b) - \hat{\mu}_t^{q_j})/se_{\hat{\mu}_t^{q_j}}(b)$, in which $\hat{\mu}_t^{q_j}(b) \equiv \hat{\mu}_t^q(b) + \hat{\alpha}_t^j(b)$ and $se_{\hat{\mu}_t^{q_j}}(b) \equiv \sqrt{\hat{\sigma}^2_{\hat{\mu}_t^q}(b) + 2\hat{\sigma}_{\hat{\mu}_t^q\hat{\alpha}_t^j}(b) + \hat{\sigma}^2_{\hat{\alpha}_t^j}(b)}$ for $b = 1, \ldots, B$. The bootstrap p-value for a two-sided test of difference between mean sales of the j-th participant and the Hotelling-inspired prediction in period t is

$$p(\mathrm{bt}_H) \equiv \sum_{b=1}^{B} \frac{\mathbf{1}\cdot\left[\left|w_t^{q_j}(b)\right| \geq \left|w_t^{q_j}\right|\right]}{B}.$$

The number of participants with bootstrap p-values larger than 0.05 is counted as the number for whom the null hypothesis of consistency with Hotelling's predictions cannot be rejected.

[11] The number of clusters–the number of group-market-sessions in the participant-deviation-from mean models or the number of participants in the other deviation-from-mean models–is considered the number of independent observations.

To represent characteristics of our experimental design, let INEXP_s or EXPER_s indicate whether subject s participates in a session as an inexperienced or experienced seller.[12] Moreover, let MARKETm, in which $m = 1, ..., 7$, indicate the replication number, or round number, of the four-period market in which a subject sells during a session.[13] In formal terms, our model with key design features of a session is this:

$$\begin{aligned} q_{smt} = &\mu_t^q + \alpha_{It}\text{INEXP}_\text{s} + \alpha_{Et}\text{EXPER}_\text{s} + \beta_{1t}\text{MARKET1} + \ldots + \beta_{7t}\text{MARKET7}+ \\ &\gamma_{I1t}\text{INEXP}_\text{s}\cdot\text{MARKET1} + \ldots + \gamma_{I7t}\text{INEXP}_\text{s}\cdot\text{MARKET7}+ \\ &\gamma_{E1t}\text{EXPER}_\text{s}\cdot\text{MARKET1} + \ldots + \gamma_{E7t}\text{EXPER}_\text{s}\cdot\text{MARKET7} + \epsilon_{smt}, \text{for } t = 1, \ldots, 4. \end{aligned}$$

Constrain the parameters in the following ways. First, $\alpha_{It}(S_I/S) + \alpha_{Et}(S_E/S) = 0$, where S_I, S_E, and S represent the number of inexperienced, experienced, and all subjects who participate in a particular type of experimental session. Second, $\beta_{1t}+\ldots+\beta_{mt}+\ldots+\beta_{7t} = 0$. Third, $\gamma_{I1t} + \ldots + \gamma_{Imt} + \ldots + \gamma_{I7t} = 0$. Fourth, $\gamma_{E1t} + \ldots + \gamma_{Emt} + \ldots + \gamma_{E7t} = 0$.[14] Fifth, $\gamma_{I1t}(S_I/S) + \gamma_{E1t}(S_E/S) = 0, \ldots, \gamma_{Imt}(S_I/S) + \gamma_{Emt}(S_E/S) = 0, \ldots,$ and $\gamma_{I6t}(S_I/S) + \gamma_{E6t}(S_E/S) = 0$. The third and fourth constraints and the fifth set of constraints imply that $\gamma_{I7t}(S_I/S) + \gamma_{E7t}(S_E/S) = 0$.

If these constraints are imposed, the least-squares estimate of the constant is the sample mean for the particular type of session. Also, the estimated constant plus the least-squares estimates of deviations from it are the sub-sample means for the relevant sub-populations in

[12] The variable INEXP_s or EXPER_s is almost always identical to a variable that indicates whether a session has exclusively inexperienced or experienced sellers. The only difference in our study is that one session with the low-r treatment has, by accident, four inexperienced and four experienced sellers.

[13] Recall that subjects initially have a practice market. Thus, MARKET1 is the first replication of the four-period market, the first market in which a seller can potentially earn income.

[14] The weight is actually 1/7 on each of the seven parameters in the second through fourth constraints for data from the low-r treatment and the control. The weight is actually 24/166 on the first six parameters and 22/166 on seventh parameter, which corresponds to MARKET7, in the second through fourth constraints for data from the high-r treatment because data for the seventh market in one session are not available.

the type of session. For example, the sub-sample mean of the quantity sold by inexperienced subjects in the second period is $\hat{\mu}_2 + \hat{\alpha}_{I2}$, the sub-sample mean of the quantity sold in the second period of the sixth market is $\hat{\mu}_2 + \hat{\beta}_{62}$, and the sub-sub-sample mean of the quantity sold by inexperienced subjects in the second period of the sixth market is $\hat{\mu}_2+\hat{\alpha}_{I2}+\hat{\beta}_{62}+\hat{\gamma}_{I62}$. Therefore, as specified, the deviations-from-mean models of our experimental design account for possibilities that experience of some subjects across sessions (EXP = 1) might affect means sales or that experience within a session, i.e., the seven-fold repetition of a 4-period market (MARKET1 = 1, ..., MARKET7 = 1), might also affect mean sales.

In the models of our experimental design, the TEST command in Stata is used to calculate values of statistics to test for differences between two deviations, for example, whether MARKET1 differs from MARKET7, or between early and late markets, namely whether (MARKET1 + MARKET2 + MARKET 3) = 3/4(MARKET4 + MARKET5 + MARKET6 + MARKET7). The statistic in Stata to test for these differences is, given bootstrapped standard errors, an asymptotic χ^2 random variable with one degree of freedom, which is equivalent to an asymptotically distristributed standard normal. Likewise, the statistic in Stata to test for a deviation from the overall mean–for example, to test whether MARKET1 differs from zero–is an asymptotic standard normal random variable. The probability that values of an asymptotically distributed standard normal are at least as extreme in absolute value as the asymptotic z-value is pr(z) in the relevant tables below.

The models of deviations from population means are estimated with the CNSREG command for the specified constraints and with the "collinear" option in Stata. The standard errors in all of the models are bootstrapped to account for possible non-normality of the error terms. The number of useable bootstrap samples is 1000.[15]

[15] If standard errors are clustered on the unique identity of subjects, a particular bootstrap sample is infreqently dropped because sales of at least one participant are not chosen and the model is not estimable.

The specification of the deviations-from-population-means model differs from the specification of the typical deviations- from-sub-population-means model. The typical model would have a constant that represents the mean of a particular sub-population, such as mean sales in the first market, and have indicator variables for all deviations of means of other sub populations from the constant, such as deviations in mean sales in the other six markets from the first market. In both types of models, one can test whether the mean of one sub population differs from the mean of another sub population, such as whether mean sales in market two differ from mean sales in market six. The test statistics are identical. However, in a deviations-from-population-mean model, one can more easily test whether the mean of a sub population differs from the mean of the overall population, such as whether mean sales in market one differ from mean sales for all seven markets.[16] In short, the deviations-from-population-mean models focus our attention, through the sample mean, on predictions of Hotelling.

5.2 Average Prices

We also test whether, as a consequence of sales of four sellers in four periods of time, average price rises over time when the resource is inter-temporarlly scarce or changes over time when the resource is intertemporally abundant. The null hypothesis is $H_0\colon \mu_t^P \geq \mu_v^P$ if the resource is intertemporally scarce or $H_0\colon \mu_t^P = \mu_v^P$ if the resource is intertemporally abundant for $t =$ 1, 2, and 3 and $v =$ t+1, ..., 4. We also test whether average prices differ from quantitative

In these rare cases, the number of sample replications is increased until 1000 of them are useable.

[16]Suppose two sub-populations of subjects make up the population, such as, male and female subjects or inexperienced and experienced subjects. A test of whether one sub-population differs from another is admittedly more intuitive than a test of whether one sub population differs from the overall population. However, the asymptotic z-statistic for the effect of one characteristic, such as inexperience or male gender, in the deviations-from-population-mean model is the same as the asymptotic z-statistic for the effect of the characteristic in the deviations-from-sub-population mean model.

predictions of our Hotelling-inspired model of four Cournot firms. In particular, the null hypothesis is $H_0\colon \mu_t^P = \mu_t^{P_H}$ for $t = 1, \ldots, 4$.

To test alternative hypotheses about the median price of the resource over time, we use the "signtest" command in Stata. The test for exceedance of a subsequent median price over a previous one is a one-sided test. The test for non-constant prices, that is, for a difference between median prices in two different periods of time, is a two-sided test. As before, the test statistic under any of the null hypotheses has a binomial distribution. In the one-sided test, the probability under the null hypothesis that the number of negative differences in prices, $\mu_t^P - \mu_v^P$, is at least as great as the number of observed negative differences is p1(bi) in our tables. In the two-sided test, the probability under the null that the number of positive differences in prices or the number of negative differences in prices is at least as great as the least likely number of observed positive or negative differences is p(bi). To test alternative hypotheses that the median price differs from the quantitative predictions for a four-seller Cournot oligopoly, we also use Stata's "signtest". The test for a difference between the median price and the prediction is also a two-sided test. The probability that the number of positive differences or negative differences between the median price and Hotelling's prediction is at least as great as the least likely number of observed positive or negative differences is p(bi_H) in our tables. The median difference in prices in our sample is 'medif' in our tables.

To test alternative hypotheses that the mean price rises under resource scarcity or changes under abundance, we ultimately use bootstrap p-values, or achieved significance levels. To do so, we start with the original sample, calculate differences in prices for each group of four sellers, regress differences in prices in any two periods on a constant. The estimated constant (cons.) is the mean difference in prices in the two periods of time. In these models we calculate cluster-robust standard errors in which the cluster is each four-seller group in a particular market and session. The cluster is chosen to account for possible correlation of sales within the four-seller group. We then calculate the square root of a Wald statistic from the original sample. Under $H_0\colon \mu_t^P \leq \mu_v^P$ when the resource is intertemporally scarce or $H_0\colon \mu_t^P = \mu_v^P$ when the resource is not for $t = 1, \ldots, 3$ and $v = t+1, \ldots, 4$, the "Wald t-statistic" from the original sample is $w_{tv}^P \equiv \hat{\mu}_{tv}^P / se_{\hat{\mu}_{tv}^P} \equiv (\hat{\mu}_t^P - \hat{\mu}_v^P)/se_{\hat{\mu}_{tv}^P}$.

We do not directly bootstrap prices because quantities sold determine prices, given demand. Instead, for the b-th bootstrap, we sample with replacement four pairs of quantities sold in periods t and v by subjects in each group-market-session, aggregate the four quantities sold in the two periods to determine prices, calculate the price differences, and regress the differences on a constant. The estimated constant (cons.) is the mean difference in prices in two periods of time. To ensure that sales in two periods pertain to the pair of sales by a particular seller in each group, market, and session, we use the same seed number in Stata to start the bootstrapping. In each bootstrap, the standard errors are calculated based on the number of clusters for the four-seller group in a particular market and session. As before, resampling continues until 1000 usable bootstrap replicates exist. The Wald t-statistic from the b-th bootstrap sample for price differences is $w_{tv}^P(b) \equiv (\hat{\mu}_{tv}^P(b) - \hat{\mu}_{tv}^P)/se_{\hat{\mu}_{tv}^P}(b)$. The terms $\hat{\mu}_{tv}^P$ and $\hat{\mu}_{tv}^P(b)$ are the estimated means of the difference in prices in time t and v in the orginal and b-th bootstrap samples. The terms $se_{\hat{\mu}_{tv}^P}$ and $se_{\hat{\mu}_{tv}^P}(b)$ are the cluster- robust standard errors of the estimated mean differences in the respective samples. The bootstrap p-values for a one-sided test of a rise in the mean price of an intertemporally scarce resource, that is, of a negative difference between the mean price in period t and any subsequent period v, is

$$p1(\mathrm{b}t_0) \equiv \sum_{b=1}^{B} \frac{\mathbf{1} \cdot [w_{tv}^P(b) \leq w_{tv}^P]}{B}.$$

The bootstrap p-value for a two-sided test of any difference between mean price in period t

and any subsequent period v when the resource is not intertemporally scarce is

$$p(\mathrm{b}t_0) \equiv \sum_{b=1}^{B} \frac{\mathbf{1} \cdot [|w_{tv}^P(b)| \geq |w_{tv}^P|]}{B}.$$

To test whether the mean price differs from Hotelling's prediction, we again use bootstrap p-values, but ones based on a different "Wald-t" statistic. We initially regress the prices that four-seller groups generated in the t-th period of time on a constant and, as before, calculate robust standard errors based on each group-market-session cluster. For the b-th bootstrap sample we select with replacement four quantities sold in period t for each cluster of four sellers in each market and session and aggregate the four sales to generate a price for the cluster. We then regress the "bootstrapped" prices on a constant and again calculate standard errors based on the clusters. Under $H_0\colon \mu_t^P = \mu_t^{P_H}$ for $t = 1, \ldots, 4$, the Wald-t statistic from the original sample is $w_t^P \equiv (\hat{\mu}_t^P - \mu_t^{P_H})/se_{\hat{\mu}_t^P}$. The Wald-$t$ statistic from the b-th bootstrap sample is $w_t^P(b) \equiv (\hat{\mu}_t^P(b) - \hat{\mu}_t^P)/se_{\hat{\mu}_t^P}(b)$. The terms $se_{\hat{\mu}_t^P}$ and $se_{\hat{\mu}_t^P}(b)$ are the cluster-robust standard errors of the estimated constants, or means, in the respective samples. The bootstrap p-value for a two-sided test of difference between the mean price and Hotelling's prediction in period t is

$$p(\mathrm{b}t_H) \equiv \sum_{b=1}^{B} \frac{\mathbf{1} \cdot [|w_t^P(b)| \geq |w_t^P|]}{B}.$$

In words, the proportion of bootstrapped t-values at least as extreme in absolute value as the original-sample t-value is p(bt_H). (Put the previous sentence as a footnote in the relevant tables?)

6 Results about Sales

6.1 Are Average Sales Consistent with Hotelling's Predictions?

6.1.1 Intertemporal Scarcity of Resource with 50 Percent Interest

Average quantities of the intertemporally scarce resource that subjects sold fell over the four periods of time in experimental sessions with the high-interest-rate treatment (Table 7). Average–median or mean–sales in the sample seem close to the quantitative predictions of Hotelling as well, especially in the first three periods.

Table 7: Predicted and Actual Sales per Seller of Scarce Resource with $r = 0.50$

period	predicted	mean	std. dev.	min	25th pctl	median	75th pctl	max
$t = 1$	42.77	40.75^g	20.61	0	27	40^g	50	100
$t = 2$	34.15	31.19^g	15.78	0	21	30^g	40	100
$t = 3$	21.23	18.02^g	13.92	0	10	20^g	25	91
$t = 4$	1.85	9.97	13.84	0	0	5	15	95

$N = 664$. An observation is the quantity sold per period by a subject in a four-seller group, market number, and session. Superscript g means that average sales are statistically $g$reater in period t than any subsequent period, as indicated by one-sided tests of medians with p-value less than 0.0001 and one-sided tests of means with p-values associated with asymptotic z-values less than 0.001, 0.000, 0.0035. Superscript H indicates that median or mean sales are statistically consistent with Hotelling's prediction.

Moreoever, we reject the null hypotheses that average sales do not fall over the four periods of time. The p-values of the bionomial and asymptotic z statistics for the one-sided sign tests of medians and means are, under the null hypotheses of non-increase, less than 0.0001. However, even though average–median or mean–sales seem close to Hotelling's predictions in the first three periods, we reject the null hypothesis that average sales equal

Hotelling's prediction in any of the four periods. Nonetheless, we are unable to reject the null hypothesis that mean sales of each of the 65 participants, or subjects with unique identities, were equal to Hotelling's prediction for 28 (0.43 percent) of the participants in the first period, 39 (0.60 percent) of them in the second period, 37 (0.57 percent) of them in the third period, and 29 (0.45 percent) of them in the fourth period. Consistent with another prediction of our model, sellers also exhausted their endowment of the intertemporally scarce resource in 661, or 99.5 percent, of the 664 markets in which they participated under the high-r treatment.

6.1.2 Intertemporal Scarcity of Resource with 25 Percent Interest

Average quantities of the intertemporally scarce resource that a subject sold with the low-interest-rate treatment also fell over time (Table 8). Average sales in the sample seem closer to Hotelling's predictions with the low-r treatment than with the high-r treatment.

Table 8: Predicted and Actual Sales per Seller of Scarce Resource with $r = 0.25$

period	predicted	mean	std. dev.	min	25th pct	median	75th pct	max
$t = 1$	35.72	35.52^{gH}	17.48	0	25	35^{gH}	45	100
$t = 2$	29.65	30.05^{gH}	15.63	0	20	30^{gH}	40	100
$t = 3$	22.06	20.23^{g}	15.25	0	10	20^{g}	30	90
$t = 4$	12.57	14.18	17.06	0	0	10	20	100

$N = 532$. An observation is the quantity sold per period by a subject in a four-seller group, market number, and session. Superscript g indicates that average sales are statistically $g$reater in period t than any subsequent period, as indicated by one-sided tests of medians with p-value less than 0.0001 and one-sided tests of means with p-values associated with asymptotic p-values less than 0.001, 0.000, 0.0035. Superscript H indicates that mean or median sales are statistically consistent with Hotelling's prediction.

We reject the null hypotheses that average sales do not fall over the four periods of time. The p-values of the bionomial and asymptotic z statistics are, under the null hypotheses of non-increase, less than 0.001 for the one-sided sign tests of medians and means.[17] Moreover, we fail to reject the null hypothesis that mean or median sales equal Hotelling's prediction in the first and second periods, even though average sales also seem close to Hotelling's prediction in the third and fourth periods. We also fail to reject the null hypothesis that mean sales of each of the 52 participants, or subjects with unique identities, equal Hotelling's prediction for 30 (0.58 percent) of the participants in the first period, 37 (0.71 percent) of them in the second period, 34 (0.65 percent) of them in the third period, and 27 (0.52 percent) of them in the fourth period. Sellers exhausted their endowment of the intertemporally scarce resource in 530, or 99.6 percent, of the 532 markets in which they participated under the low-r treatment.

6.1.3 Intertemporal Abundance of Resource with 50 Percent Interest

The mean of sales in the first period was one unit less than the mean of sales in the second or third period (Table 9). The mean of sales in the fourth period was four units less than the mean of sales in the second and third periods and three units less than the mean of sales in the first period. Average quantities that a subject sold did not fall over the first three periods but fell slightly in the fourth period if the resource was intertemporally abundant (Table 9). The median of sales equaled the prediction of the static model of four Cournot firms in each of the three initial periods. The mean of sales in each of three initial periods

[17]These conclusions are based on one-sided z-tests with bootstrapped standard errors, resampling that is based on the identity of subjects and, thereby, that generates clustered standard errors, and p-values less than 0.001, 0.000, 0.0035 for the low-r treatments. We reach the same conclusion about mean sales falling over the four periods with z-tests from models of seemingly unrelated regressions with standard errors clustered on the identity of the subject.

was close to the prediction but 4.4 units less than the prediction in the final period (Table 9). Average sales in the sample seem close to Hotelling's quantitative predictions of constant sales for a resource that is inter-temporally abundant.

Table 9: Predicted and Actual Sales per Seller of Abundant Resource with $r = 0.50$

period	predicted	mean	std. dev.	min	25th pct	median	75th pct	max
$t = 1$	60.00	58.65^{sH}	21.72	0	42	60^{sH}	75	110
$t = 2$	60.00	59.67^{sH}	20.52	0	45	60^{sH}	75	150
$t = 3$	60.00	59.68^{sH}	20.68	0	45	60^{sH}	70	180
$t = 4$	60.00	55.57	21.68	0	40	56	69	150

$N = 280$. Superscript s indicates that average sales in the first three periods do not statistically differ from each other, even if $\alpha = 0.1$ for two-sided tests. Superscript H indicates that average sales in period t do not statistically differ from Hotelling's prediction of 60 units, even if $\alpha = 0.2$ for two-sided tests

We fail to reject the null hypothesis that median sales of the intertemporally abundant resource are the same in each of the three initial periods. We reach this conclusion based on two-sided p-values that range between 0.50 and 1.00 for signs tests. We reject, however, the null hypothesis that median sales in the first, second, and third period are each the same as median sales in the fourth period based on two-sided p-values of 0.0101, 0.0026, or 0.0197 for the respective sign tests.

Similarly, we fail to reject the null hypothesis that mean sales of the abundant resource are the same in each of the three initial periods of time, given two-sided p-values of 0.536, 0.667, 0.997 for asymptotic z-statistics whose standard errors were bootstrapped with replications clustered on subject identity. Furthermore, mean sales in the first period do not statistically differ from mean sales in the fourth period because the two-sided p-value is 0.323. If the level of statistical significance is 0.10, we conclude that mean sales in the second period and third period differ from mean sales in the fourth period because the two-sided p-values are 0.083 and 0.081. Sellers did not exhaust their endowment of the intertemporally abundant resource in 241, or 86.1 percent, of the 280 4-period markets in which they participated.

In short, the average sales fall over time if the resource is intertertemporally scarce but do not fall over time in three of the four periods if the resource is intertemporally abundant.

6.2 How Do Mean Sales of Males and Females Compare?

6.2.1 Intertemporal Scarcity of Resource with 50 Percent Interest

Mean sales of male sellers were higher in the first three periods and, as a result, lower in the fourth period than the mean sales of female sellers in the corresponding periods (Table 10). In Table 10 the estimated constant is the sample mean for all genders. The sum of the estimated constant and the estimated coefficient for MALE or FEMALE is, except for rounding differences, the mean sales of the associated gender. Mean sales of males were closer to predictions of the four-seller Cournot adaptation of Hotelling's model than mean sales of females were (Table 10).

The differences between mean sales of males and females are statistically significant in the second and fourth periods, but the significance is weak in the second period (Table 10). Mean sales of males and females do not signficantly differ in the first and third periods (Table 10). Equally important is that mean sales of females are signficantly different from predictions of the four-seller Cournot model in the second, third, and fourth periods (Table 10). Mean sales of male sellers are signficantly different from Hotelling's predictions in the third and fourth periods, but the significance in the third period is weak (Table 10).

Table 10: Mean Sales and Deviations by Gender and Participant with a High-r Treatment

period	variable	pred.	mean	coef.	bse	$p(\mathrm{a}z_0)$	$p(\mathrm{a}z_H)$	rse	$p(\mathrm{bt}_0)$	$p(\mathrm{bt}_H)$
$t = 1$	Constant			40.75	1.82	0.000	0.268	1.81	0.000	0.330
	MALE	42.77	41.33	0.58	1.30	0.658	0.584	1.35	0.669	0.633
	FEMALE		39.87	-0.88	1.98		0.161	2.05		0.174
	Constant (and ID1 ... ID65)			40.75	0.79	0.000	0.011	0.69	0.000	0.017
$t = 2$	Constant	34.15		31.19	1.13	0.000	0.009	1.13	0.000	0.021
	MALE		32.69	1.50	0.84	0.064	0.372	0.84	0.093	0.402
	FEMALE		28.90	-2.29	1.28		0.0001	1.28		0.000
	Constant (and ID1 ... ID65)			31.19	1.13	0.000	0.009	1.13	0.000	0.021
$t = 3$	Constant	21.23	18.02	18.02	1.11	0.000	0.004	1.07	0.000	0.009
	MALE		18.40	0.39	0.81	0.630	0.074	0.81	0.654	0.085
	FEMALE		17.42	-0.59	1.23		0.005	1.23		0.014
$t = 4$	Constant	1.85	9.97	9.97	0.98	0.000	0.000	0.98	0.000	0.000
	MALE		7.55	-2.42	0.79	0.002	0.000	0.83	0.004	0.000
	FEMALE		13.67	3.69	1.21		0.000	1.27		0.000

$N = 401 + 263 = 664$ per time period by a male or female in a group of four sellers, market number, and session. The estimated coefficient (coef.) for MALE or FEMALE is the mean deviation of the gender's sales from the estimated constant, or sample mean. Bootstrapped and robust standard errors (bse and rse) are clustered on the identity of participants. Probabilities of z values more extreme than asymptotic z-statistics calculated under null hypotheses that a coefficient equals zero and that a mean of a gender equals Hotelling's prediction (pred.) are $p(\mathrm{a}z_0)$ and $p(\mathrm{a}z_H)$. Proportions of bootstrapped t values more extreme than the original t statistics calulated under the same respective null hypotheses are $p(\mathrm{bt}_0)$ and $p(\mathrm{bt}_H)$.

6.2.2 Intertemporal Scarcity of Resource with 25 Percent Interest

Mean sales of male sellers were lower in the first and second periods but higher in the third and fourth periods than the mean sales of female sellers in the corresponding periods (Table 11). In Table 11 the estimated constant is the sample mean for both genders. The sum of the estimated constant and cofficient for MALE or FEMALE is, except for rounding differences, the mean sales of the associated gender. Mean sales of males were quite close to predictions of the four-seller Cournot model and closer than mean sales of females were in the first three periods.

The difference between mean sales of males and females is statistically significant in the second and fourth periods (Table 11). Mean sales of males and females do not signficantly differ in the first period (Table 11). More importantly, means sales of females are signficantly different from predictions of the four-seller Cournot model in the second, third, and fourth periods. Mean sales of male sellers are signficantly different from the predictions only in the fourth period.

6.2.3 Intertemporal Abundance of Resource with 50 Percent Interest

Mean sales of male sellers were higher in the first and second periods but lower in the third and fourth periods than the mean sales of female sellers in the corresponding periods (Table 12). In Table 12 the sum of the estimates of constant and the cofficient for MALE or FEMALE is, except for rounding differences, the mean sales of the associated gender. Mean sales of males in the first period were close to the prediction of the static model of four Cournot firms as were the mean sales of each gender in the second and third periods. Mean sales of each gender were not as close to the prediction in the fourth period as in previous periods (Table 12).

Table 11: Mean Sales and Deviations by Gender with Low-r Treatment

period	variable	pred.	mean	coef.	bse	$p(\mathrm{a}z_0)$	$p(\mathrm{a}z_H)$	rse	$p(\mathrm{b}t_0)$	$p(\mathrm{b}t_H)$
	Constant	35.72		35.52	1.47	0.000	0.894	1.47	0.000	0.893
$t = 1$	MALE		34.38	-1.15	0.62	0.064	0.441	0.63	0.096	0.480
	FEMALE		39.83	4.31	2.33		0.086	2.37		0.133
	Constant	29.65		30.05	1.25	0.000	0.749	1.28	0.000	0.747
$t = 2$	MALE		28.73	-1.32	0.84	0.118	0.462	0.82	0.272	0.469
	FEMALE		34.99	4.94	3.16		0.157	3.07		0.380
	Constant	22.06		20.23	1.23	0.000	0.137	1.24	0.000	0.146
$t = 3$	MALE		21.83	1.59	0.56	0.004	0.871	0.56	0.036	0.881
	FEMALE		14.26	-5.97	2.10		0.001	2.09		0.066
	Constant	12.57		14.18	1.54	0.000	0.297	1.58	0.000	0.325
$t = 4$	MALE		15.05	0.87	0.75	0.248	0.171	0.74	0.288	0.189
	FEMALE		10.92	-3.26	2.82		0.584	2.78		0.631

$N = 420 + 112 = 532$ sales per time period by a male or female subject in a group of four sellers, market number, and session. The estimated coefficient (coef.) for MALE or FEMALE is the mean deviation of the gender's sales from the estimated constant, or sample mean. Bootstrapped and robust standard errors (bse and rse) are clustered on the unique identity of subjects. Probabilities of z-values more extreme than asymptotic z-statistics calculated under null hypotheses that a coefficient equals $\underline{0}$ and that a mean of one or both genders equals $\underline{H}$otelling's prediction (pred.) are $p(\mathrm{a}z_0)$ and $p(\mathrm{a}z_H)$. Proportions of bootstrapped t values more extreme than the original t-statistics calculated under the same respective null hypotheses are $p(\mathrm{b}t_0)$ and $p(\mathrm{b}t_H)$.

Table 12: Mean Sales and Deviations by Gender with High-r Control

period	variable	pred.	mean	coef.	bse	$p(\mathrm{a}z_0)$	$p(\mathrm{a}z_H)$	rse	$p(\mathrm{bt}_0)$	$p(\mathrm{bt}_H)$
	Constant	60		58.65	2.92	0.000	0.643	2.94	0.000	0.633
$t = 1$	MALE		61.89	3.25	2.85	0.255	0.653	2.94	0.288	0.651
	FEMALE		55.40	-3.25	2.85		0.245	2.94		0.288
	Constant	60		59.67	2.43	0.000	0.891	2.38	0.000	0.886
$t = 2$	MALE		60.93	1.26	2.32	0.587	0.788	2.38	0.591	0.791
	FEMALE		58.41	-1.26	2.32		0.626	2.38		0.622
	Constant	60		59.68	1.89	0.000	0.864	1.850	0.000	0.885
$t = 3$	MALE		58.64	-1.03	1.82	0.570	0.579	1.850	0.582	0.619
	FEMALE		60.71	1.03	1.82		0.800	1.850		0.795
	Constant	60		55.57	2.15	0.000	0.039	2.208	0.000	0.078
$t = 4$	MALE		54.87	-0.70	2.22	0.752	0.102	2.208	0.756	0.163
	FEMALE		56.27	0.70	2.22		0.220	2.208		0.263

$N = 140 + 140 = 280$ sales per time period by a male or female subject in a group of four sellers, market number, and session. The estimated coefficient (coef.) for MALE or FEMALE is the mean deviation of the gender's sales from the estimated constant, or sample mean. Bootstrapped and robust standard errors (bse and rse) are clustered on the unique identity of subjects. Probabilities of z-values more extreme than asymptotic z-statistics calculated under null hypotheses that a coefficient equals zero and that a mean of a gender equals $\underline{H}$otelling's prediction (pred.) are $p(\mathrm{a}z_0)$ and $p(\mathrm{a}z_H)$. Proportions of bootstrapped t values more extreme than the original sample t statistics calculated under the same respective null hypotheses are $p(\mathrm{bt}_0)$ and $p(\mathrm{bt}_H)$.

Mean sales of males do not statistically differ from mean sales of females in any period (Table 12). Neither the mean sales of males nor the mean sales of females is signficantly different from the prediction of the four-seller Cournot model in any period, although the mean sales of all sellers does differ from the prediction in the fourth period (Table 12).

6.3 Do Mean Sales Vary During and Across Sessions?

Could our results be systematically affected by our experimental design? We describe differences in sample means and test for systematic deviations in sales during seven four-period markets in a session and, for some subjects, in multiple sessions. Recall that each market has four time periods and the seven markets with potential financial remunerations are replications of a warm-up market without remuneration.

6.3.1 Intertemporal Scarcity of Resource with 50 Percent Interest

In (Table 13) the sum of the estimated constant and any other estimated coefficient equals, except for rounding error, the mean for the indicated stratum. Mean sales in the first and second periods usually rose during a session after the first market (Table 13). Mean sales in the third period did not exhibit a trend during a session. Mean sales in the fourth period fell over the seven markets during a session. Mean sales of inexperienced sellers were slightly lower in the first and second periods and thus slightly higher in the third and fourth periods than mean sales of experienced sellers (Table 13).

Deviations of stratal means from the overall mean sales in any time period are rarely statistically significant, however (Table 13). Significant deviations exist in only 3 of 28 unique cases at $\alpha = 0.05$ and 3 additional cases at $\alpha = 0.10$. These six signficiant deviations occur within sessions. Mean sales during inexperienced sessions do not significantly differ

Table 13: Stratal Means and Deviations from Overall Mean of Sales with High-r Treatment ($N = 664$)

variable	mean	coef.	bse	z-value	p-value	mean	coef.	bse	z-value	p-value
	First Period (t = 1)					Second Period (t = 2)				
Constant		40.75	1.83	22.28	0.000		31.19	1.15	27.19	0.000
MARKET1	36.41	-4.35	1.48	-2.94	0.003	30.08^e	-1.11	1.13	-0.98	0.330
MARKET2	40.39	-0.37	1.44	-0.25	0.799	32.05^e	0.86	1.26	0.68	0.495
MARKET3	42.35	1.60	1.38	1.16	0.247	28.20^e	-2.99	1.20	-2.49	0.013
MARKET4	42.53	1.78	1.19	1.49	0.136	31.05	-0.14	1.15	-0.12	0.904
MARKET5	41.03	0.28	1.26	0.22	0.825	32.86	1.67	1.18	1.42	0.156
MARKET6	41.36	0.61	1.66	0.37	0.711	31.97	0.78	1.46	0.53	0.593
MARKET7	41.23	0.48	1.94	0.25	0.808	32.19	1.00	1.48	0.68	0.496
INEXP	40.16	-0.60	1.05	-0.57	0.57	30.76	-0.43	0.58	-0.74	0.46
EXPER	41.73	0.97	1.72	0.57	0.57	31.90	0.71	0.95	0.74	0.46
	Third Period (t = 3)					Fourth Period (t =4)				
Constant		18.02	1.09	16.48	0.000		9.97	1.08	9.26	0.000
MARKET1	18.98	0.96	1.23	0.78	0.433	14.27^e	4.30	1.36	3.15	0.002
MARKET2	16.74	-1.28	0.93	-1.38	0.169	10.70^e	0.73	1.07	0.68	0.497
MARKET3	19.19	1.17	1.10	1.07	0.286	10.26^e	0.29	1.16	0.25	0.804
MARKET4	17.76	-0.25	1.17	-0.22	0.828	8.66	-1.32	0.79	-1.67	0.094
MARKET5	17.71	-0.31	1.01	-0.30	0.762	8.29	-1.68	0.91	-1.84	0.066
MARKET6	17.03	-0.98	1.26	-0.78	0.436	9.64	-0.34	1.31	-0.26	0.797
MARKET7	18.76	0.75	1.18	0.63	0.528	7.82	-2.15	1.13	-1.91	0.056
INEXP	18.48	0.46	0.68	0.69	0.493	10.50	0.52	0.69	0.76	0.446
EXPER	17.26	-0.76	1.10	-0.69	0.493	9.12	-0.85	1.12	-0.76	0.446

The sum of the estimates of the constant and coefficient (coef.) for a stratal variable equals, except for rounding error, the mean of the variable. Each of the four deviation-from-mean models is also estimated with 14 pairs of interacted stratal variables, such as MARKET1·INEXP. Bootstrapped standard errors (bse) are clustered on the unique identity of subjects. The probability of observing a value as extreme as the z-value from an asymptotically standard normal is the p-value. Superscript e indicates that the mean of the sales of the three early markets statistically differ from the mean of the sales of the four late markets.

from mean sales during experienced sessions in any period of time. Moreover and for related reasons, deviations of stratal means from each other within a session are often statistically insignificant. Finally, mean sales in early markets of a session, the three initial replications, do not signficantly differ from means sales in late markets of a session, the four subsequent replications, in the first and third periods of time.[18] In short, the market number within a session rarely affects, in a statistically significant way, mean sales of the intertemporally scarce resource with an interest rate of 50 percent and repeated selling of the resource in successive sessions never does (Table 13).

6.3.2 Intertemporal Scarcity of Resource with 25 Percent Interest

Mean sales in any market number seemed to randomly differ from sales in other market numbers in the first three periods (Table 14). In other words, mean sales during the first three periods varied without any discernible pattern over the seven markets in a session. However, mean sales in the last period usually increased during a session. Mean sales of inexperienced sellers were slightly lower in the first and second periods and thus slightly higher in the third and fourth periods than respective mean sales of experienced sellers (Table 14).

Almost all deviations of stratal means from the overall mean of sales in any time period are not statistically significant. One statistically significant deviation of the 28 unique deviations exists if $\alpha = 0.05$ and two more exist if $\alpha = 0.10$. All three signficant differences occur within a session. As a result, sales in one market during a session usually do not significantly differ from sales in other markets. However, as occurs with the high-r treatment, the mean of sales

[18]The asymptotic chi-square statistics with one degree of freedom under the null hypothesis of no difference between sales in early and late markets are 2.12 and 0.33 for the first and third period and the associated p-values are 0.145 and 0.564.

Table 14: Stratal Means and Deviations from Overall Mean of Sales with Low-r Treatment ($N = 532$)

variable	mean	coef.	bse	z-value	p-value	mean	coef.	bse	z-value	p-value
	First Period (t = 1)					Second Period (t = 2)				
Constant		35.52	1.46	24.38	0.000		30.05	1.38	21.74	0.000
MARKET1	34.45	-1.08	1.80	-0.60	0.549	32.61^e	2.55	1.69	1.51	0.131
MARKET2	37.43	1.91	1.60	1.19	0.233	31.55^e	1.50	1.74	0.86	0.389
MARKET3	34.88	-0.64	1.29	-0.50	0.619	29.29^e	-0.76	1.40	-0.54	0.587
MARKET4	33.72	-1.80	1.22	-1.47	0.141	31.13	1.08	1.29	0.84	0.402
MARKET5	37.46	1.94	1.56	1.24	0.214	27.28	-2.77	1.29	-2.15	0.032
MARKET6	33.84	-1.68	1.02	-1.65	0.099	30.49	0.44	1.16	0.38	0.707
MARKET7	36.88	1.36	1.61	0.84	0.400	28.01	-2.04	1.15	-1.77	0.077
INEXP	35.32	-0.21	0.87	-0.24	0.813	29.72	-0.33	0.70	-0.48	0.633
EXPER	35.97	0.45	1.88	0.24	0.813	30.77	0.72	1.52	0.48	0.633
	Third Period (t = 3)					Fourth Period (t = 4)				
Constant		20.23	1.31	15.41	0.000		14.18	1.57	9.03	0.000
MARKET1	21.41	1.17	1.53	0.77	0.443	11.54^e	-2.64	1.69	-1.56	0.118
MARKET2	18.49	-1.75	1.34	-1.31	0.192	12.53^e	-1.65	1.32	-1.25	0.211
MARKET3	22.14	1.91	1.28	1.50	0.135	13.68^e	-0.49	1.22	-0.41	0.684
MARKET4	19.74	-0.50	1.26	-0.39	0.694	15.38	1.20	1.17	1.03	0.303
MARKET5	19.55	-0.68	1.36	-0.50	0.617	15.64	1.47	1.15	1.27	0.203
MARKET6	20.46	0.23	1.22	0.19	0.852	15.21	1.03	1.29	0.80	0.422
MARKET7	19.84	-0.39	1.50	-0.26	0.794	15.26	1.08	1.43	0.76	0.448
INEXP	20.32	0.09	0.60	0.14	0.886	14.63	0.45	0.89	0.50	0.617
EXPER	20.05	-0.19	1.29	-0.14	0.886	13.21	-0.97	1.94	-0.50	0.617

The sum of the estimates of the constant and coefficient (coef.) for a stratal variable equals, except for rounding error, the mean of the variable. Each of the four deviation-from-mean models is also estimated with 14 pairs of interacted stratal variables, such as MARKET1·INEXP. Bootstrapped standard errors (bse) are clustered on the unique identity of subjects. The probability of observing a value as extreme as the z-value from an asymptotically standard normal is the p-value. Superscript e indicates that the mean of the sales of the three early markets statistically differ from the mean of the sales of the four late markets.

in the three early markets signficantly differs from the mean of sales in the four late markets in the second ($p = 0.046$) and fourth ($p = 0.053$) periods (Table 14). As also occurs with the high-r treatment, mean sales during inexperienced sessions do not significantly differ from mean sales during experienced sessions in any period of time with the low-r treatment (Table 14).

6.3.3 Intertemporal Abundance of Resource with 50 Percent Interest

Mean sales in the first period varied without any discernible pattern over the seven markets of a session (Table 15). Mean sales in the second and third periods were higher at the start of a session than at the end. Mean sales in the fourth period tended to increase over the seven replications in a session. The deviation of mean sales in the first market from the overall mean is the largest deviation in the second and fourth periods, the second largest one in the third period, and the third largest in the first period of time. The bootstrapped standard errors of the mean deviations are largest in the first market in all time periods. These errors also tend to become smaller over the seven markets in all time periods (Table 15).

In spite of variations of subsample means during the seven markets of a session, the mean deviations in sales usually do not significantly differ from the overall mean in any time period (Table 15). The mean of sales in a market number, or replication number, significantly differs from the overall mean in only 5 of the 21 market numbers during the first three periods at $\alpha = 0.05$ and 2 of the 7 market numbers in the fourth period at $\alpha = 0.10$. Moroever, the mean of sales in early markets significantly differs from the mean of sales in late markets in only the fourth period of time (p-value $= 0.007$). In other words, the mean of sales in the first three markets does not significantly differ from mean of sales in the last four markets in the first (p-value $= 0.61$), second (p-value $= 0.60$), and third (p-value $= 0.21$) periods

Table 15: Stratal Means and Deviations from Overall Mean of Sales with High-r Control ($N = 280$)

variable	mean	coef.	bse	z-value	p-value	mean	coef.	bse	z-value	p-value
	First Period (t = 1)					Second Period (t = 2)				
Constant		58.65	2.85	20.6	0.000		59.67	2.29	26.0	0.000
MARKET1	56.83	-1.82	2.62	-0.70	0.487	65.20	5.53	3.73	1.48	0.138
MARKET2	57.60	-1.05	2.58	-0.41	0.685	57.58	-2.09	2.33	-0.90	0.370
MARKET3	63.85	5.20	2.17	2.39	0.017	58.38	-1.29	2.46	-0.53	0.599
MARKET4	56.40	-2.25	2.38	-0.94	0.345	62.55	2.88	2.05	1.41	0.160
MARKET5	58.25	-0.40	1.76	-0.23	0.822	59.48	-0.19	2.37	-0.08	0.935
MARKET6	57.73	-0.92	1.63	-0.56	0.572	59.65	-0.02	1.98	-0.01	0.993
MARKET7	59.88	1.23	2.12	0.58	0.562	54.85	-4.82	1.87	-2.58	0.010
INEXP	58.93	0.28	1.65	0.17	0.864	60.65	0.98	1.37	0.72	0.474
EXPER	57.99	-0.66	3.85	-0.17	0.864	57.38	-2.29	3.19	-0.72	0.474
	Third Period (t = 3)					Fourth Period (t = 4)				
Constant		59.68	1.86	32.1	0.000		55.57	2.22	25.1	0.000
MARKET1	63.85	4.18	4.41	0.95	0.344	48.30^e	-7.27	4.73	-1.54	0.124
MARKET2	59.33	-0.35	2.26	-0.15	0.877	54.33^e	-1.25	2.58	-0.48	0.628
MARKET3	61.78	2.10	2.59	0.81	0.418	50.75^e	-4.82	2.54	-1.90	0.058
MARKET4	64.20	4.53	2.31	1.96	0.050	59.20	3.63	2.59	1.40	0.162
MARKET5	57.98	-1.70	2.28	-0.75	0.455	57.15	1.58	2.28	0.69	0.489
MARKET6	54.88	-4.80	1.85	-2.59	0.010	58.55	2.98	1.92	1.55	0.120
MARKET7	55.73	-3.95	1.64	-2.40	0.016	60.73	5.15	2.75	1.87	0.061
INEXP	60.05	0.37	0.87	0.42	0.671	55.75	0.18	1.23	0.15	0.885
EXPER	58.81	-0.87	2.04	-0.42	0.671	55.15	-0.42	2.87	-0.15	

The sum of the estimates of the constant and coefficient (coef.) for a stratal variable equals, except for rounding error, the mean of the variable. Each of the four deviation-from-mean models is also estimated with 14 pairs of interacted stratal variables, such as MARKET1·INEXP. Bootstrapped standard errors (bse) are clustered on the unique identity of subjects. The probability of observing a value as extreme as the z-value from an asymptotically standard normal is the p-value. Superscript e indicates that the mean of the sales of the three early markets statistically differs from the mean of the sales of the four late markets.

of time. Finally, as also occurs with the two treatments, mean sales during inexperienced sessions do not significantly differ from mean sales during experienced sessions in any period of time with the high-r control (Table 15).

7 Results about Prices

7.1 Intertemporal Scarcity of Resource with 50 Percent Interest

Average prices of the intertemporally scarce resource with a 50 percent interest rate rose over time (Figure 1). Mean and median prices were greater than the predicted prices in the first, second, and third periods but were less than the predicted price in the fourth period (Figure 1).[19] The smallest and largest standard deviations are 26 and 40 experimental credits per unit of the resource in the third and first periods (Table 20 in Appendix D.1). The smallest and largest interquartile ranges are 34 and 55 experimental credits per unit of the resource in the third and first periods (Figure 1).

Mean and median prices are signficantly higher in a subsequent time period than any previous one.[20] However, mean and median prices significantly differ in a statistical sense from the quantitative predictions of the Hotelling-inspired model (Table 16).

7.2 Intertemporal Scarcity of Resource with 25 Percent Interest

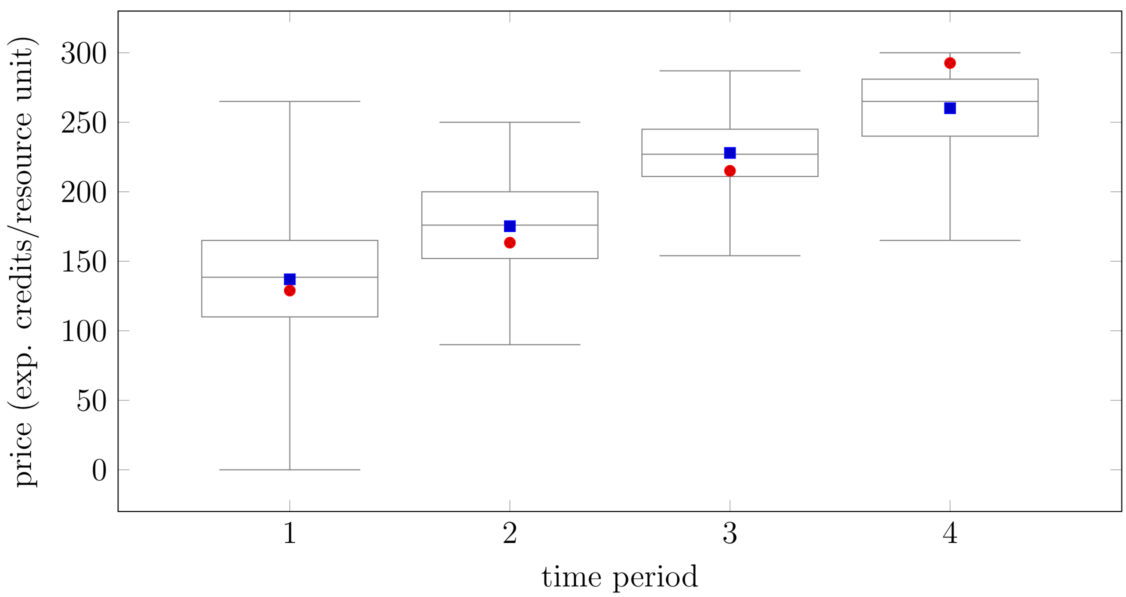
Average prices rose over time (Figure 2). In fact, mean and median prices in each period were close to the predicted prices, particularly in the first two periods (Figure 2).[21] The

[19]See Table 20 in Appendix D.1 for the data on which the box-whisker plot is based.

[20]Inferences are based on probabilities less than 0.000 for one-sided tests with asymptotic z-values and with bootstrapped t-values. See Table 21 in Appendix D.1 for details.

[21]See Table 22 in Appendix D.2 for the data on which the box-whisker plot is based.

Figure 1: Box-Whisker Plot of Prices of Scarce Resource for $r = 0.50$



$n = 166$. An observation is the equilibrium price (experimental credits per unit of the resource) at any time period in a four-seller group, market number, and session. Blue squares are means and red circles are predicted prices. The horizontal bar in each box is the median.

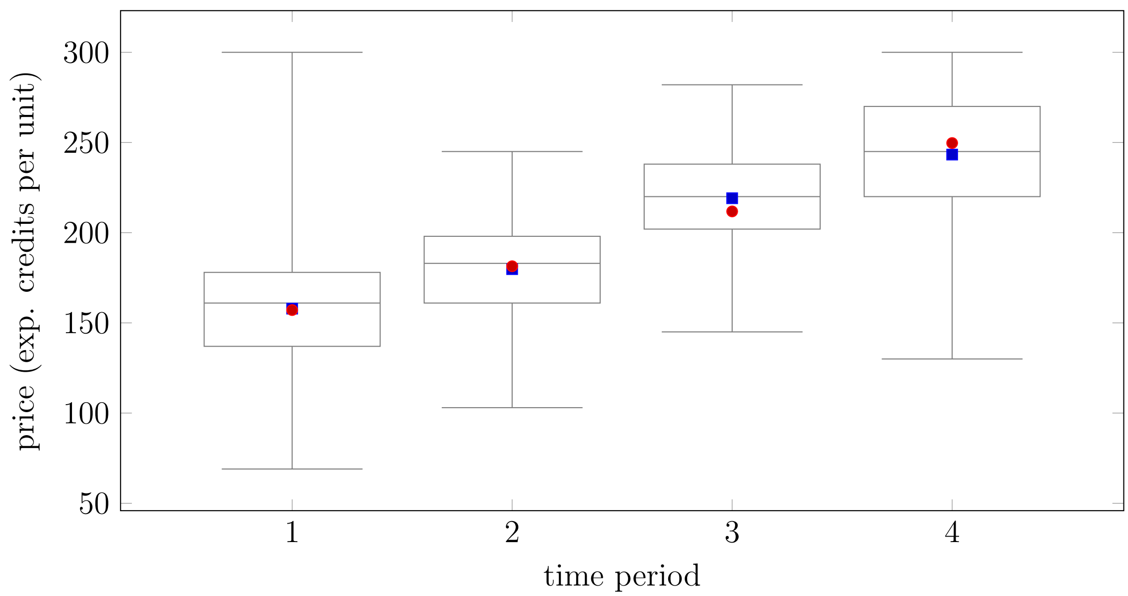
Table 16: Tests for Differences between Predicted and Average Prices of Scarce Resource for $r = 0.50$

period	pred.	mean	cons.	bse	p(z)	rse	p(bt_H)	median	p(Bi_H)
$t = 1$	128.92	137.02	8.10	2.569	0.001	3.106	0.0000	138.5	0.004
$t = 2$	163.38	175.24	11.86	2.074	0.000	2.505	0.000	176	0.000
$t = 3$	215.08	227.94	12.86	1.876	0.000	2.037	0.000	227	0.000
$t = 4$	292.62	260.11	-32.51	1.779	0.000	2.305	0.000	265	0.000

$n = 166$. An observation is the equilibrium price (experimental credits per unit of the resource) at a period of time in a four-seller group, market number, and session. The estimated constant (cons.) is the difference between the mean price and the predicted (pred.) price. Bootstrapped standard errors (bse) and robust standard errors (rse) are used, along with 'cons.', to calculate z- and t-values. The probability that values of an asymptotically distributed standard normal are at least as extreme in absolute value as the asymptotic z-value is p(z). The proportion of bootstrapped t-values at least as extreme in absolute value as the original-sample t-value is p(bt). In the two-sided sign test, the binomial probability that positive or negative differences in actual and predicted prices are at least as unlikely as the less likely number of observed positive or negative differences is pr(Bi).

smallest and largest standard deviations of prices were 25 and 34 experimental credits per unit of the resource in the third and first periods (Table 22 in Appendix **??**). The smallest and largest interquartile ranges were 36 and 50 experimental credits per unit of the resource in the third and fourth periods (Figure **??**).

Figure 2: Predicted and Box-Whisper Plot of Prices of Scarce Resource for $r = 0.25$



$n = 133$. An observation is the equilibrium price (experimental credits per unit of the resource) at a period of time in a four-seller group, market number, and session. The blue squares are sample means and the red circles are predicted prices of the sub-game prefect equilibrium.

Mean and median prices are significantly higher in each successive period of time.[22] Moreover, mean and median prices in the first and second periods do not significantly differ, in a statistical sense, from the quantitative predictions of the Hotelling-inspired model of four Cournot sellers (Table 17). However, mean and median prices are signficantly different from Hotelling's prediction in the third period. Although the mean price does significantly differ from the predicted price in the fourth period, the median price does not significantly

[22]See Table 23 in AppendixD.2 for details.

differ it (Table 17).

Table 17: Tests for Differences between Predicted and Average Prices of Scarce Resource for $r = 0.25$

period	pred.	mean	cons.	bse	pr(z)	rse	p(bt_H)	median	p(bi_H)
$t = 1$	157.13	157.90	0.772	2.64	0.770	2.92	0.704	161	0.298
$t = 2$	181.41	179.80	-1.61	2.45	0.511	2.40	0.373	183	0.603
$t = 3$	211.76	219.07	7.31	2.46	0.003	2.20	0.000	220	0.037
$t = 4$	249.70	243.29	-6.41	2.66	0.016	2.76	0.003	245	0.386

$n = 133$. An observation is the equilibrium price (experimental credits per unit of the resource) at a period of time in a four-seller group, market number, and session. The estimated constant (cons.) is the difference between the mean price and the predicted (pred.) price. Bootstrapped standard errors (bse) and robust standard errors (rse) are used to calculate z- and t-values. The probability that values of an asymptotically distributed standard normal are at least as extreme in absolute value as the asymptotic z-value is pr(z). The proportion of bootstrapped t-values at least as extreme in absolute value as the original-sample t-value is pr(b t). In the two-sided sign test, the binomial probability that positive or negative differences in actual and predicted prices are at least as unlikely as the less likely number of observed positive or negative differences is pr(bi).

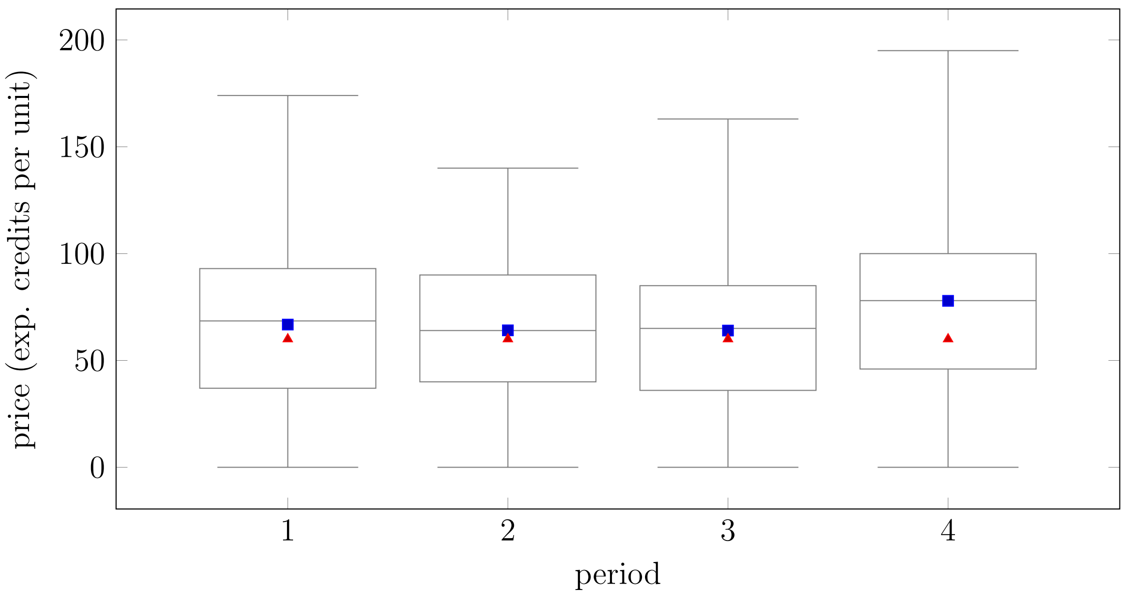
7.3 Intertemporal Abundance of Resource with 50 Percent Interest

Mean and median prices of the intertemporally abundant resource are moderately close to the theoretical predictions in the first period, close to the predictions in the second and third periods, and not close to them in the fourth period (Figure predactprice-ar50).[23]. Deviations from the mean and median prices are large too.

Mean and median prices are not significantly different from each other in the three initial periods of time (Table 18). The mean price in the fourth period significantly differs from

[23]See Table 24 in Appendix D.3

Figure 3: Predicted and Actual Prices of Abundant Resource for $r = 0.50$



$n = 70$. An observation is the equilibrium number of experimental credits per unit of the resource at a period of time in a four-seller group, market number, and session. The red triangles are the predicted prices and the blue squares are the sample means.

each mean price in the first, second, and third periods, as the probabilities associated with the asymptotic z-value and bootstrapped t-values indicate. The median price in the fourth period does not, however, significantly differ from any median price in the three initial periods (Table 18).

Table 18: Statistics and Probabilities for Tests of Changes in Price of Abundant Resource for $r = 0.50$

periods	variable	cons.	bse	pr(z)	rse	pr(b t)	mdnd	pr(bi)
$t = 1, 2$	dprices12	2.71	3.93	0.490	4.96	0.399	0.5	0.904
$t = 1, 3$	dprices13	2.79	4.45	0.531	6.48	0.474	-3.5	0.630
$t = 1, 4$	dprices14	-11.13	5.46	0.042	6.71	0.015	-14.0	0.282
$t = 2, 3$	dprices23	0.071	4.13	0.986	5.60	0.984	0.0	1.000
$t = 2, 4$	dprices24	-13.84	5.22	0.008	6.87	0.002	-5.0	0.630
$t = 3, 4$	dprices34	-13.91	4.77	0.004	7.47	0.001	-9.0	0.550

$n = 70$. An observation is the difference between prices (experimental credits per unit of the resource) in two periods of time for a four-seller group, market number, and session. The estimated constant (cons.) is the mean difference in prices in two periods of time. Bootstrapped standard errors and robust standard errors (rse) are used to calculate z- and t-values. The probability that values of an asymptotically distributed standard normal are at least as extreme in absolute value as the asymptotic z-value is pr(z). The proportion of bootstrapped t-values at least as extreme in absolute value as the original-sample t-value is pr(b t). The median difference in prices is 'mdnd'. In the two-sided sign test, the binomial probability that positive or negative differences in prices are at least as unlikely as the less likely number of observed positive or negative differences is pr(bi).

Moreover, mean and median prices in the first, second, and third periods do not significantly differ, in a statistical sense, from the quantitative predictions of the static model of four Cournot sellers (Table 19). However, mean and median prices are signficantly different from the prediction of the static model in the fourth period (Table 19).

Table 19: Tests for Differences between Predicted and Average Prices of Abundant Resource for $r = 0.50$

period	pred.	mean	cons.	bse	pr(z)	rse	pr(bt)	median	pr(bi)
$t = 1$	60	66.81	6.81	4.15	0.101	4.85	0.065	68.5	0.120
$t = 2$	60	64.10	4.10	3.61	0.256	4.30	0.198	64	0.470
$t = 3$	60	64.03	4.03	3.41	0.238	4.75	0.265	65	0.268
$t = 4$	60	77.94	17.94	3.92	0.000	5.27	0.000	78	0.008

$n = 70$. An observation is the equilibrium price (experimental credits per unit of the resource) in a period of time for a four-seller group, market number, and session. The estimated constant (cons.) is the difference between the mean price and the predicted (pred.) price. Bootstrapped standard errors (bse) and robust standard errors (rse) are used to calculate asymptotic z- and original-sample t-values. The probability that values of an asymptotically distributed standard normal are at least as extreme in absolute value as the asymptotic z-value is pr(z). The proportion of bootstrapped t-values at least as extreme in absolute value as the original-sample t-value is pr(b t). In a two-sided sign test the binomial probability that positive or negative differences in actual and predicted prices are at least as unlikely as the less likely number of observed positive or negative differences in the median and predicted price is pr(bi).

8 Discussion

To what extent are our results consistent with Hotelling's rule and other predictions of a discrete-time version of a Hotelling- inspired model with four Cournot sellers of a non-renewable resource? Degree of consistency is, like beauty, in the eye of the beholder. The null hypotheses that average sales and associated prices in our experimental markets equal the sales and prices that give rise to Hotelling's rule cannot be rejected in the half of the time periods when the resource is scarce with a low interest rate and the initial three periods when the resource is abundant. Obversely, however, the null hypotheses of equality with quantitative predictions are rejected in all periods when the resource is scarce with a high interest rate, half of the periods of the periods when the resource is scarce with a low interest rate, and the last period when the resource is abundant. A majority of rejections of null hypotheses about these quantitative predictions might lead some to conclude inconsistency with Hotelling's rule, as van Veldhuizen and Sonnemans (2018) do in light of their statistical results.

Our average results are unambiguosly consistent, however, with qualitative predictions of our Hotelling-inspired model. In our experiments, as in theory, average individual sales fall and market price rises over time when the resource is inter-temporally scarce. Moreover, the mean and median price of the scarce resource is initially lower but subsequently higher when the interest rate is high than when the interest rate is low. However, when the resource is inter-temporally abundant and its shadow price is zero, average sales and price do not usually change over time, as predicted. Furthermore, mean and median prices of the intertemporally abundant resource are significantly lower, as predicted, than the mean and median prices of the intertemporarlly scarce resource. Finally, individual sellers exhaust their reserves if the resource is intertemporally scarce but do not exhaust them 86 percent of the time if the resource is abundant.

The consistency of average sales and prices with qualitative predictions of Hotelling's model has two proximate causes. First, the mean sales of approximately half of all sellers in the high-interest-rate and low-interest-rate treatments do not statistically differ from Hotelling's predictions for a Cournot seller among three other sellers and, thereby, contribute to average individual behavior that is usually consistent with Hotelling's hypothesis. Second, although variation in quantities sold by other sellers is large, individuals who sell more than the *ex-ante* subgame perfect equilibrium quantities seem to counter balance and be counter balanced by those who sell less than such quantities in many instances. Behavior in static Cournot experiments has also been consistent, on average, with theoretical predictions even while it has exhibited variability around the average (e.g., Holt 1985, pp. 320-323).

Is the consistency between average behavior of sellers in our experiments and qualitative behavior predicted from a Hotelling-inspired model replicable? In recent laboratory experiments with Cournot duopolists who receive a high or low endowment of a non- renewable, inter-temporally scarce resource and who can sell the resource over six periods, mean sales also fall over time (van Veldhuizen and Sonnemans 2018, p. 503). These results resemble the decreases in mean sales under our two treatments. Our experimental results are different, however, because they are based on four rather than two sellers, a high and low interest rate rather than a single one, and a control with inter-temporal abundance, in addition to treatments with scarcity. As such, our results represent new, expanded consistency of average experiemental behavior with qualitative predictions of Hotelling's theory.

Does the consistency of average sales and prices with qualitative predictions of Hotelling's model depend on our experiment design? On the one hand, average sales in our experiments are not likely to be significantly influenced by the market number, or replication number, within a session or whether the session has exclusively experienced sellers. That is, mean sales of the resource, regardless of the interest rate or the inter-temporal scarcity, usually does not significantly vary with successive markets, or replications, within a session and never significantly varies with experience of sellers across sessions. Moreover, average quantities and prices of the inter-temporally abundant resource do not change nor depend on inter-temporal tradeoffs, as predicted, even though the interest rate is 50 percent and half of the content in our written instructions is devoted to explanations of compounding interest at a rate of 50 percent per period. Priming, or influencing behavior in a laboratory by including instructional content that is not relevant for monetary payoffs, has been observed in other economic experiments, however (e.g., Benjamin, Choi, and Strickland 2010; Horton, Rand, and Zeckhauser 2011).

On the other hand, our experimental design has numerous features chosen to increase the likelihood of consistency between sales and theoretical predictions. First, each experimental market comprises four sellers because two sellers tend to collude but four tend to sell at or slightly above non-collusive, Cournot amounts in static Cournot markets (Huck, Normann, and Oechssler 2004). Finitely repeated games with four players are also not particularly conducive environments for cooperation. The time paths of average sales in our experiments do not reflect significant collusive or perfectly competitive behavior. Second, the low and high interest rates in our experiements were chosen for salience in recognition that real-word sellers of oil or other non-newable resources make decisions with multi-million dollar consquences. Third, sellers were recruited to be mathematically sophisticated and often business or economics oriented just as we presume real-world sellers are. Fourth, the inter-temporally scarce endowment of 100 units was partly chosen to permit sellers to think in percentages of the resource and reduce costs of mental processing. Fifth, the inter-temporarlly abundant endowment of 300 units was 60 units, or 25 percent, larger than the minimum endowment of 240 units for inter-temporal scarcity to exist. The large endowment reflects the strong possibility that inter-temporal scarcity has functionally not existed or at least not been salient in markets for some non-renewable resources, such as oil (e.g., Hart and Spiro 2011).

Writing more than 60 years ago, Muth (1961) observed that one major conclusion of studies of expectation data are that “averages of expectations in industries are more accurate than naive models and as accurate as elaborate equation systems, although there are considerable cross-sectional differences of opinion”. He then hypothesized that “expectations of firms (or, more generally, the subjective probability distribution of outcomes) tend to be distributed, for the same information set, about the prediction of the theory (or the ‘objective’ probability distributions of outcomes)” (Muth 1961, p. 316). The average expectations of our subjects, as revealed by their quantity choices, seem consistent with Muth’s observation and hypothesis.

Neither our version nor any other version of Hotelling’s model generates predictions about differences in behavior of male and female sellers. In other laboratory experiments, however, male and female subjects make systematically different choices about social preferences, competition, and risk-taking (e.g., Croson and Gneezy 2009; Niederle 2015). Female sellers do occasionally behave differently from male sellers in our experiments. The differences do not exhibit a universal pattern, however. Also, the differences are not statistically significant in all time periods when the resource is inter-temporally abundant and in half of the time periods when the resource is scarce. Given that Niederle (2015) detects only small differences in risk-aversion by gender, the statistical power of our tests of differences between male and female selling might be low for lack of sufficently large sizes of the subsamples.

9 Conclusion

Our findings prompt us to ask several questions or sets of them for future research. First, which features of our experimental design induce average behavior and prices to be consistent with qualitative predictions of Hotelling's model? Would the average behavior of less mathematically adept and less business-oriented sellers also be consistent with the qualitative predictions? Would professional commodity traders behave, on average, in accordance with the qualitative predictions and would they generate less dispersion in sales and prices? Which new features of an experimental design would induce average behavior and prices to be more consistent with quantitative predictions of the model?

Second, would a futures market for the resource, in addition to the spot market, provide information that reduces the variation in individual behavior and bring average behavior closer to Hotelling's quantitative predictions? How does one create a futures market that operates with a spot market in a laboratory? Although Hotelling argues that an owner of a mine would be indifferent between earning marginal rent from selling of a unit of his resource in the current period and earning marginal rent plus interest on the rent from selling it in the next period (Hotelling 1931, p. 140), he does not describe the institutional details or process through which the owner acquires information to successfully maximize the present-value wealth of the mine. A futures market conveys information that helps traders adjust and readjust their sales. In so doing, future markets enable sellers to make fewer strategic mistakes themselves or mitigate mistakes of others through arbitrage. Creation and operation of a futures market within a laboratory would be challenging tasks. But, once successfully implemented, the futures market would provide feedback that subjects would likely use and, thereby, reduce the variance of their choices.

Third, we observe many sellers whose choices deviate from the subgame perfect equilibrium choices of the Cournot-market game. If sales deviate from the orginal, subgame perfect equilibrium, market prices will probably deviate from Hotelling's price path too. To maximize their wealth sellers should not return to the original price path but instead find an updated one. The updated price path is based on an assumption that sellers will behave rationally after any previous mistake(s) and will return to a Cournot-Nash equilibrium. If sellers initially deviate from the orginal subgame perfect equilibrium and prices deviate from the orginal predictions, to what extent are subsequent sales and associated prices in the laboratory actually optimal, actually consistent with the updated subgame perfect equilibrium? Answering this question for four sellers would be challenging.

Fourth, how would sellers behave if they also chose when to exhaust their resource endowment? In the models of Hotelling (1931) miners choose how long to extract the exhuastible resource to maximize their present-value wealth from its sale. Given the demand function and zero marginal cost of extraction in our experimental markets, four periods of time is the optimal stopping time for four Cournot sellers.

The importance of Hotelling's models, rule, and extensions of his models in economic thought cannot be understated. The rule is derived and explained in all graduate-level textbooks on natural resource economics (e.g., Conrad 2010) and most upper-division undergraduate ones (e.g., Tietenberg and Lewis 2018). The consistency of our average results with qualitative predictions provides a renewed justification for continued teaching of the subject.

We designed our experimental markets to match several aspects of real-world markets, namely that subjects be relatively sophisticated sellers who understand intertemporal tradeoffs and that tradeoffs be salient and strongly incentivized. Given our experimental design, the consistency of our average results with qualitative predictions of the theory also has an important lesson for understanding prices over time in real-world markets for non-renewable resources: inter-temporal scarcity or abundance matters.

References

Baumeister, C., and L. Kilian. 2016. "Forty Years of Oil Price Fluctuations: Why the Price of Oil May Still Surprise Us." *Journal of Economic Perspectives* 30(1):139–160.

Benjamin, D.J., J.J. Choi, and A.J. Strickland. 2010. "Social Identity and Preferences." *American Economic Review* 100:1913–1928.

Cameron, A.C., and D.L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50:317–352.

Casari, M., and C.R. Plott. 2003. "Decentralized Management of Common Property Resources: Experiments with a Centuries-Old Institution." *Journal of Economic Behavior and Organization* 51:217–247.

Cochran, G.W., and W.G. Cochran. 1989. *Statistical Methods, 8th Ed.*. Ames, IA: Iowa State University Press.

Conrad, J.M. 2010. *Resource Economics, 2nd Edition*. New York, NY: Cambridge University Press.

Croson, R., and U. Gneezy. 2009. "Gender Differences in Preferences." *Journal of Economic Literature* 47:1–27.

Cummings, R.G. 1969. "Some Extensions of the Economic Theory of Exhaustible Resources." *Western Economic Journal* 7:201–210.

Duffy, J. 2015. "Macroeconomics: A Survey of Laboratory Research." In J. H. Kagel and A. E. Roth, eds. *The Handbook of Experimental Economics, Volume 2*. Princeton University Press, pp. 1–90.

Efron, B., and R.J. Tibshirani. 1994. *An Introduction to the Bootstrap*. Boca Raton, FL: Chapman and Hall/CRC, an imprint of Taylor and Francis Group, LLC.

Fischbacher, U. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics* 10:171–178.

Fitzgerald, T. 2013. "Frackonomics: Some Economics of Hydraulic Fracturing." *Case Western Reserve Law Review* 63:1337–1362.

Friedman, D., and S. Sunder. 1994. *Experimental Methods: A Primer for Economists*. New York, NY: Cambridge University Press.

Gordon, R.L. 1967. "A Reinterpretation of the Pure Theory of Exhaustion." *Journal of Political Economy* 75:274–286.

Hanson, D.A. 1980. "Increasing Extraction Costs and Resource Prices: Some Further Results." *The Bell Journal of Economics* 11:335–342.

Hart, R., and D. Spiro. 2011. "The Elephant in Hotelling's Room." *Energy Policy* 39:7834–7838.

Heal, G. 1976. "The Relationship between Price and Extraction Cost for a Resource with a Backstop Technology." *The Bell Journal of Economics* 7:371–378.

Hey, J.D., and V. Dardanoni. 1988. "Optimal Consumption Under Uncertainty: An Experimental Investigation." *The Economic Journal* 98:105–116.

Holt, C.A. 1985. "An Experimental Test of the Consistent-Conjectures Hypothesis." *American Economic Review* 75:314–325.

Horton, J.J., D.G. Rand, and R.J. Zeckhauser. 2011. "The Online Laboratory: Conducting Experiments in a Real Labor Market." *Experimental Economics* 14:399–425.

Hotelling, H. 1931. "The Economics of Exhaustible Resources." *Journal of Political Economy* 39:137–175.

Huck, S., H.T. Normann, and J. Oechssler. 2004. "Two Are Few and Four Are Many: Number Effects in Experimental Oligopolies." *Journal of Economic Behavior and Organization* 53:435–446.

Kelley, H., and D. Friedman. 2002. "Learning to Forecast Price." *Economic Inquiry* 40:556–573.

Lewis, T.R., and R. Schmalensee. 1980. "On Oligopolistic Markets for Nonrenewable Natural Resources." *Review of Economics and Statistics* 95:475–491.

Livernois, J. 2009. "On the Empirical Significance of the Hotelling Rule." *Review of Environmental Economics and Policy* 3:22–41.

Livernois, J., and P. Martin. 2001. "Price, Scarcity Rent, and a Modified r per cent Rule for Non-Renewable Resources." *Canadian Journal of Economics* 34:827–845.

Meissner, T. 2016. "Intertemporal Consumption and Debt Aversion." *Experimental Economics* 19:281–298.

Miller, M.H., and C.W. Upton. 1985. "A Test of the Hotelling Valuation Principle." *Journal of Political Economy* 93:1–25.

Miller, R.M., C.R. Plott, and V.L. Smith. 1977. "Intertemporal Competitive Equilibrium: An Empirical Study of Speculation." *Quarterly Journal of Economics* 91:599–624.

Muth, J.F. 1961. "Rational Expectations and the Theory of Price Movements." *Econometrica* 29:315–335.

Neumann, C., and M. Erlei. 2014. "Price Formation of Exhaustible Resources: An Experimental Investigation of the Hotelling Rule." Unpublished, Unpublished manuscript, Energy Research Center of Lower Saxony and Clausthal University of Technology.

Niederle, M. 2015. "Gender." In J. H. Kagel and A. E. Roth, eds. *The Handbook of Experimental Economics, Volume 2*. Princeton University Press, pp. 481–562.

Pindyck, R.S. 1978. "The Optimal Exploration and Production of Nonrenewable Resources." *Journal of Political Economy* 86:841–861.

Salant, S.W. 1976. "Exhaustible Resources and Industrial Structure: A Nash-Cournot Approach to the World Oil Market." *Journal of Political Economy* 84:1079–1093.

Schulze, W.D. 1974. "The Optimal Use of Non-Renewable Resources: The Theory of Extraction." *Journal of Environmental Economics and Management* 1(1):53–73.

Smith, V.L., G.L. Suchanek, and A.W. Williams. 1988. "Bubbles, Crashes, and Endogenous Expectations in Experimental Spot Asset Markets." *Econometrica* 56:1119–1151.

Stiglitz, J.E. 1976. "Monopoly and the Rate of Extraction of Exhaustible Resources." *American Economic Review* 66:655–661.

Stoop, J.T.R., D.P. van Soest, and J. Vyrasktekova. 2013. "A Tale of Two Carrots: The Effectiveness of Multiple Reward Stages in a Common Pool Resource Game." In J. A. List and M. K. Price, eds. *Handbook on Experimental Economics and the Environment*. Edward Elgar, pp. 291–318.

Suits, D.B. 1984. "Dummy Variables: Mechanics V. Interpretation." *Review of Economics and Statistics* 66:177–180.

Sunder, S. 1995. "Experimental Asset Markets: A Survey." In J. H. Kagel and A. E. Roth, eds. *The Handbook of Experimental Economics*. Princeton University Press, pp. 445–500.

Tietenberg, T., and L. Lewis. 2018. *Environmental and Natural Resource Economics, 11th Edition*. New York and London: Routledge, Taylor and Francis Group.

van Veldhuizen, R., and J. Sonnemans. 2018. "Nonrenewable Resources, Strategic Behavior and the Hotelling Rule: An Experiment." *Journal of Industrial Economics* 66:481–516.

Wooldridge, J.M. 2010. *Econometric Analysis of Cross Section and Panel Data, 2nd Ed.*. Cambridge, MA: The Massachusetts Institute of Technology Press.

A Instructions for Participants

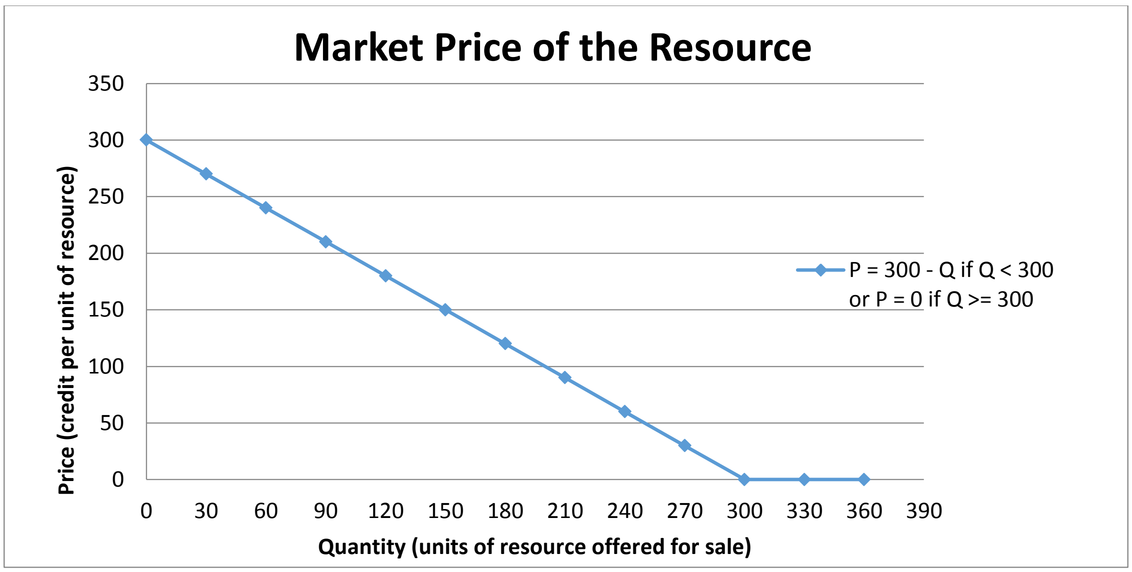
Thank you for participating in the experiment today! As a courtesy to me and the other participants, I ask you to observe a few rules: (i) focus your attention on the experiment rather than reading, texting, or other activities, (ii) do not talk while the experiment is ongoing, (iii) do not use other programs on the computers, and (iv) do not look at other people's computer screens.

The purpose of our economic experiment is to learn how people sell over time a resource that is available in fixed amounts and cannot be reproduced. To study this decision-making, you will offer your resource for sale over time in a market containing yourself and three other subjects. At the start of a market, each of you will have 100 units of the resource to sell in exchange for credits over 4 periods of trading. You will decide how much, if any, of your resource to sell in each period of trading. You will sell your resource through the software on your computer.

The software will report to you the units of the resource available to you and among the other sellers at the start of each period. In any period you may sell none, all, or some of the resource you still have. For example, if you still have 40 units of the resource, you may sell as little as zero and as much as 40 units in the current period. However, any remaining resource that you still you own at the end of the 4 periods cannot be converted into money or carried over to the next market; any unsold resource becomes worthless after the four periods of trading.

Your sales revenue each period depends on how many units you and the three other participants offer to sell in that period and the price of the resource. In particular, your sales revenue in any period will be the market price of the resource times the amount of the resource that you offer for sale. For example, if the price of the resource is 40 experimental credits per unit and you sell 25 units, you earn 1000 credits. The computer program will calculate and report to you your sales revenue in each period.

The price of the resource in each period depends on the total amount of the resource that is offered for sale. In particular, the price (P), or credit per unit, for a total quantity (Q) of the resource offered for sale is represented in the computer by this equation: P = 300 – Q, although prices never go below 0. Thus, for example, if you offer 35 units and the other three sellers offer an additional 85 units, the total amount offered would be 120 units and the price would be 180 (= 300 – 120) credits per unit of the resource. Or, if you and other owners offer a total of only 60 units, the price will be 240 (= 300 – 60) credits per unit. Or, if owners offer a total of 210 units, the price will be 90 (= 300 – 210) credits per unit. But, if owners offer at least 300 units for sale, the price will be zero. The market price of the resource will be revealed to you after all participants decide how much of their resource to sell. The figure below shows how the market price depends on the total quantity of the resource offered for sale in a trading period.



To help you decide how much to sell, the software running on your terminal has a calculator that will indicate your hypothetical sales revenue if you enter how many units you will sell and how many units you believe the other three sellers in your market will sell. These calculations are completely hypothetical and you can enter as many possibilities as you wish each period.

Your sales revenue will be deposited at the end of each period in a virtual bank that pays interest on your balance at a rate of 50 percent per period in future periods. For example, suppose you earn 1000 credits in the first period. At the end of the second period, you will be paid 500 credits in interest because 1000 x 0.50 = 500. Suppose you also sell some of your remaining resource in the second period and earn 2000 credits. Thus, at the end of the second period the balance in your virtual bank account would be 3500.00 [= (1000 + 1000 x 0.5) + 2000 = 3500.00] credits. This pattern of earning sales revenue in a period and being paid interest on the balance of your account in the next period occurs throughout the market. You will be paid interest at the end of the fourth period for the balance that you carry over from the third period. Your accumulated interest and final balance – principal plus accumulated interest – at the end of the fourth period if you deposit various amounts of credits in a prior period are shown here:

Amount of Deposit (Credits)	Prior Period When Deposit Is Made:	Accumulated Interest at End of Market (Credits):	Balance at the End of Market (Credits):
1	Period 1	2.375	3.375
	Period 2	1.25	2.25
	Period 3	0.50	1.50
2	Period 1	4.75	6.75
	Period 2	2.50	4.50
	Period 3	1.00	3.00
5	Period 1	11.875	16.875
	Period 2	6.25	11.25
	Period 3	2.50	7.50
1,000	Period 1	2,375	3,375
	Period 2	1,250	2,250
	Period 3	500	1,500
2,000	Period 1	4,750	6,750
	Period 2	2,500	4,500
	Period 3	1,000	3,000
5,000	Period 1	11,875	16,875
	Period 2	6,250	11,250
	Period 3	2,500	7,500

In addition to calculating hypothetical revenue and informing you about your actual sales revenue in each period, the computer program will also report to you the actual balance of your virtual bank account at the end of each period and four-period market. At the end of the fourth period, the final period of each market, a fee of 25,000 credits will be deducted from your virtual bank balance to determine your final net balance. However, you will only collect credits from two markets that will be randomly chosen by the computer at the end of the experiment. You will collect the credits that you earned in the two randomly chosen markets before you depart and the credits will be converted into dollars at a ratio of one dollar per 2,000 credits. For example, if the sum of your final net balances from the two randomly selected markets is 40,000 credits, you will be paid $20.00 in addition to the participation payment. Please do your best to maximize your final net balance at the end of each market.

You will participate in several markets. The first one is a practice market. You will be randomly assigned to a new market when your current market ends. So, you are unlikely to be in a market with the exact same participants as the previous market you were in. Furthermore, you will not be able to identify who your fellow sellers are in the current market. You will have 90 seconds per period in the practice market and 45 seconds per period in all subsequent markets to make your decision about how much to sell. If you do not enter a quantity to sell within the time limit, you will sell none in that period. The practice market will not count towards your earnings.

Please return these instructors to us at the end of the experiment.

B Quiz

1. T or F. The revenue that I earn from sale of my resource in a particular period equals the market price of the resource multiplied by the amount of my resource that I offered for sale.
2. T or F. The market price of the resource, the price that buyers pay for the resource, will increase as the amount of the resource offered for sale increases.
3. T or F. Any unsold amount of your resource at the end of a market will not increase your final balance.
4. T or F. Ten dollars earned from sale of my resource in the first period and deposited in my virtual bank account is worth less at the end of the fourth period than fifteen dollars earned from sale of my resource in the fourth period.
5. Fill in the blank. If, in period 1, I sell 5 units and the other 3 sellers in my market each sell 10 units, I earn __________ that period in revenue from the sale of my five units.
6. Fill in the blank. Suppose I earn 5 credits in period 1 and 20 credits in period 3 from selling my resource. If I earn no credits from any other sale, my final balance, interest included, at the end of this market is _______.

C Theory

We focus on three behaviors: subgame perfect equilibrium; collusion, in which the oligopolists cooperate and act jointly like a monopolist; and price-taking, in which the oligopolists ignore their market power and jointly act like a competitive industry. With subgame-perfect behavior, within a period, in equilibrium every firm optimally produces the same amount; for collusive and price-taking behavior, we assume that every firm produces the same amount, although asymmetric equilibria would be possible.

Proof of Prop. 1: We approach the firms' problems as static maximization problems, an approach that we show below is valid. As $q_{s,t}$ denotes s's quantity choice in period t, let $q_{-s,t} \equiv \sum_{j \neq s} q_{j,t}$ denote the total quantity choice of players other than s.

Player s's profit from a $t = 1$ perspective is

$$\sum_{t=1}^{T} \delta^{t-1} q_{s,t}(A - q_{s,t} - q_{-s,t}),$$

where $\delta = 1/(1+r)$ normalizes s's profit to first-period units, and s's choices are subject to $T + 1$ constraints:

$$\forall t \, q_{s,t} \geq 0$$
$$\sum_{t=0}^{T-1} q_{s,t} \leq x_0.$$

The Lagrangean for s's problem is

$$\mathcal{L} = \sum_{t=1}^{T} \delta^{t-1} q_{s,t}(A - q_{s,t} - q_{-s,t}) + \lambda \left(x_0 - \sum_{t=1}^{T} q_{s,t} \right).$$

The conditions are

$$\begin{aligned}
\mathcal{L}_{q_{s,1}} &= A - 2q_{s,1} - q_{-s,1} - \lambda \leq 0 & q_{s,1} &\geq 0 & q_{s,1}(A - 2q_{s,1} - q_{-s,1} - \lambda) &= 0 \\
&\cdots & &\cdots & &\cdots \\
\mathcal{L}_{q_{s,T}} &= \delta^{T-1}(A - 2q_{s,T} - q_{-s,T-1}) - \lambda \leq 0 & q_{s,T} &\geq 0 & \delta^{T-1}x_{s,T}(A - 2q_{s,1} - q_{-s,T} - \lambda) &= 0 \\
\sum_{t=1}^{T} q_{s,t} &- x_0 \leq 0 & \lambda &\geq 0 & \lambda\left(\sum_{t=1}^{T} q_{s,t} - x_0\right) &= 0
\end{aligned}$$

Resource scarcity: Assume first that $\lambda = 0$ and s's endowment is not exhausted. The condition $\delta^{t-1}q_{s,t}(A - 2q_{s,t} - q_{-s,t} - \lambda) = 0$ is only satisfied if $q_{s,t} > 0$, so the first-order conditions $\mathcal{L}_{q_{i,t}}$ all hold with equality. Therefore for every s and t, $q_{s,t} = (A - q_{-s,t})/2$ and symmetry in strategies $q_{1,t} = \cdots = q^{N,t}$ implies

$$q_{s,t} = \frac{A}{N+1}. \tag{1}$$

Strategies are symmetric because the first-order conditions for player j's problem ($j \neq s$) is the same as player i's: $A - 2q_{j,t} - q_{-j,t} = 0$, so

$$q_{s,t} + \sum_{k=1}^{N} q_{k,t} = 2q_{s,t} + q_{-s,t} = A = 2q_{j,t} + q_{-j,t} = q_{j,t} + \sum_{k=1}^{N} q_{k,t}.$$

If $TA/(N+1) \leq x$, the equation solution does not violate s's quantity constraint, and so is the unique solution to $\mathcal{L}$. That means the Nash equilibrium $q_{s,T} = \frac{A}{N+1}$ given behavior in $t = 1, \ldots, T-1$ is unique and must be a NE of these period T subgames. Induction implies $q_{s,T-1}, \ldots, q_{s,1}$ are the unique SPE Nash equilibrium strategies.

If resources are scarce: Now consider the case in which $x < \frac{TA}{N+1}$. Then $\lambda > 0$, and the first-order conditions are related by

$$(A - 2q_{s,1} - q_{-s,1}) = \delta^{t-1}(A - 2q_{s,t} - q_{-s,t}). \tag{2}$$

A similar argument to the previous one shows that strategies are symmetric, so

$$\begin{aligned} A - (N+1)q_{s,1} &= \delta(A - (N+1)q_{s,2}) \\ q_{s,t+1} &= \frac{q_{s,t}}{\delta} - \left(\frac{1-\delta}{\delta}\right)\left(\frac{A}{N+1}\right), \end{aligned} \tag{3}$$

provided $q_{s,t+1} \geq 0$. If that is the case, then $q_{s,1}$ can be found using

$$q_{s,t} = \frac{q_{s,1}}{\delta^{t-1}} + \frac{A}{N+1} - \frac{A}{\delta^{t-1}(N+1)} \tag{4}$$

so if all non-negativity constraints are slack

$$\begin{aligned} q_{s,1} + q_{s,2} + \cdots + q_{s,T} &= x \\ q_{s,1}\left(1 + \frac{1}{\delta} + \cdots + \frac{1}{\delta^{T-1}}\right) + \left(\frac{TA}{N+1}\right) - \left(1 + \frac{1}{\delta} + \cdots + \frac{1}{\delta^{T-1}}\right)\left(\frac{A}{N+1}\right) &= x \\ q_{s,1}(1 + \delta + \cdots + \delta^{T-1}) - (1 + \cdots + \delta^{T-1})\left(\frac{A}{N+1}\right) &= \delta^{T-1}\left(x_0 - \frac{TA}{N+1}\right), \end{aligned}$$

yielding

$$q_{s,1} = \frac{A}{N+1} - \left(\frac{\delta^{T-1}}{\sum_{t=0}^{T-1} \delta^t}\right)\left(\frac{TA}{N+1} - x\right). \tag{5}$$

Again each period's solution is unique, so the Nash equilibrium is subgame perfect.

Optimal stopping at $T^* \geq T$: To verify that the non-negativity constraints $q_{s,t+1} \geq 0$ for $t \leq T$ do not bind, it is sufficient to check that $q_{s,T} \geq 0$. From equations (4) and (5),

$$\begin{aligned}
q_{s,T} &= \frac{1}{\delta^{T-1}} \left(\frac{A}{N-1} - \left(\frac{\delta^{T-1}}{\sum_{t=0}^{T-1} \delta^t} \right) \left(\frac{TA}{N+1} - x \right) \right) - \frac{(1-\delta^{T-1})A}{\delta^{T-1}(N+1)} \\
&= \frac{A}{N+1} - \left(\frac{1}{\sum_{t=0}^{T-1} \delta^t} \right) \left(\frac{TA}{N+1} - x \right) \\
&= \frac{x(N+1) + A \sum_{t=0}^{T-1} \delta^t - TA}{(N+1) \sum_{t=0}^{T-1} \delta^t},
\end{aligned}$$

giving the condition

$$Q(N+1) \geq A \sum_{t=0}^{T-1} (1-\delta^t).$$

Price-taking behavior: To construct the competitive equilibrium, let $\bar{x} = N x_0$ be the total market supply of the good. We will examine aggregate behavior as N becomes large, holding $\bar{x}$ constant, so $x_0 = \bar{x}/N$.

Aggregate supply is $Q_t = N q_{s,t}$, which is

$$Q_1 = \frac{NA}{N-1} - \left(\frac{\delta^{T-1}}{\sum_{t=0}^{T-1} \delta^t} \right) \left(\frac{TNA}{N+1} - \frac{N\bar{x}}{N} \right).$$

Because in fact N firms are replicating the competive equilibrium,

$$q_{s,1} = \frac{1}{N}\left(\lim_{N\to\infty} Q_1\right) = \frac{A}{N} - \left(\frac{\delta^{T-1}}{\sum_{t=0}^{T-1}\delta^t}\right)\left(\frac{TA - Nx_0}{N}\right).$$

This provides a starting value, and subsequent periods follow

$$\begin{aligned} q_{s,t+1} &= \frac{1}{N}\left(\lim_{N\to\infty} Q_{t+1}\right) \\ &= \frac{1}{N}\left(\lim_{N\to\infty} \frac{Q_t}{\delta} - \left(\frac{1-\delta}{\delta}\right)\left(\frac{NA}{N+1}\right)\right) \\ &= \frac{q_{i,t}}{\delta} - \left(\frac{1-\delta}{\delta}\right)\left(\frac{A}{N}\right). \end{aligned}$$

Hence under price-taking behavior, the quantities sold are

$$q^*_{s,1} = \frac{A}{N} - \left(\frac{r}{\sum_{t=1}^{T} r^t}\right)\left(\frac{TA}{N} - x\right) \quad \text{and} \quad q^*_{s,t+1} = q^*_{s,t} - r\left(\frac{A}{N} - q^*_{s,t}\right).$$

Collusive behavior: The monopolist's solution is given by equations (5) and (3), letting $N = 1$ and the monopolist's quantity stock be the market quantity stock (i.e., the monopoly acts as if its stock is Nx). The oligopolists will each produce $1/N$ of that solution if they act as a joint monopolist. This implies that under collusive behavior, the quantities sold are

$$q^*_{s,1} = \frac{A}{2N} - \left(\frac{r}{\sum_{t=1}^{T} r^t}\right)\left(\frac{TA}{2N} - x\right) \quad \text{and} \quad q^*_{s,t+1} = q^*_{s,t} - r\left(\frac{A}{2N} - q^*_{s,t}\right).$$

D Predicted and Distribution of Actual Prices

D.1 Scarce Resource with High Interest Rate

Table 20: Predicted and Distribution of Actual Prices of Scarce Resource for $r = 0.50$

period	pred.	mean	std. dev.	min	25 pctl	median	75 pctl	max
$t = 1$	128.92	137.02	40.02	0	110	138.5	165	265
$t = 2$	163.38	175.24	32.28	90	152	176	200	250
$t = 3$	215.08	227.94	26.24	154	211	227	245	287
$t = 4$	292.62	260.11	29.69	165	240	265	281	300

$n = 166$. An observation is the price (experimental credits per unit of the resource) at a period of time in a four-seller group, market number, and session.

Table 21: Information for Tests of Increases in Prices of Scarce Resource for $r = 0.50$

periods	variable	cons.	bse	pr(z)	rse	pr(bt)	mdnd	pr(bi)
$t = 1, 2$	dprices12	-38.22	3.92	0.000	4.67	0.000	-35	0.000
$t = 1, 3$	dprices13	-90.92	3.90	0.000	4.48	0.000	-92.5	0.000
$t = 1, 4$	dprices14	-123.09	3.71	0.000	4.69	0.000	-127	0.000
$t = 2, 3$	dprices23	-52.70	3.06	0.000	3.60	0.000	-54.5	0.000
$t = 2, 4$	dprices24	-84.87	3.10	0.000	3.96	0.000	-86.5	0.000
$t = 3, 4$	dprices34	-32.17	2.60	0.000	2.98	0.000	-34	0.000

$n = 166$. An observation is the difference between prices (experimental credits per unit of the resource) in two periods of time for a four-seller group, market number, and session. The estimated constant (cons.) is the mean difference in two prices. The median difference is 'mdnd'.

Table 22: Predicted and Distribution of Actual Prices of Scarce Resource for $r = 0.25$

period	pred.	mean	std. dev.	min	25 pctl	median	75 pctl	max
$t = 1$	157.13	157.90	33.68	69	137	161	178	300
$t = 2$	181.41	179.80	27.71	103	161	183	198	245
$t = 3$	211.76	219.07	25.34	145	202	220	238	282
$t = 4$	249.70	243.29	31.84	130	220	245	270	300

$n = 133$. An observation is the equilibrium price (experimental credits per unit of the resource) at a period of time in a four-seller group, market number, and session.

Table 23: Statistics and Probabilities to Test for Increase in Price of Scarce Resource for $r = 0.25$

periods	variable	cons.	bse	pr(z)	rse	pr(bt)	mdnd	pr(bin)
$t = 1, 2$	dprices12	-21.89	3.80	0.000	4.28	0.000	-19	0.000
$t = 1, 3$	dprices13	-61.17	4.34	0.000	4.54	0.000	-60	0.000
$t = 1, 4$	dprices14	-85.38	4.55	0.000	4.68	0.000	-85	0.000
$t = 2, 3$	dprices23	-39.27	4.08	0.000	3.28	0.000	-40	0.000
$t = 2, 4$	dprices24	-63.49	4.37	0.000	4.55	0.000	-65	0.000
$t = 3, 4$	dprices34	-24.22	3.83	0.000	3.83	0.000	-28	0.000

$n = 133$. An observation is the difference between prices (experimental credits per unit of the resource) in two periods of time for a four-seller group, market number, and session. The estimated constant (cons.) is the mean difference in prices in two periods of time. The median difference is 'mdnd'.

D.2 Scarce Resource with Low Interest Rate

D.3 Abundant Resource with High Interest Rate

Table 24: Predicted and Actual Prices of Abundant Resource for $r = 0.50$

period	pred.	mean	std. dev.	min	25 pctl	median	75 pctl	max
$t = 1$	60	66.81	40.61	0	37	68.5	93	174
$t = 2$	60	64.10	35.97	0	40	64	90	140
$t = 3$	60	64.03	39.78	0	36	65	85	163
$t = 4$	60	77.94	44.09	0	46	78	100	195

$n = 70$. An observation is the equilibrium number of experimental credits per unit of the resource in a period of time for a four-seller group, market number, and session.