COORDINATION AND UNCERTAINTY

BY DEAN JENS

A dissertation submitted to the

School of Graduate Studies

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Economics

Written under the direction of

Barry Sopher

and approved by

New Brunswick, New Jersey

October, 2018

ABSTRACT OF THE DISSERTATION

Coordination and Uncertainty

by Dean Jens

Dissertation Director: Barry Sopher

Even when economic agents have common interests, it can be difficult for them to coordinate on a commonly-accepted outcome. I use both human subject experimentation and theory to shed light on how coordination takes place and where and why it might be prone to fail. In a human subject experiment, I find that the coordination problem leads to inertia in groups — once coordination has taken place, agents avoid trying to change the outcome — and that, with exogenous uncertainty and differing risk preferences, the less risk-averse agent tends to defer to the more risk-averse agent, though in a state of potential loss, conflicting notions of risk-aversion can make coordination even more difficult. In the second chapter, I construct a theoretical model of storable consumption goods, in which markets for goods with high income-elasticity are shown to be prone to "runs" in which consumers, in the face of otherwise adequate supply, store up for a shortage, thereby creating the very shortage they feared; it would be in everybody's interest to avoid this outcome, and yet avoiding it requires a great deal of mutual trust and coordination. In the final chapter, an information-theoretic elaboration of this idea is presented, in which agents form robust but diffuse beliefs about what other agents believe other agents believe, and I note that cutting off the chain of reasoning at the

second level favors “risk-dominant” equilibria over “risk-dominated” equilibria, but can also favor non-equilibrium outcomes in situations in which equilibrium strategies are only best-responses to fairly precise actions by other agents.

Table of Contents

Chapter 1

Coordination with Uncertainty

1.1 Background

1.1.1 Coordination and strategic uncertainty

While theoretical game theory has made great strides in analyzing normative solution concepts in games,[1] these solution concepts encounter occasional problems in describing the strategic interaction of people. One class of problems stems from the failure of the assumption of common knowledge of rationality, but another class falls under the rubric of "equilibrium selection"; if a game has two theoretically sound equilibria, the agents need a way to coordinate on one equilibrium.

If two agents are driving cars toward each other, they need to decide whether to drive on the right hand side of the street or the left hand side (figure 1.1). Any pure action by either player is part of a strict equilbrium, and thus strategic equilibrium theory provides no guidance to either player as to which action to take. Indeed, the two pure-strategy equilibria survive attempts at equilibrium selection (e.g. Harsanyi and Selten (1988)); while either will be self-fulfilling if both agents expect it to obtain, coordination on an equilibrium requires either luck or some exogenous source of beliefs.

Figure 1.2 represents a game with a Pareto-dominant equilibrium in which players encounter a different problem with coordination. The narrative is that two hunters are individually capable of hunting hares, but may bring down a stag if they coordinate on that choice; they would each prefer their share of a stag to a hare, but the consequences

[1]This is reviewed in Hillas and Kohlberg (2002); Van Damme (2002).

of a coordination failure are more severe for an agent who chooses the bigger game. Harsanyi and Selten (1988) and Harsanyi (1995) come to different conclusions as to which equilibrium is (in some sense) the "right" one.

In a final illustrative example, figure 1.3 shows the "battle of the sexes", in which each player has the same action set and has a strong preference to coordinate on the same action as the other player, but also has a (weaker) preference for playing one action or the other, with the two players differing in this last preference. The two-player game in this experiment (figure 1.5 on page 8) seems superficially to resemble figure 1.1, but if the players have different risk preferences the strategic situation is closer to that of figure 1.3. Of course, agents with different risk preference may be expected to respond differently to the strategic risk inherent in not knowing what one's partner will choose. There is thus a coupling between an agent's preference between the two (pure) equilibria and its response to the possibility that its partner would choose differently.

In a game with one player, this strategic uncertainty is not an issue. It is still the case that subjects in the lab, as well as economic agents in the field, exhibit violations of rationality that cannot be explained by unusual transitive preferences. There have been attempts to expand the utility-maximization framework (e.g. Gul and Pesendorfer (2001)), and attempts to more deeply model possible explanations for violations of utility maximization, such as bounded rationality. One popular model for boundedly rational agents is that of a finite-state automaton; the finite number of states limits the ability of the agent to reliably choose the "best" option available according to some underlying preferences. (See, for example, Salant (2011).)

An experimental setting allows us to observe human subjects making decisions in a comparatively informationally sparse environment. In this study, we present the subjects with a simple stochastic game. The explicit information entailed in this environment is well attuned to the finite automaton construct, though there is a great deal

		player 2	
		left	right
player 1	left	1,1	$-1,-1$
	right	$-1,-1$	1,1

Figure 1.1: A simple game with multiple equilibria.

		player 2	
		stag	hare
player 1	stag	2,2	$-2,1$
	hare	$1,-2$	1,1

Figure 1.2: Jean-Jacques Rousseau's "stag hunt"

of implicit information and room for heuristic decision making remaining.

1.1.2 Stochastic Games

In 1953, Lloyd Shapley introduced into the literature the concept of a "stochastic game", a game with multiple stages and a "state space" such that the actions available to players in a stage may depend on the state in that stage (Shapley (1953)). Payoffs to players will be (possibly discounted) sums of stage game payoffs, which in general depend on the state in that stage as well as the action profile played in that stage. The state in a period is random, with a distribution determined by the state in the previous stage, as well as the action profile in that stage. Such games can model economic situations with repeated interactions in which actions affect state variables; for example, fishermen drawing fish from a common fishery might affect the quality of the fishery

		player 2	
		ballet	ball game
player 1	ballet	3,2	1,1
	ball game	0,0	2,3

Figure 1.3: "Battle of the sexes." There are two Pareto-optimal essential equilibria, but the preferences over the equilibria are in conflict, making coordination more difficult.

going forward by their fishing practices, subject to some variability that is outside their ability to precisely forecast. Shapley proved the existence of an equilibrium in a context that is somewhat narrower than usual modern interest, but existence results have been generalized substantially in the last six decades (Mertens (2002); Vieille (2002)).

A pure strategy in such a setting could be quite complicated, but of particular interest are stationary Markov strategies, in which a player's action is a function only of the state. For the duration of this paper, we shall only be concerned with stationary Markov strategies.

1.1.3 Individual Decision Making

Traditional game theory investigates the strategic consequences of rational utility maximization; it has been well established, however, that human agents do not always behave as rational expected utility maximizers. This is especially true in situations that are novel and those that involve probabilities.

For binary lotteries over gains, human agents generally exhibit risk aversion to a degree that seems severe if risk aversion is measured relative to the agent's total wealth. Indeed, there is some evidence of risk-seeking behavior in the face of "losses", with agents preferring risks of larger losses to certainties of the same expected loss. It is also the case that, given series of lotteries, human agents routinely fail to aggregate them according to laws of probability; for example, a lottery presenting a subject with a 1/3 probability of winning the chance to then have a 50/50 chance of winning a $5 gain tends to be treated differently from a single, atomic 1/6 probability of a $5 gain. Even the simplest of stochastic games will necessarily incorporate many features that may distinguish the actual behavior of subjects from expected valuation maximization.

1.1.4 Coordination

There are many experimental results in which subjects are presented with games in which there are many Nash equilibria that are either Pareto ranked (e.g. “minimum effort” games; cf. Van Huyck, Battalio, and Beil (1990)) or such that agents have different preferences over them (e.g. “battle of the sexes”). Outcomes in either case can depend on a variety of details. In genuinely one-shot games with multiple equilibria, there is frequently no particular reason to expect that any equilibrium will be attained if there is no common basis for beliefs as to which equilibrium should obtain.

If an agent has clear, concrete expectations for the actions of other agents, the agent may view the choice of action as a decision problem; if there is one other agent with possible actions C and D, and the agent assigns a 70% chance to action C, then choosing action A is tantamount to choosing a 70% chance of the results of action profile (A,C) and a 30% chance of the outcome of (A,D). In practice, a crisp probability of this sort is likely not to be at hand; a result generally attributed to Ellsberg (1961) shows that people have preferences for “certain uncertainties” over less quantifiable uncertainties, as are exhibited in situations of strategic uncertainty.

1.1.5 Other Related Literature

Cooper, DeJong, Forsythe, and Ross (1990) investigate equilibrium selection in play of small games with rematching after each round; in each case play does converge to a Nash equilibrium, but which equilibrium obtains can be affected by payoffs that should, from a theoretical standpoint, be irrelevant to strategic stability. Pareto dominated equilibria often result.

Mookherjee and Sopher (1994, 1997) examine learning behavior in repeated constant-sum games; these are similar to a zero-difference game in that the folk theorem has very little “bite”. They find that experienced payoffs drive behavior to a greater extent than deep strategizing. Camerer and Ho (1999) examine a more general learning model state 1

A	B
0	a

state 2

A	B
b	c

A	B
p_1	p_3

A	B
p_4	p_2

Figure 1.4: Transition probabilities given are for transition to state 2.

and find support for some influence of higher-order strategic behavior, but with an important contribution from experienced feedback.

1.2 A Stochastic Game Experiment

1.2.1 One player game

Consider the following one-player stochastic "game" (decision problem): there are two states, state 1 and state 2, and two actions in each state, which will be denoted A and B in both states, as in figure 1.4. State 2 has generally higher payoffs than state 1 ($b > 0$, $c > a$), and action A in either state results in a higher probability that the next stage will be in state 2 rather than state 1; however, action B results in a higher payoff ($a > 0$, $c > b$) in the current stage. Of particular interest is the case with $p_3 = 0$ and $p_4 = 1$; in each state, the agent is asked to choose between a payoff that will leave the state unchanged versus a different payoff with a lottery over continuation values.

Valuations for different states and strategies are derived at the top of page 7 for stationary Markov pure strategies with discount factor δ. In particular, for $p_3 = 0$, $p_4 = 1$, and $\delta \to 1$, AB gives an expected average payoff of $cp_1(1 + (p_1 - p_2))^{-1}$, while the other three strategies end up stuck in one state repeatedly getting the same payoff. For any value of δ, p_1, $p_2 \in (0, 1)$, and a, b, and c, there will be at least one strategy that is optimal in both state 1 and state 2.

As $\delta \to 1$, the law of large numbers kicks in, and discounted payoffs converge in

The pure stationary Markov strategies for the decision problem have expected values that can be expressed recursively

strategy	state 1	state 2
AA	$0+\delta p_1 c_2+\delta(1-p_1)c_1$	$b+\delta c_2+\delta(1-p_4)(c_1-c_2)$
AB	$0+\delta p_1 c_2+\delta(1-p_1)c_1$	$c+\delta p_2 c_2+\delta(1-p_2)c_1$
BA	$a+\delta c_1+p_3\delta(c_2-c_1)$	$b+\delta c_2+\delta(1-p_4)(c_1-c_2)$
BB	$a+\delta c_1+p_3\delta(c_2-c_1)$	$c+\delta p_2 c_2+\delta(1-p_2)c_1$

For interior values of δ and p_i, we can solve and multiply by $1-\delta$:

	state 1	state 2
AA	$\delta p_1(1+\delta p_1-\delta p_4)^{-1}b$	$(1-\delta(1-p_4)(1+\delta p_1-\delta p_4)^{-1})b$
AB	$c\delta p_1(1+\delta(p_1-p_2))^{-1}$	$c(1-\delta(1-p_1))(1+\delta(p_1-p_2))^{-1}$
BA	$a+p_3\delta(1-\delta p_4+p_3\delta)^{-1}(b-a)$	$b-\delta(1-p_4)(1-\delta p_4+p_3\delta)^{-1}(b-a)$
BB	$a+p_3\delta(1-\delta p_2+\delta p_3)^{-1}(c-a)$	$c-\delta(1-p_2)(c-a)(1-\delta p_2+\delta p_3)^{-1}$

For $p_3=0$ and $p_4=1$, these reduce to

strategy	state 1	state 2
AA	$(1-\delta(1-p_1))^{-1}\delta p_1 b$	b
AB	$c\delta p_1(1+\delta(p_1-p_2))^{-1}$	$c(1-\delta(1-p_1))(1+\delta(p_1-p_2))^{-1}$
BA	a	b
BB	a	$c+\delta(a-c)(1-p_2)(1-\delta p_2)^{-1}$

probability. Classical treatments of risk-aversion, then, leave asymptotically little room for non-trivial risk-aversion as long as agents compound lotteries and only care about the distribution of total final payoffs. A long strain of results, however, shows that subjects presented with sequences of risky choices may not aggregate them according to the laws of probability (Allais (1953)). Yaari (1987) constructs a model of choice in which he replaces the assumption that compound lotteries are handled in this fashion with a different assumption, leading him to a model that differs from expected utility; applying his model here amounts to allowing subjects to act as though p_1 and p_2 had different values than they do here. If $\min\{c,b\} > \max\{a,0\}$, the value of an agent playing any strategy will be weakly higher in state 2 than in state 1, and strictly so for δ, p_1, and $p_2 \in (0,1)$. Under these circumstances, it seems natural to frame a move from state 1 to state 2 as a gain and a move the other direction as a loss; in this case a risk averse agent will use a lower effective value for both p_1 and p_2; a risk-seeking agent will maximize an expression raising those numbers. If agents are risk averse when faced state 1

	A	B
A	0,0	0,0
B	0,0	a,a

state 2

	A	B
A	b,b	b,b
B	b,b	c,c

	A	B
A	p_1	p_3
B	p_3	p_3

	A	B
A	p_4	p_2
B	p_2	p_2

Figure 1.5: Transition probabilities given are for transition to state 2.

with potential gains but risk-seeking when faced with losses, as suggested by Kahneman and Tversky, they might maximize an expression that lowers p_1 and raises p_2.

A model closer to expected utility maximization would allow for attempted "smoothing" between stage-game payoffs, even where actual "consumption" may not be tied (temporally) to those payoffs. Agents would seek to optimize the expected net present value of a function of the payoffs. This, too, is a simple change from the derived expressions; risk-aversion in this framework would amount, for example, to using lower effective values of b and c.

While a setup with $p_3 = 0$ and $p_4 = 1$ is attractive in that the choice at each decision stage is one between a pure value and a lottery, it is unattractive in that some of the strategies do not induce ergodic outcomes, in which case it becomes hard to distinguish between full strategies. Accordingly, in this study we use interior values for these probabilities, allowing the agents to compare different lotteries instead of a lottery with a certainty.

1.2.2 Two player game

Now expand this to the two-person coordination game in figure 1.5 with $p_1 > p_3$, $p_4 > p_2$, $c > b > a > 0$; in each state, if the two players coordinate, always playing the same action, then they face the same decision problem as before, choosing between a safe immediate payoff or a chance at higher future payoffs in state 1, and between a safe payoff now or a higher payoff with a risk of lower payoffs in the future.[2] However, if they make different choices of action (one choosing A and the other choosing B), they get the worst of both: in state 1, they get no payoff now nor any chance of moving up to the higher continuation value, while in state 2, they get the lower available payoff now and the risk of moving to the lower continuation value.

As with many coordination games, each player needs to make an assessment of the other player's likely action, based on beliefs of that player's own preferences as well as that player's strategizing and higher order beliefs. In this game the decisions depend explicitly on preferences over exogenous risk factors, instead of just on uncertainty created by the challenge of coordination.

By construction of the game, if one player could publicly commit to an action, that same action would dominate[3] the other action for the other player. Two players with different risk preferences who fail to coordinate, then, are failing to coordinate for strategic reasons — either because they failed to anticipate each other's actions, or because they hope to affect the other player's choice of action in future rounds.

1.3 Procedure

The experiment was programmed and conducted with the software z-Tree (which is described in Fischbacher (2007)) and conducted in the Gregory Wachtler Experimental Economics Laboratory (room 107 in Scott Hall at Rutgers University). After being given instructions, the 12 subjects in the first session were presented with the one-player game, which they played for four different values of the parameters, which were chosen randomly; the parameters are given in figure 1.6. Each of the four rounds used

[2] There is some framing in the phrasing here, with continuation values measured against what would be earned from remaining in a given state. A perfectly rational player might no more view a transition from state 2 to state 1 as a loss than staying in state 1 when there was a positive chance of moving to state 2, especially where $p_1 + p_2 = 1$. I expect, and even intend, for subjects to frame the decision as I have in the text, and for any differences between lotteries over gains and lotteries over losses to generate a difference in behavior between state 1 and state 2.

[3] In a first order stochastic sense, not a game theoretic sense.

round	a	b	c	p1	p2	p3	p4	B in st. 1	B in st. 2	stages
1	4	9	9	0.76	0.77	0.29	0.29	0.5909	0.1250	14
2	4	8	13	0.72	0.76	0.28	0.28	0.3803	0.2623	11
3	4	8	13	0.77	0.76	0.24	0.29	0.2927	0.2903	6
4	7	8	13	0.78	0.79	0.28	0.29	0.8387	0.5517	5
5	6	8	12	0.85	0.70	0.26	0.24	0.7411	0.1071	21
6	7	7	12	0.80	0.69	0.28	0.19	0.6923	0.6481	11
7	9	8	12	0.83	0.71	0.29	0.17	0.9712	0.5385	47
8	8	11	11	0.84	0.74	0.32	0.21	0.7284	0.0108	29
9	7	10	12	0.79	0.66	0.33	0.18	0.8529	0.1333	37
10	9	9	12	0.77	0.75	0.28	0.25	0.9867	0.6800	23
11	6	11	13	0.75	0.72	0.35	0.20	0.5682	0.3846	7
12	7	9	12	0.77	0.67	0.29	0.18	0.6303	0.3517	31
13	5	7	11	0.85	0.68	0.28	0.16	0.3957	0.2552	26
14	7	8	14	0.77	0.67	0.34	0.22	0.6506	0.5410	29
15	7	8	10	0.73	0.69	0.30	0.31	0.7852	0.1518	24
16	7	10	10	0.75	0.68	0.22	0.31	0.5000	0.0660	10
17	7	7	13	0.73	0.67	0.25	0.27	0.8212	0.6287	28

Figure 1.6: The parameters for each round, along with the fraction of choices in each state in which the subject chose B in that round.

previous stage	state 1		state 2	
action	count	B%	count	B%
A	227	34.80	444	7.66
B	604	87.58	181	78.45

Figure 1.7: Descriptive statistics for one-player games

a random stopping rule that imposed a 5% chance of ending the round at each stage. After those four rounds, subjects were then put in pairs, in which they played the two-player version, allowing us to study the ability of subjects to coordinate. In each round the players were randomly matched by the software with a partner, with different values of parameters; the same 5% stopping rule was used as in the one-player game.

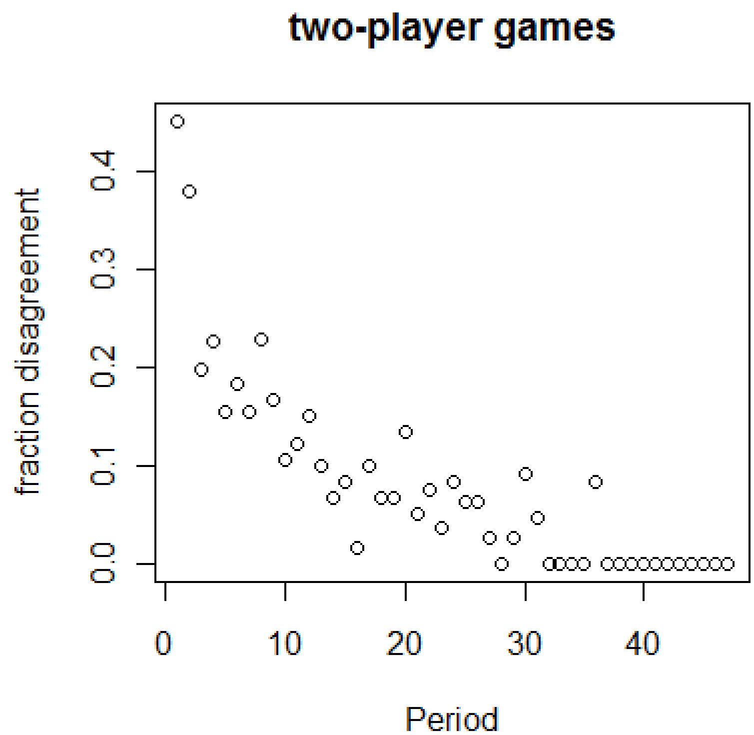
In the second session, 20 subjects were first exposed to the two-player game, which they played 4 times, and then to the one-player game, which they played 3 times.

previous stage		state 1		state 2	
own	partner	count	B%	count	B%
A	A	344	9.59	1018	1.57
A	B	134	34.33	87	21.84
B	A	134	51.49	87	73.56
B	B	1326	99.40	392	97.45

Figure 1.8: Descriptive statistics for two-player games

round	A/A	A/B	B/A	B/B	B in st. 1	B in st. 2	stages
1	6.203	4.322	**6.292**	5.303	0.5909	0.1250	14
2	5.472	6.107	5.828	**6.306**	0.3803	0.2623	11
3	5.536	**6.325**	5.648	6.052	0.2927	0.2903	6
4	5.558	6.249	7.450	**8.496**	0.8387	0.5517	5
5	5.951	6.365	6.941	**7.541**	0.7411	0.1071	21
6	5.371	6.223	7.000	**8.369**	0.6923	0.6481	11
7	6.308	6.254	8.434	**9.826**	0.9712	0.5385	47
8	8.380	5.660	**9.648**	8.863	0.7284	0.0108	29
9	7.725	6.309	**8.760**	8.583	0.8529	0.1333	37
10	6.461	5.876	9.000	**9.776**	0.9867	0.6800	23
11	8.228	6.403	**8.904**	8.182	0.5682	0.3846	7
12	6.912	6.190	8.110	**8.432**	0.6303	0.3517	31
13	5.599	5.908	6.137	**6.659**	0.3957	0.2552	26
14	5.908	7.222	7.555	**9.240**	0.6506	0.5410	29
15	5.345	4.957	7.453	**7.863**	0.7852	0.1518	24
16	6.741	5.059	**8.133**	7.693	0.5000	0.0660	10
17	4.854	6.533	7.000	**8.542**	0.8212	0.6287	28

Figure 1.9: Expected payoff from different pure stationary Markov strategies for the realized parameters of each round.



	state 1, one-player Estimate	one-player Std. Error	state 2, one-player Estimate	one-player Std. Error
(Intercept)	-1.5729	0.5873	-1.0321	0.8446
A/A−7	-0.5682	3.1863	2.2610	3.5780
A/B−7	-0.8367	2.2355	-0.9049	2.5752
B/A−7	-0.5401	5.2249	-4.6501	5.9522
B/B−7	0.9531	3.9823	3.8065	4.5279
	state 1, two-player Estimate	two-player Std. Error	state 2, two-player Estimate	two-player Std. Error
(Intercept)	-1.8443	0.2504	-2.3733	0.2358
A/A−7	-2.4520	0.7642	-0.4217	0.6047
A/B−7	-0.3261	0.3533	-0.3414	0.3089
B/A−7	2.9508	1.0319	-0.7747	0.8202
B/B−7	-0.6482	0.6264	1.1683	0.5062

Figure 1.10: Logit coefficients, regressing on theoretical values of pure strategies, subtracting 7 (approximately the mean value) to reduce the standard error on the intercept term. Negative intercepts indicate a preference for A when expected payoffs for different strategies are comparable.

1.4 Results

Some basic statistics are presented in figures 1.7 and 1.8; player responses in the tables exclude the first stage at which a state is encountered in a round; for example, if the first three stages of a round are in states 1, 1, and 2, respectively, stages 1 and 3 are excluded. This is so that the results can be broken down into categories by which action was played in the previous occurence of that state. Subjects in one-player games rarely switch from B to A in state 1 or from A to B in state 2; switches in the other direction are somewhat more common.

In the two-player game, switches out of coordination are even less likely; of the four possible switches, the only one that happens in more than 3% of the opportunities is the attempt to switch to B when the players previously coordinated on A in state 1; note that A to B in state 1 is also the most common switch in the one-player game.

There are 134 stages in which the players are in state 1, will be in state 1 at least one more time in that round, and play different actions; there are 87 similar occurrences in state 2. In state 1 the player who played B switches almost half of the time at the next opportunity, but the player who played A attempts to switch more than 1/3 of the time as well. In state 2, the players both tend to "dig in their heels" a bit more, switching about a quarter of the time. This can result in a sustained period of non-coordination; see figure 1.13 for an illustration.

Figure 1.11 shows one of the only three situations in which partners managed to switch directly from playing the same action in one stage to both playing the other action in the next occurrence of the same state. The subjects switch from playing B in stage 6 to, when they finally get back into state 2 in stage 13, both playing A, on which they coordinate for the remainder of that round. This game is an outlier in most of the logit models, and motivates the inclusion of a variable that captures the number of stages subjects spend in state 1 between recurrences of state 2; the significance of the

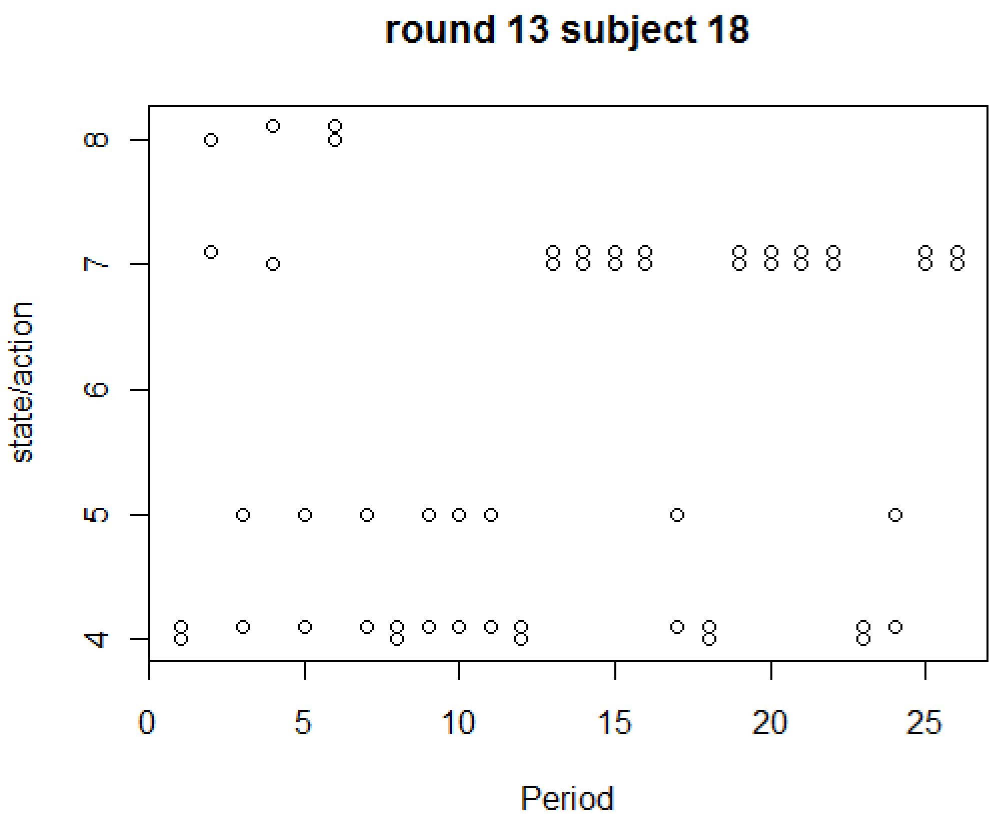


Figure 1.11: Round 13, subjects 18 and 25.

coefficient on this variable demonstrates that agents are at least in part responding to experienced probabilities and not just the numbers presented.

1.4.1 Stationary Markov strategies

It was hypothesized that subjects in one-player games would play stationary Markov strategies. A rational agent basing its play solely on the data presented and not on experience in the game would have no reason to change its mind during play. It was also hypothesized that, in two-player games, both subjects would play the same stationary Markov strategy after a "coordinating period".

Of 1456 occasions in which subjects played a one-player game having been in the same state previously, the subject played the same action 1229 times, viz. 84.4% of the time. Even beyond 20 periods, however, this only goes up to 88.9%. On the other hand, of 3522 stages in which subjects played a two-player game having been in the same state previously, the subject played the same action as the previous occasion 3302 times, viz. 93.8% of the time; beyond 20 periods, this goes up to 1178/1204, or 97.8%. (In the first 8 periods, the ratio is 87.1%.) This suggests that in fact the two-player games are converging to a pure stationary Markov equilibrium more quickly than the one-player games are.

In another, more formal test of these hypotheses, likelihood ratio tests were performed. A logit is fit against an indicator variable indicating either the previous action (for the one-player game) or the previous pair of actions (for the two-player game; this results in 3 dummy variables in addition to the intercept term), along with

- indicator variables for the subjects
- the 7 parameters describing the particular game
- both of these sets of explanatory variables.

These fits are done separately for each state, and separately for games with one player

		all 38	no dummies	no 7 params	just prev A
one player	log-likelihood	-6.821	-43.134	-16.599	-49.153
Period> 20	p-values	2.08×10^{-5}	0.0993	3.23×10^{-4}	NA
one player					
Period> 10	p-values	3.55×10^{-15}	.127	1.69×10^{-10}	NA
two player	log-likelihood	-21.714	-42.610	-24.234	-45.262
Period> 20	p-values	0.1479	0.6228	0.0889	NA
1p, > 20	p-values	2.38×10^{-3}	2.49×10^{-3}	.1214	(state 2)
1p, > 10	p-values	1.65×10^{-7}	3.63×10^{-7}	.0508	(state 2)
2p, > 20	p-values	.256	.242	.413	(state 2)

Figure 1.12: Likelihood ratio tests as described in the text. For one-player games, even after 20 periods, the previous action by itself is not the best predictor of the subsequent action, but for two-player games the data are more consistent with the assertion that convergence to a stationary markov strategy has taken place. This is true both in state 1 and in state 2.

and games with two players excluding the first 20 stages of each round; because this excludes so many one-player games altogether, it is also done for one-player games excluding only the first 10 stages. If the subjects are playing stationary Markov strategies at this point, actions should be entirely predictable from the previous action in that state; other explanatory variables should not improve the prediction. Results are in figure 1.12. While subjects do not deviate at a significant level from two-player equilibrium, perhaps because coordination is valuable and difficult to achieve, they do, even after 20 periods, switch their actions, even in predictable ways.

1.4.2 Hypotheses on differing risk preferences

It was hypothesized that subjects with different risk preferences would take longer to coordinate than subjects whose risk preferences were similar. It was of general interest to see how coordination took place, and whether there was an asymmetry in the resolution of disagreements, i.e. whether subjects who preferred B (the immediate option) were more or less likely to defer to subjects who preferred A than vice versa.

In state 2, there were 16 occasions when a disagreement was followed by A/A,

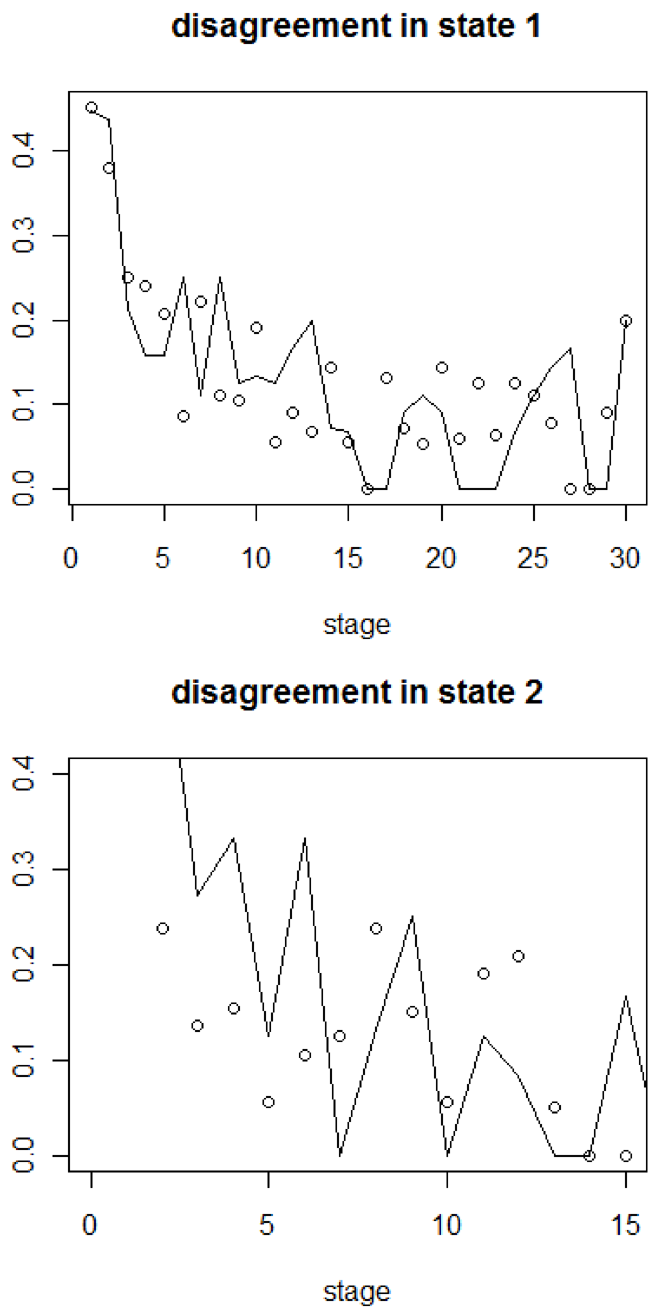


Figure 1.13: The circles represent the fraction of subject pairs that disagree in a given stage in a two-person game in which two subjects with similar behavior in the one-person game are matched. The lines represent games in which the partners exhibit different one-person game behavior.

and 6 when it was followed by B/B. In state 1, the respective figures were 31 and 8.[4] For a symmetric binomial distribution, these have two-tailed p-values of 0.0525 and 2.94×10^{-4}, respectively.

Figure 1.13 shows no discernible tendency to take longer to coordinate in state 1 on the basis of mismatched preferences, but state 2 does appear to exhibit such a tendency. In periods 2–6, the mismatched pairs disagree on 21 out of 61 occasions, while the similar pairs disagree 13 times out of 93. A random sample of 61 items from a population of 154 will include at least 21 out of 34 specially designated items with a probability of 8.30×10^{-13}.

Figure 1.10 demonstrates a generic tendency by subjects to prefer A in situations in which an expected payoff maximizing agent would be indifferent.

1.5 Discussion and Conclusion

We have presented subjects with a fairly simple environment in which we can study the ability of subjects to coordinate on whether to take a large payoff now or take a chance of receiving a higher payoff in the future. Subjects had a tendency to prefer to enhance their expected continuation value, especially in the two-player game, in which differences tended to be resolved in that direction; if one agent wanted to take the higher payoff immediately, especially in state 2, but the other agent wanted to "play it safe", increasing the likelihood of staying in state 2, the former agent would typically defer to the latter's choice.

In the context of single-agent decision theory, a reluctance to risk moving from a profitable state to an unprofitable state may be related to the concept of "loss aversion" (Kahneman and Tversky (1979)). It is interesting to note that an apparent loss aversion seems to be more pronounced in this case when agents are attempting to coordinate than when they do not face that challenge; the subject who is less loss-averse tends to

[4] See page 20 for an exhaustive breakdown on which action pairs followed each other in what quantity.

defer to the subject who is more loss-averse.

The challenge of coordination also makes it more costly for a subject to experiment; the subject pairs thus lock more quickly into a pure stationary Markov strategy, while individual players exhibit more of a learning behavior, responding to experience by maintaining a willingness to change strategies even late in a round. Economic agents in the field may experience similar kinds of lock-in, in which agents find it difficult to make even Pareto-improving changes if the cost of coordinating the change seems likely to exceed the benefit from the switch.

While it would be interesting to attempt to infer effective one-player game parameters that would lead to the behavior in the two-player games, and perhaps follow that back to effective beliefs about what one's partner is likely to play, this would seem to require more data than have been gathered in this experiment.

action	previous	count
state 1, one-player		
A	-	57
B	-	51
A	A	148
B	A	79
A	B	75
B	B	529
state 2, one-player		
A	-	65
B	-	43
A	A	410
B	A	34
A	B	39
B	B	142

action	partner	prev A	prevPA	count
state 1, two-player				
A	A	-	-	26
A	B	-	-	32
B	A	-	-	32
B	B	-	-	52
A	A	A	A	280
A	B	A	A	31
B	A	A	A	31
B	B	A	A	2
A	A	A	B	41
A	B	A	B	47
B	A	A	B	24
B	B	A	B	22
A	A	B	A	41
A	B	B	A	24
B	A	B	A	47
B	B	B	A	22
A	B	B	B	8
B	A	B	B	8
B	B	B	B	1310
state 2, two-player				
A	A	-	-	54
A	B	-	-	22
B	A	-	-	22
B	B	-	-	44
A	A	A	A	986
A	B	A	A	16
B	A	A	A	16
A	A	A	B	20
A	B	A	B	48
B	A	A	B	3
B	B	A	B	16
A	A	B	A	20
A	B	B	A	3
B	A	B	A	48
B	B	B	A	16
A	A	B	B	4
A	B	B	B	6
B	A	B	B	6
B	B	B	B	376

Figure 1.14: Extended descriptive statistics.

Appendix: Instructions for Subjects

Introduction

You are about to participate in an experiment in the economics of decision making. Various research foundations have provided the funding for this research. The research is designed to study how people make decisions when facing uncertainty, both individually and in small groups. At the end of the experiment you will be paid for your participation, as outlined in the following instructions.

Background

Many decisions we make, especially financial decisions, require long-term considerations, but the future is inherently uncertain. While you may not have complete control over the future, however, you can often have some influence; for example, your car might break down even if you have maintenance done on it, or it might not break down even if you don't, but it's more likely to have a breakdown if you are behind on your regular maintenance.

In this experiment, you will be making a series of simple choices about a machine; you can make light use of it, in which case it produce lower profits for you in the short-run but is more likely to be more productive in the future, or you can make heavy use of it, making it produce more now. Your final payoff will depend on its total production.

Instructions

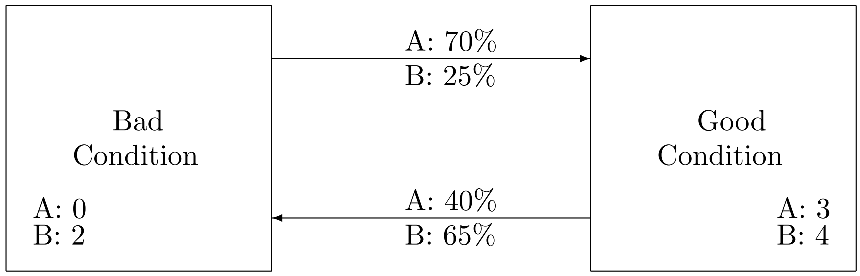
This experiment will be divided into several rounds; in each round you will make a series of binary choices (A or B). For each choice you will be awarded a certain number of an in-experiment currency unit called shillings, but your decision will also affect the number of shillings that each choice will be worth in later stages of the same round.

At the beginning of each round you will learn something about your machine for that round. It will have two possible conditions: it may be in good condition or bad condition. The round will be divided into stages; the condition of the machine may change from one stage to the next. If the machine is in good condition, it might wear down to bad condition for the next stage, but, because some regular maintenance is being done on the machine, it is also possible for a machine in bad condition to move back into good condition. In each stage you will be given two options as to how to use the machine. Choice A will generally produce less in the current stage than choice B, but it is also generally the case that a machine is more likely to go from bad condition to good condition with choice A, and is more likely to go from good condition to bad condition with choice B.

Figure 1.15 gives an example. In this example if the machine is in bad condition, choosing A will produce 0 shillings in this stage, but creates a 70% chance that the machine will be in good condition in the next stage; choosing B will produce 2 shillings in this stage, but creates only a 25% chance that the machine will be in good condition in the next stage. If the machine is in good condition, choosing A will produce 3 shillings in this stage, and creates a 40% chance that the machine will fall into bad condition for the next stage; choosing B will produce 4 shillings in this stage, but creates a 65% chance that the machine will fall into bad condition for the next stage.

In each round the immediate payouts (i.e. productivity) will be given in whole numbers; there is a payout level for choice A when the machine is in bad condition, choice B when the machine is in bad condition, choice A when the machine is in good condition, and choice B when the machine is in good condition. Similarly, for each choice in each condition, you are provided with the probability that that choice would cause the machine to change conditions in the next stage, i.e. change from bad condition to good condition or from good condition to bad condition.

Throughout the round, all eight of these numbers will always be displayed near the top left of the screen. During the experiment, when the machine is in bad condition,



A	B		A	B
0	2	payouts	3	4
0.7	0.25	switch prob	0.4	0.65

Figure 1.15: An example of the table you could see at the top of the screen to tell you about your machine, along with a diagram demonstrating what the numbers mean.

the four numbers related your immediate options will be in red, while the numbers related to the machine in good condition will be in gray. When the machine is in good condition, the four numbers related your current choice will be in green, while the numbers related to the machine in bad condition will be in gray.

Different machines are different. Some produce a lot more when they are in good condition than when they are in bad condition, while for others the difference is small. Some are more sensitive than others to how hard they are run; some might be a lot more likely to deteriorate under high levels of production than others. It is important for you to decide how to trade off these factors; it may make sense to treat some machines differently from others.

The first screen capture on page 26 shows an example; in this example, the machine is in bad condition, so the payoffs (and some of the labels) are in red; those payoffs are 0 for A and 2 for B, as noted both in the main screen and above it; if the player chooses A there is a 75% chance that the machine will be in good condition for the next stage, which would offer 2 for A and 4 for B (as shown in gray above the main screen), but if the player chooses B there is only a 25% chance of switching.

At some point, the machine will abruptly stop producing profits. It is, unfortunately, prone to being hit by lightning, which happens with a 5% probability in each stage. This probability is independent of whether the machine is in good condition or bad condition. When that happens, the round ends; your total payoff from the round will be the total of the payoffs received in each stage of the round up to that point.

Partnership machines

In some rounds, you will be matched with one other person who is also participating in the experiment. In each such round, you will be randomly matched; your matching in one round is not related to your matching from another round. In these rounds you have to operate the machine in coordination with your partner. In each stage you will have to make your choices without knowledge of your partner's choice. As with the single-user machine, different short-term profits and probabilities will be associated with the two choices, so that if you both choose A you will tend to get a lower payoff than if you both choose B, but will be more likely to have a machine in good condition in the next round. If you choose A when your partner chooses B, or you choose B when your partner chooses A, the machine will produce the lower possible amount in this stage (as though you chose A) but will be more likely to be in bad condition in the next stage (as though you chose B).

As with the single-user machine, the 8 numbers that describe the machine will appear on near the top left of the screen. The numbers associated with the current condition of the machine will also be indicated to you in each stage in tables as explained in figure 1.16. The second screen capture on page 26 shows an example; in this case, the players are offered the green payoffs, which are 2 if either player chooses A and 4 if both choose B, as noted both in the grid to the left and in the bar on top; if both players choose A there is a 25% chance of switching to the red payoffs in the next stage, while if either player chooses B there is a 75% chance of switching to the red payoffs for

		partner's choice	
		A	B
your	A	2	2
choice	**B**	**2**	4

		partner's choice	
		A	B
your	A	0.25	0.75
choice	**B**	**0.75**	0.75

Figure 1.16: For this set of tables, which would appear on the left side of the screen for a partnership machine, if you choose B and your partner chooses A, you get a payoff of 2 and a 75% chance of switching to the other set of payoffs. In the experiment, the entries in the table may be different from the numbers here, but in each case your payoff and switching probability is determined by the row of the choice you make and the column of the choice your partner makes.

the next stage, which would offer payoffs of 0 and 2 (as shown in gray above the main screen).

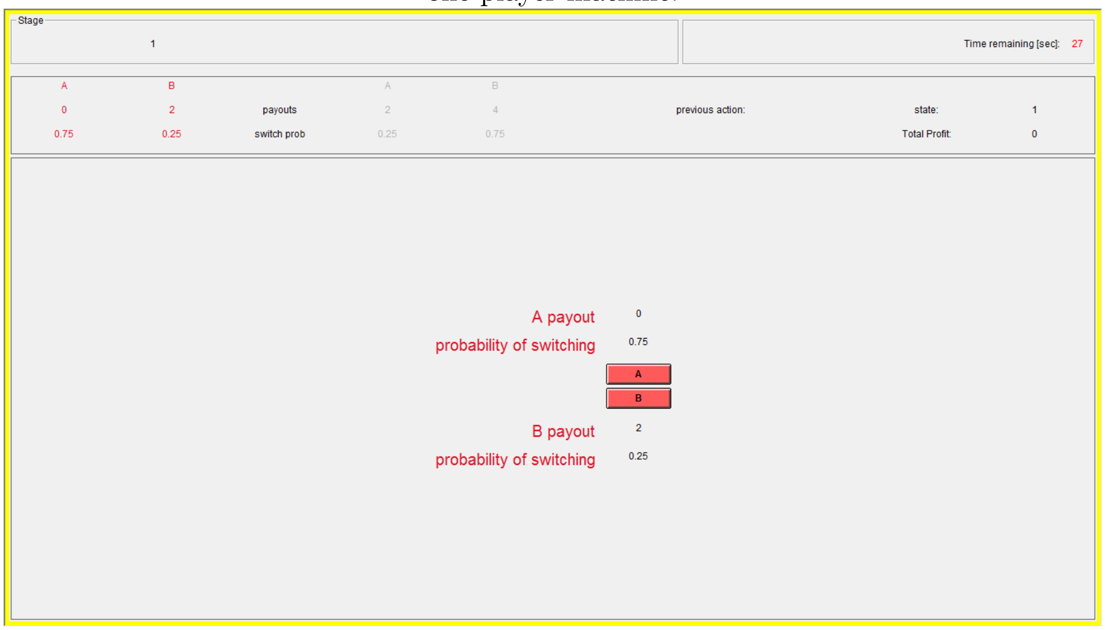
Payment

At the end of the experiment, your total earnings from each round (i.e. the total profits from all stages of all rounds) will be tallied up; you will be paid at a rate of $1 for every 100 shillings you have accumulate. (This is in addition to the $5 participation fee.) You will be paid your earnings in cash before leaving the experiment.

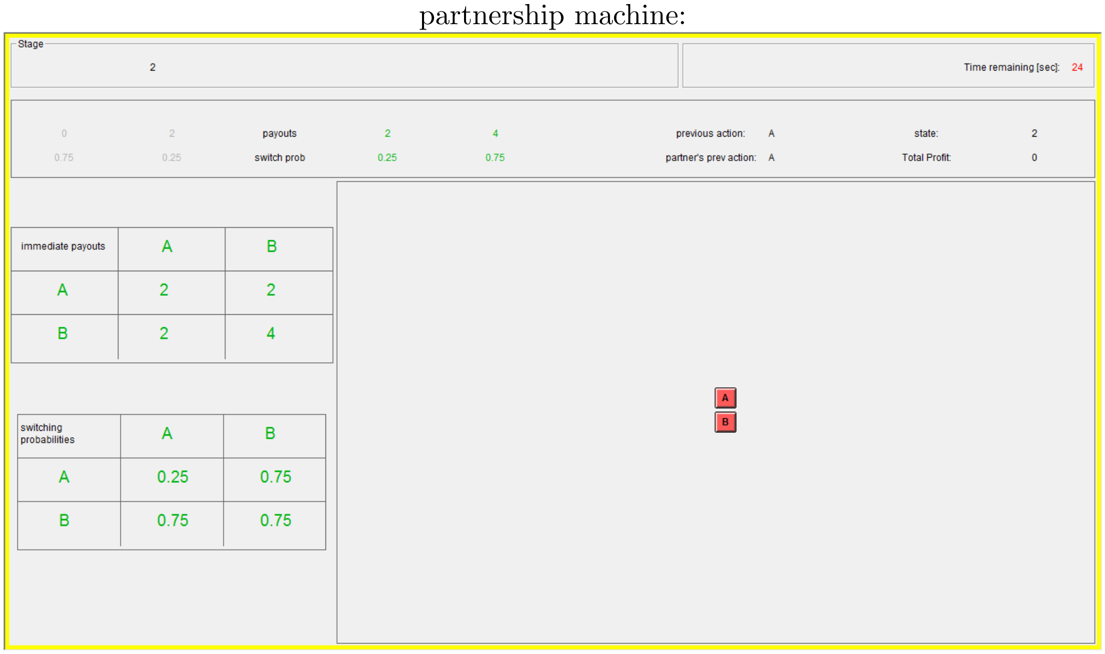
Summary

In the experiment you will make a series of choices between immediate payoffs and possible later payoffs, either alone or with a partner. Whether you are successful in earning profits in each round will depend both upon the decisions that you and the person you are paired with make, as well as on some degree of chance.

one-player machine:



partnership machine:



Chapter 2

Shortages and Runs

2.1 Introduction

In 1973, a congressman from a forested district attempted to raise concerns that the government had made insufficient efforts to procure paper products. On December 19, 1973, Johnny Carson's Tonight Show picked up on one of his releases, and made some jokes about an impending toilet paper shortage.[1] In response, shoppers the next day bought out the entire supply of toilet paper at many stores; the empty store shelves confirmed to other shoppers that, indeed, there was a shortage of toilet paper, and as soon as the store resupplied the run continued.

In times of actual or perceived shortage, it is common to want to stockpile something just in case, even if it is not needed right now. This, naturally, exacerbates the shortage, and may even be self-fulfilling. Gasoline is perhaps a more common example than toilet paper; if nobody tried to fill their gas tank precautionarily, there would be no shortage, but if an individual needs gas tomorrow, and is afraid people who do not need it tomorrow will take what is available, then that person is helping create the shortage today. In the event of a temporary supply disruption, this effect can prolong the shortage after supply comes back on line; if agents could coordinate, at that point everyone would be able to buy and consume at (or near) the "normal" level, but an expectation that the shortage will continue while consumers replenish their depleted

[1] The story is told at the priceonomics website, among other places; there is a youtube video of some of Carson's remarks.

stocks brings about the behavior that prolongs it.

This is related to the idea of liquidity premium; in some way it is its flip side. If there is a liquid market in which to sell an item, it can gain a liquidity premium: if I believe there is some chance that I will have a sudden need for any random other thing for which I can trade it (possibly indirectly), then by buying the item I'm buying that service as well as the item's use value itself. On the other hand, if there is a liquid market in which to *buy* an item, that can *lower* the value of the asset — as is perhaps more starkly noted when that market *does not* exist. The value of having some (storable) item now that may be hard, expensive, or even impossible to buy when I might suddenly need it includes a premium for the chance that I will need it a great deal at some point in the future and be unable to acquire it easily then. If I can get it when I need it, there is less value to acquiring it in advance.

The existence of shortages is a puzzle for pure classical economics; it requires that the price not increase to match quantity demanded with quantity supplied. Institutional factors of many kinds may prevent these adjustments — the employees of a price-offering retailer may not notice that the inventories are getting low until the shelves are empty, and those employees with the authority to raise prices may learn about it even later; the firm may have reputational reasons not to raise prices, if consumers are likely to punish the firm more for raising prices than for running out of stock; there may even be local regulations on prices, either hard caps or more ambiguous warnings about "price-gouging", that prevent the price from rising. Even if two consumer markets have the same instituional frictions on the supply-side, consumers may respond more forcefully to a prospective shortage in one market than another. Differences in both demand-side behavior and supply-side behavior further affect how agents form their expectations of the future, which then potentially feeds back into the behavior of both buyers and sellers. This paper seeks to address the question of what characteristics make a market more likely to get spooked and which characteristics make it more likely to recover. In some situations the answers to these questions may suggest policy responses, such as changing rationing rules imposed by policy-makers or adjusting formal price regulations; even where there are no policy responses, however, they are of interest to economic forecasters and anybody who wants to plan for the future.

Nichols and Zeckhauser (1977); Hendel, Dudine, and Lizzeri (2006); Bayer (2010); Mitraille and Thille (2014) study the stockpiling of storable goods in situations in which prices may differ from period to period. Their models have a small number of periods, strategic sellers, and markets that clear; this paper calls attention to an instability that exists in markets with inflexible prices and rationing, and which requires an infinite horizon.

Cavallo, Cavallo, and Rigobon (2014) study the retail market in Chile after its 2010 earthquake and Japan after its 2011 earthquake, where they note that prices changed very little even as shortages were severe and somewhat long-lasting — shortages in Chile were most severe a full two months after the initial disruption, and "a significant share of goods remained out of stock after six months." They especially note that "emergency" goods are more likely than others to stock-out after the earthquake, while "perishable" goods are less likely to do so. Roth (2008) identifies qualities that a market should have to be considered well-functioning; the one I am studying here he calls "thickness", but in financial literature especially it is "market liquidity". Several strains of theoretical literature are related to the emergence of market liquidity and its effect on the terms of trade of assets. Harrison and Kreps (1978) is seminal in a line of literature in which the ability of agents to trade an asset increases its value; in that model agents trade because they have different beliefs, and Morris (1996) characterizes the sets of beliefs they must hold to make the effect non-trivial. Other literatures depend on other asymmetries between agents to generate trade. Kiyotaki and Wright (1989, 1993) introduce a microfounded model of money by showing how the existence of bilateral markets with institutional barriers to multilateral trading can generate a liquidity premium for an asset. In those papers agents have different production capabilities and consumption preferences. Most of the financial economics papers on this topic use idiosyncratic "liquidity shocks" to motivate trading, as the agents suffering the shock seek to sell to the agents not suffering the shock. Rocheteau and Weill (2011) reviews a class of search models for assets in which the value of the asset is increased by the ease of selling it in the face of a liquidity shock. Duffie, Gârleanu, and Pedersen (2005, 2007) model a dealership-based market for financial markets and explore the attendant dynamics of price shocks. Vayanos and Wang (2011) provides a general three-period model (in the style of Diamond and Dybvig (1983)) to comprise several earlier similar models that explore the effects of different market imperfections on liquidity.

2.2 Model

My focus in this paper is the behavior of the demand side of the market, which will be assumed to have the convexity properties necessary for the separation theorems; accordingly, the demand side can be modeled separately from the supply side. The supply side, in fact, will not be modeled; the aims of the paper are intended not to be contingent on a particular model of the supply side of the market. Predictions for general market behavior will take supply behavior as input; policy recommendations will be made in terms of the supply behaviors that policymakers should attempt to cultivate.

2.2.1 Model

Each of a continuum of consumers receives an endowment of money m at each period t and uses it to buy b units of a storable consumption good; the consumer consumes c units of the good and $m_t - b_t p_t$ units of unmodeled goods, so that the utility in a given period is $u(c_t) - b_t p_t$. These other goods (and the quasilinear form of utility) keeps the marginal utility of money fixed and makes storing money undesirable.[2] While the consumption good can be stored, fraction λ of it depreciates in each period; $\lambda \to 1$ and $\lambda \to 0$ represent natural limits in which the good becomes nonstorable or nonperishable. Agents discount a period "flow" utility with a discount factor of δ. It is assumed that $\delta \in (0,1)$ and $\lambda \in [0,1]$. Letting C_t denote the amount of the good held at the beginning of the period, $b_t \geq 0$ the amount purchased in period t, and $c_t \in [0, C_t + b_t]$ the amount then consumed in period t, the agent's optimization problem is to maximize

$$\mathbb{E}_t \sum_{j=t}^{\infty} \delta^{j-t}(u(c_j) + m_j - b_j p_j)$$

subject to

$$C_{t+1} = (1-\lambda)(C_t + b_t - c_t)$$

$$b_t \geq 0$$

$$c_t \in [0, C_t + b_t]$$

where each agent takes the price p_t as given. u is strictly increasing, strictly concave, and twice continuously differentiable.

As we move to general equilibrium, it will *not* be assumed that p_t is necessarily allowed to rise to clear the market. In those situations in which p_t is held below the market-clearing price, there will be a shortage on the market; in the event of shortage, agents are assigned random priorities, based on which some agents get their chosen b_t, leaving the rest of the agents with $b_t = 0$.

---

[2] For example, it may be impossible to get a positive real return on stored money, whereupon the constant marginal utility of money and the discount factor make it suboptimal to store money.

The first order conditions for the optimizing consumer give

$$u'(c_t) \begin{cases} = p_t & b_t > 0 \\ \leq p_t & b_t = 0 \end{cases} \tag{2.1a}$$

$$u'(c_t) \begin{cases} \geq \delta(1-\lambda)\mathbb{E}_t\{u'(c_{t+1})\} & c_t = C_t + b_t \\ = \delta(1-\lambda)\mathbb{E}_t\{u'(c_{t+1})\} & c_t < C_t + b_t \end{cases} \tag{2.1b}$$

for agents choosing b_t and c_t to optimize expected discounted utility. (The b in (2.1a) is the agent's chosen b conditional on being able to purchase, while the b in (2.1b) is 0 if the agent was unable to buy.)

2.2.2 Equilibrium

We now consider the nature of equilibrium in this setting. We will consider "equilibrium" to be a sequence of prices, rationing levels, and quantities available for sale, starting at a particular time, along with a distribution of stockpiles with which agents enter the first period in which prices and rationing levels are given. Conditions for equilibrium will be that the quantities that optimizing agents buy in each period, with knowledge of current and future prices and rationing and current stockpiles, equal the quantity that is available for sale in that period.

A more natural definition of equilibrium might depend on specifying each agent's sequence of responses to that agent's history separately, and in using this simpler notion of "equilibrium" I am tacitly relying on a couple of properties of the model. One is that quantities and probabilities are specified exactly, such that the quantity demanded is equal to its expected value; one can think of a continuum of agents, and use the exact law of large numbers; alternatively, we could have a finite number of agents subject to the sequential service constraint of bank-run models, e.g. Wallace (1988), in which the macroeconomic quantity is exact and the probability results from a random ordering of which agents get to buy first. With this latter approach we still need to assume that the probability of being the last agent to buy and getting a "partially filled" order is 0, or at least that we are willing to neglect it. The other important tacit property is the serial independence of rationing; agents do not care much about their history beyond the level of stockpiles with which they enter a period; it does not affect future buying opportunities (conditional on the macroeconomic sequences).

If we are given an equilibrium starting at time t, then optimizing agents will buy and consume at time t, leading to a distribution of stockpiles entering time $t+1$; that distribution of stockpiles, along with the sequences of prices, rationing levels, and quantities available for sale starting at time $t+1$, will necessarily also be an equilibrium. We are not requiring the converse; an equilibrium beginning at time t need not be the natural consequence of any equilibrium beginning at time $t-1$. We are interested in evaluating when a (possibly measure-zero) "sunspot" event could cause agents to coordinate on a self-fulfilling macroeconomic course of action starting from time t, regardless of what expectations caused them to get to time t in a given state.

Formally, an equilibrium is

- A starting period $t \in \mathbb{Z}$
- A sequence of rationing levels $\{\pi_{t+k}\}_{k\geq 0}$ with $\pi_{t+k} \in [0,1]$
- A sequence of prices $\{p_{t+k}\}_{k\geq 0}$ with $p_{t+k} > 0$
- A sequence of quantities $\{q_{t+k}\}_{k\geq 0}$ with $q_{t+k} > 0$

such that

- Consumers choose functions for consumption $c : \mathbb{R} \to \mathbb{R}^+$ $(C_t \mapsto c_t)$ and purchases $b : \mathbb{R} \to \mathbb{R}^+$ $(C_t \mapsto b_t)$ to satisfy (2.1) for all times;
- $$\pi_t = \min\left\{1, \frac{q_t}{\langle b_t \rangle}\right\}$$

  for each t , where $\langle b_t \rangle$ represents an average over agents.

Note that with our strictly concave utility function, all consumers will have the same unique optimal strategy conditional on the macro variables; the equilibrium will be symmetric in this sense. Frequently we will specify

$$u = \frac{1}{1-\gamma} c^{1-\gamma} \qquad 0 < \gamma \neq 1$$

Note also that our attention is entirely on the demand side; there is no model of the supply side.

2.2.3 Optimizing Behavior

The approach to characterizing optimizing behavior by the consumer is as follows: in each period, the consumer will consume in such a quantity that the marginal utility of consumption is equal to the shadow value of the consumer's remaining stockpile. In a period in which the consumer is able to purchase, this shadow value is less than or equal to the price p_t at which that purchase would take place, with complementary slackness: the shadow value is equal to the price if the consumer purchases a positive quantity, and may only be less that the price if the consumer chooses not to purchase. In a period in which the consumer is able to purchase the consumer will also purchase for *future* consumption if there is a possibility of rationing to prepare for. The evolution of shadow value μ when an optimizing agent is unable to buy is determined by a straightforward Euler equation, and the agent can therefore determine what the optimal conditional consumption in each future period would be as long as the agent is unable to purchase again between time t and that future period.[3] The agent will aim to end period t with enough of a stockpile to be able to consume that amount.

[3] There will be no equilibrium in which $\mu \leq 0$ at any time for any agent who was able to buy at some previous point; a positive mass of such agents would demand an unlimited quantity at that previous point, so quantity demanded would not be positive and finite.

If each agent knows (and takes as given) the entire sequence of π_t and p_t, an optimizing agent's marginal utility of consumption will obey

$$\mu_t = \delta(1-\lambda)\left[\pi_{t+1}p_{t+1} + (1-\pi_{t+1})\mu_{t+1}\right]$$

where the μ_{t+1} on the right hand side is conditional on not being able to buy at time $t+1$; while the agent is unable to buy, μ will evolve according to

$$\mu_{t+1} = \frac{\delta^{-1}(1-\lambda)^{-1}}{(1-\pi_{t+1})}\mu_t - \frac{\pi_{t+1}}{(1-\pi_{t+1})}p_{t+1} \tag{2.2}$$

and μ_t will drop to p_t when the agent is able to buy.[4]

If we recursively define an effective discount factor

$$D_t = (1-\pi_t)\delta(1-\lambda)D_{t-1} \tag{2.3}$$

and if all $\mu_t \geq p_t$, then at time t an agent that last bought at time $t-j$ has

$$D_t\mu_{t,t-j} = D_{t-j}p_{t-j} - \sum_{i=0}^{j-1}\frac{\pi_{t-i}}{(1-\pi_{t-i})}D_{t-i}p_{t-i} \tag{2.4}$$

and consumes $c_{t,t-j}$ such that $u'(c_{t,t-j}) = \mu_{t,t-j}$.

For u strictly concave, $\mu = u'(c)$ provides a one-to-one mapping between c and μ that holds at all times; this can be combined with the above relationships to determine what the agent's consumption pattern would look like if the agent never succeeds in buying again. If that is the case, then the agent will simply consume the amount in stock, adjusting for the deterioration of stocks over time:

$$(1-\lambda)^{-1}C_{t+1} = \sum_{j=1}^{\infty}\frac{c_{t+j,t}}{(1-\lambda)^j} \tag{2.5}$$

where $c_{t+j,t+1} = c_{t+j,t}$ if the agent does not buy at time $t+1$. If the agent does get to buy again, μ becomes p, and the sequence of c_t if the agent cannot buy after *that*

---

[4] If μ can dip below p an agent would have to be allowed to sell some of the stockpile at the market price in order for these formulas to hold; note that positive quantities imply $\pi_{t+1} > 0$, and that future conditional values of μ are irrelevant if $\pi = 1$ between now and then, so we will frequently suppose that $\pi \in (0, 1)$.

changes; an agent who does buy, having last bought at $t-k$, buys

$$\sum_{j=0}^{\infty}(1-\lambda)^{-j}\left(c^*_{t+j,t}-c^*_{t+j,t-k}\right)$$

(enough to match the new optimal consumption path). The fraction of agents in this group is

$$\pi_{t-k}(1-\pi_{t-k+1})\cdots(1-\pi_{t-1})\pi_t$$

so the total quantity purchased in a period is

$$\sum_{j=0}^{\infty}(1-\lambda)^{-j}c^*_{t+j,t}\pi_t-\sum_{k=1}^{\infty}\sum_{j=0}^{\infty}(1-\lambda)^{-j}\left(c^*_{t+j,t-k}\right)\pi_{t-k}(1-\pi_{t-k+1})\cdots(1-\pi_{t-1})\pi_t \quad (2.6)$$

$$=\sum_{k=1}^{\infty}\sum_{j=0}^{\infty}(1-\lambda)^{-j}\left(c^*_{t+j,t}-c^*_{t+j,t-k}\right)\pi_{t-k}(1-\pi_{t-k+1})\cdots(1-\pi_{t-1})\pi_t \quad (2.7)$$

When studying the recovery of a system from a shortage, it is useful to note that if $\pi_{t+j}=1$ for some $j\geq 1$, then $c^*_{t+j,t-k}=0$ for all $k\geq 0$, so the sum over j effectively gets cut off.

For much of this paper, we will work with HARA utility, for which $\mu=c^{-\gamma}$ for some fixed γ, in which case

$$c^{-\gamma}_{t+j,t}=D^{-1}_{t+j}D_t\mu_t-D^{-1}_{t+j}\sum_{k=1}^{k=j}\frac{\pi_{t+k}}{(1-\pi_{t+k})}D_{t+k}p_{t+k}$$

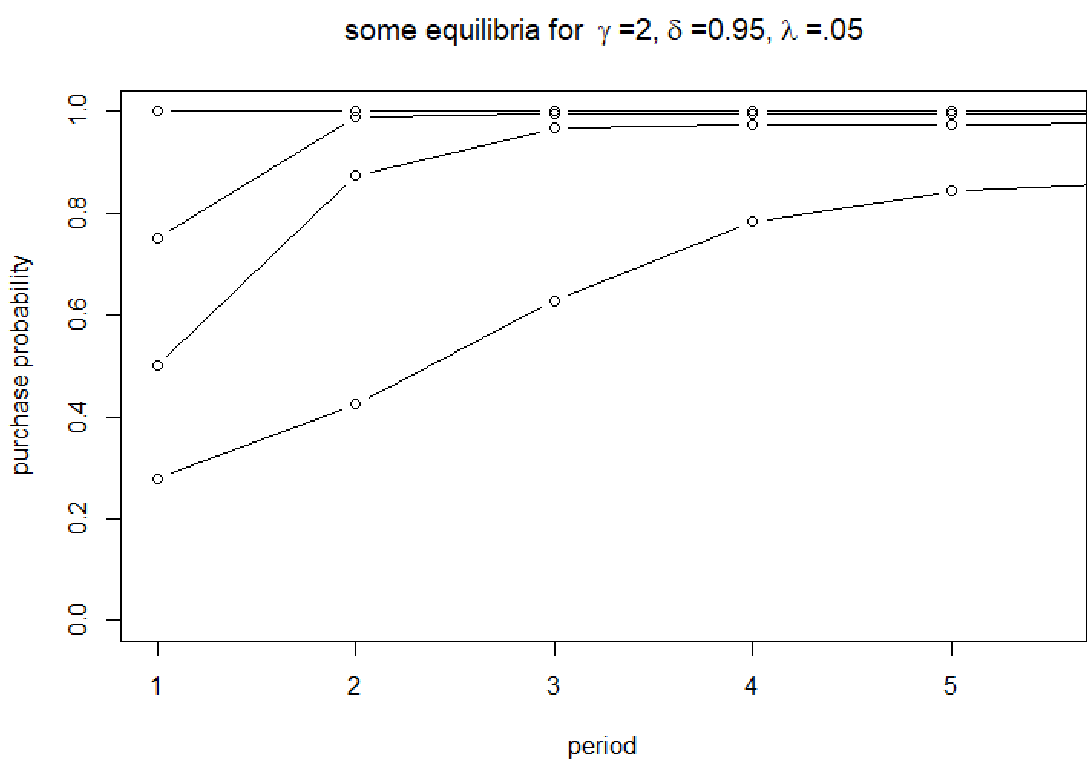
2.3 Markets with Multiple Equilibria

2.3.1 Illustrative Example

Consider the scenario outlined in figure 2.1, where agents know that prices will be constant, a such a level that if they could reliably buy at that price they would buy and consume (in each period) a quantity that is normalized to 1. The agents also anticipate being unable to reliably buy, however; the probability of being able to buy in each of the first six periods is shown in the top row. The average agent enters period 2 with a quantity of .685 stored up from the previous period (some agents have 2.458

rationing:	0.5	0.875	0.9671	0.9737	0.9752	0.9763
prices:	1	1	1	1	1	1
quantities:	1	1	1	1	1	1
stockpile:	0.5	0.5543	0.5453	0.533	0.5207	0.5086
consumption:	0.5	0.9208	0.9814	0.985	0.9857	0.9862
stocktarget:	2	1.618	1.561	1.545	1.532	1.519
conditional	1	0.7324	0.173	0.02698	0.004039	0.00059
consumptions:		1	0.4829	0.08337	0.01251	0.001829
			1	0.4422	0.07272	0.01065
				1	0.4322	0.06911
					1	0.4239
						1

Figure 2.1: $\delta = 0.95$, $\lambda = 0.05$, $\gamma = 2$. The first 50% of the consumers that get to the store each buy twice as much as usual, so the store runs out; if they are unable to buy again in the first three periods, they consume 17.3% of the classical equilibrium amount in period 3.



while others have 0); an agent able to buy in the second period, anticipating the rest of the sequence of probabilities, seeks to end the period with 2.037 in storage and to consume 1 in period 2, so the average agent wishes to buy 2.352; as only 42.5% of agents successfully buy, they buy a total quantity of $2.352 \times .4253 = 1$.

2.3.2 Main Result

Proposition 1. *Suppose*

- $u = (c^{(1-\gamma)})/(1-\gamma)$ *for some* $\gamma > 0, \gamma \neq 1$
- $\lambda > 0$
- *the quantity is positive, and is the same in every period*
- *the price is positive, and is the same in every period*
- *the price and quantity are such that* $\pi_{t+1} = \pi_{t+2} = \cdots = 1$ *completes an equilibrium*

then

1. *there are other equilibria with the same quantities and prices, but with* $\pi_{t+k} < 1 \forall k \geq 1$, *if* $\gamma > 1$
2. *there are no other equilibria with the same quantities and prices, but with* $\pi_{t+1} < 1$, *if* $\gamma < 1$.

We call the $\pi = 1$ equilibrium "the classical equilibrium".

If $\gamma < 1$, the classical equilibrium is "stable" in this sense; if $\gamma > 1$, however, the economy is at risk of a sunspot shortage — there will be patterns of behavior such that a shortage results entirely from self-fulfilling expectations.

Lambda

Note that with prices constant, an agent who is able to buy will consume such that the marginal cost of utility is equal to the price, while any agent who is unable to buy will have a strictly higher marginal cost of utility, which is to say a lower consumption. In any equilibrium as described, if there is rationing, then the agents are on average consuming less than they are buying.

One consequence of this is that rationing equilibria of the sort just described require $\lambda > 0$. If $\lambda = 0$, the aggregate stockpiles of the agents never decrease. With $\delta < 1$, in order for that to be consistent with optimizing behavior, the $\pi_t \leq \bar{\pi} < 1$ have to be bounded away from 1, which means that stokpiles are in fact growing without bound, which ultimately becomes inconsistent with its being optimal not to consume some of it.

2.3.3 Summary of Proof

Given $\gamma > 1$, there are values of $\pi_{t+1} < 1$ such that there is a sequence $\pi_{t+2}, \pi_{t+3}, \cdots$ such that total quantity demanded in each period is equal to the value it would obtain if all the π were 1.

The optimizing behavior as §2.2.3 defines a mapping from sequences of π to sequences of quantities demanded; each step $\pi \mapsto \mu \mapsto c^* \mapsto \tilde{q}$ is smooth. A compact (with the ℓ^p topology) subset K of the space of sequences $\pi_{t+2}, \pi_{t+3}, \cdots$ is constructed on which $\sum_{i=2}^{\infty}(\tilde{q}_i - 1)^2$ is defined and smooth, and attains its minimum on K only where $\tilde{q}_i = 1$ for all $i \geq 1$.

The bulk of the work is showing that $-\partial(\tilde{q}_1, \cdots, \tilde{q}_u)/\partial(\pi_{t+2}, \pi_{t+3}, \cdots, \pi_{t+u+1})$ is positive definite on the interior of K; some attention is also required to constructing K to ensure that the minimum of S on K is attained on the interior.

Details are relegated to the appendix (page 47).

2.3.4 Intuition of the Result

In this section we illustrate the dynamics of a market in a self-fulfilling shortage using approximations that hold for $\pi \approx 1$.

If $\epsilon = 1 - \pi_{t+1}$,

$$\mu_{t+1} = \frac{\delta^{-1}(1-\lambda)^{-1}\mu_t - p_{t+1}}{\epsilon} + p_{t+1}$$

If

$$\epsilon \ll \frac{\delta^{-1}\mu_t}{p_{t+1}} - 1,$$

then

$$\mu_{t+1} \approx \frac{\delta^{-1}(1-\lambda)^{-1}\mu_t}{\epsilon}$$

If prices are constant and the probability of rationing is much less than $\delta^{-1} - 1$, then the top-up level is

$$\approx p^{-1/\gamma} + \delta^{1/\gamma}(1-\lambda)^{1/\gamma-1}\epsilon_{t+1}^{1/\gamma}p^{-1/\gamma}$$

(but somewhat higher for positive ϵ) with purchases

$$\approx p^{-1/\gamma}(1-\epsilon_t)\left(1 + \delta^{1/\gamma}(1-\lambda)^{1/\gamma-1}\left(\epsilon_{t+1}^{1/\gamma} - (1-\lambda)\epsilon_t^{1/\gamma}\right)\right)$$

$\approx p^{-1/\gamma}+$

$$p^{-1/\gamma}\left(\delta^{1/\gamma}(1-\lambda)^{1/\gamma-1}\left(\epsilon_{t+1}^{1/\gamma} - (1-\lambda)\epsilon_t^{1/\gamma}\right) - \epsilon_t\right)$$

The quantities purchased are more or less constant and equal to $p^{-1/\gamma}$ if

$$(1-\lambda)^{-1}\epsilon_{t+1}^{1/\gamma} = \delta^{-1/\gamma}(1-\lambda)^{-1/\gamma}\epsilon_t + \epsilon_t^{1/\gamma} \tag{2.8}$$

which gives

$$\frac{\epsilon_{t+1}}{\epsilon_t} = \left(\delta^{-1/\gamma}(1-\lambda)^{1-1/\gamma}\epsilon_t^{1-1/\gamma} + (1-\lambda)\right)^{\gamma}$$

For $\gamma > 1$ and $\lambda > 0$,[5] as $\epsilon_t \to 0$,

$$\frac{\epsilon_{t+1}}{\epsilon_t} = (1-\lambda)^{\gamma}$$

[5] For $\lambda = 0$, the expression stays bigger than 1 for small positive ϵ.

while for $\gamma < 1$, if ϵ_t is small, $\epsilon_{t+1} > \epsilon_t$; in that case there is no asymptotically non-rationing equilibrium with positive rationing. (If ϵ_t is too small, the right hand side exceeds 1.)

In words, an anticipated shortage at time $t+1$ results in higher quantity demanded at time t, creating a shortage at time t; if $\gamma > 1$ the shortage created at time t is larger than the shortage anticipated at the next period. If $\gamma < 1$, then a shortage at time t will only occur in response to the expectation of ever more severe shortages. Within a finite number of periods, there is no $\pi_{t+k+1} \in [0,1]$ that would cause each agent to demand a high enough quantity to support the preceding sequence of values of π.

2.3.5 Welfare analysis

Consider agents with ex ante total utility of

$$\mathbb{E}_t \sum_{j=t}^{\infty} \delta^{j-t}(u(c_j) - b_j p_j)$$

with u strictly concave, increasing, and differentiable in c. u is a strictly convex function of μ (e.g. from (Rockafellar, 1970, theorem 23.5)); for example, in the CRRA framework in which we are mostly working, current consumption utility is

$$u = \frac{1}{1-\gamma} c^{1-\gamma} = \frac{1}{1-\gamma} \mu^{1-1/\gamma}$$

Proposition 2. *Consider two equilibria with*

- *the same quantities*
- *the same prices*
- $\pi = 1$ *at all times in one of the equilibria.*

The the equilibrium with $\pi = 1$ *Pareto dominates that with* $\pi < 1$*.*

The sums

$$\mathbb{E}_t \sum_{j=t}^{\infty} \delta^{j-t} u_j \text{ and } \mathbb{E}_t \sum_{j=t}^{\infty} \delta^{j-t} b_j p_j$$

converge separately. As long as $\mu_t \geq p_t$, $\mathbb{E}_t \sum_{j=t}^{\infty} \delta^{j-t} u_j$ is lower than it would be in the classical equilibrium; on the other hand, as long as agents start out symmetrically, $\mathbb{E}_t \sum_{j=t}^{\infty} \delta^{j-t} b_j p_j$ is the same for any two equilibria in which agents buy (in aggregate) the same quantity (matching period-by-period).

For fixed sequences of p and π,

$$T_t = \sum_{j=0}^{\infty} \frac{c_{t+j,t}}{(1-\lambda)^j} (2.5)$$

is strictly convex and strictly decreasing in μ_t — each term in the sum is — thus μ_t is stictly convex and strictly decreasing as a function of T_t

As long as $\delta^{-1}(1-\lambda)^{-1}\mu_t > p_{t+1}$ and $\delta^{-1}(1-\lambda)^{-1}p_{t+k} > p_{t+k+1}$ for all k, T_t is monotonically decreasing[6] in future π, for any fixed μ_t; the entire curve shifts down if any π increases toward 1. Thus μ_t as a function of T_t also shifts down (because μ is strictly decreasing in T); at least for small enough stockpiles that $\mu_t > \delta(1-\lambda)p_{t+1}$, agents coming in with stockpiles will also have a reduction in welfare from shortages.

2.4 Supply response: eliminating rationing

The result above applies if prices are unresponsive to shortages. If prices rise, then quantity demanded goes down; if buyers anticipate that prices *would* rise sufficiently if there *were* a shortage, then they also anticipate that the shortage will cease if it begins, and the bad equilibria are eliminated. To be clear, this is ultimately off-path; prices do not actually change, but it is commonly believed that they would if there were rationing.[7]

Our primary interest is in $q = 1$ at all times, so "price flexibility" will not be p_{t+k} as a function of q_{t+k} as it typically is. Instead, we allow p_{t+k} to be a function of the

---

[6]Strictly decreasing in π_{t+l} if $\pi_{t+k} < 1 \forall k : 1 \leq k < l$; otherwise it is unchanged.

[7]This simplifies the welfare analysis by obviating the concern that higher prices hurt consumers, and need to be balanced against reliability. This hypothetical welfare concern does, however, introduce a question of policy credibility: if there is a policy that allows prices to rise substantially, it may be harder to establish the common belief that it would be followed. It is worth noting, therefore, that the shortage in figure 2.1 can be avoided by threatening to raise the price by 6%.

sequence of π. Given a sequence of such functions (one function for each k), a sequence of π implies a sequence of p. Our equilibrium concept thus changes somewhat for this section: instead of defining the equilibrium as sequences of π, q, and real numbers p that satisfy demand-side constraints, an equilibrium is a sequence of π, q, and functions p that, evaluated on the sequence π, give a sequence of real numbers p that satisfy the requirements of equilibrium from §2.2.2. In exploring the structure of equilibria, we still think of q and p as given and seek sequences of π that satisfy the criteria jointly with them.

There are many sequences of functions p that we could explore, even restricting ourselves to those that are continuously differentiable and for which there is an equilibrium with $\pi_{t+k} = 1$ and $q_{t+k} = 1$ for all k. (We will restrict ourselves to that class of functions.) A formally simple subclass of such functions is those that depend only on rationing in the previous period, $p_{t+k}(\pi_{t+k-1})$, with $p(1) = 1$ and p decreasing in π_{t+k-1} (i.e. increasing as π decreases). An economically more intuitive simple subclass is that for which each p_{t+k} is a function of C_{t+k}, with $p(0) = 1$ and p increasing in C_{t+k}.

As before, the existence of equilibria in which $\pi_{t+k} < 1$ but $\pi_{t+k} \to 1$ as $k \to \infty$ will hinge on the behavior of

$$\frac{\partial q_{t+k+j}}{\partial \pi_{t+k}}$$

for different values of j when the various π are near 1. Because the total quantity demanded $\tilde{q}$ is smooth in p and π and p is continuously differentiable in π,

$$\frac{\partial \tilde{q}_t}{\partial \pi_{t+k}} = \left.\frac{\partial \tilde{q}_t}{\partial \pi_{t+k}}\right|_p + \sum_m \left.\frac{\partial \tilde{q}_t}{\partial p_{t+m}}\right|_\pi \frac{\partial p_{t+m}}{\partial \pi_{t+k}}$$

as long as the sum is well-defined (its terms absolute summable). If the sum is much smaller (in absolute size) than $\partial \tilde{q}_t / \partial \pi_{t+1}$, then it will not affect whether there is such an equilibrium; price-sensitivity that is big enough to matter has to diverge as $\pi \to 1$ at least as quickly as $\partial \tilde{q}_t / \partial \pi_{t+1}$ and $\partial C_{t+1} / \partial \pi_{t+1}$ do. We therefore concentrate on p as a function of C.

2.4.1 Price Response to Stockpiles

If $\epsilon = 1 - \pi_{t+1}$,

$$\mu_{t+1} = \frac{\delta^{-1}(1-\lambda)^{-1}\mu_t - p_{t+1}}{\epsilon} + p_{t+1}$$

If

$$\epsilon \ll \frac{\delta^{-1}\mu_t}{p_{t+1}} - 1,$$

then

$$\mu_{t+1} \approx \frac{\delta^{-1}(1-\lambda)^{-1}\mu_t}{\epsilon}$$

If prices are constant and the probability of rationing is much less than $\delta^{-1} - 1$, then the top-up level is

$$\approx p^{-1/\gamma} + \delta^{1/\gamma}(1-\lambda)^{1/\gamma - 1}\epsilon_{t+1}^{1/\gamma} p^{-1/\gamma}$$

with purchases

$$\approx (1-\epsilon_t)\left(p^{-1/\gamma}\left(1 + \delta^{1/\gamma}(1-\lambda)^{1/\gamma-1}\epsilon_{t+1}^{1/\gamma}\right) - C_t\right)$$

Suppose p is a function of C; for present purposes, $p = 1 + \alpha C$. Now the quantity purchased q_t obeys

$$q_t + \epsilon_t q_t \approx 1 - \frac{\alpha}{\gamma} C_t + \delta^{1/\gamma}(1-\lambda)^{1/\gamma-1}\epsilon_{t+1}^{1/\gamma} - C_t$$

with $q_t = 1$ if

$$\epsilon_t \approx \delta^{1/\gamma}(1-\lambda)^{1/\gamma-1}\epsilon_{t+1}^{1/\gamma} - \frac{\alpha+\gamma}{\gamma} C_t$$

Now, suppose again that this is all anticipated by optimizing agents, and that $C_t \approx \delta^{1/\gamma}(1-\lambda)^{1/\gamma}\epsilon_t^{1/\gamma}$, with $\gamma > 1$ such that (with $\epsilon \approx 0$) $\epsilon \ll \epsilon^{1/\gamma}$, then

$$\epsilon_{t+1}^{1/\gamma} \approx \frac{\alpha+\gamma}{\gamma}(1-\lambda)\epsilon_t^{1/\gamma},$$

such that there is, as for $\gamma < 1$, no asymptotically vanishing rationing equilibrium if

$$\frac{\alpha+\gamma}{\gamma}(1-\lambda) > 1,$$

which is to say if

$$\alpha > \left((1-\lambda)^{-1} - 1\right)\gamma \tag{2.9}$$

If agents anticipate that (for whatever reason) prices would rise sufficiently quickly if stockpiles rise, then these equilibria go away.

2.5 Modeling questions

2.5.1 Interpretting γ

We are working with a fairly specific functional form for utility: additively separable between money and the consumption good, and linear in money, such that the marginal utility of wealth is constant. The results generalize somewhat, and in intuitive ways.

First, note that if the marginal utility of wealth is η, $u'(c) \leq \eta p$ with equality for $c > 0$; we have fixed $\eta = 1$ on the grounds that it is unlikely to change very much in a shortage, but the γ parameter may be easier to interpret in terms of how consumption changes with income. If η is allowed to vary then γ is proportional to the elasticity of consumption with respect to the marginal utility of wealth, to wit

$$c^* = \arg\max_c m + u(c) - \eta(m + c/p + I) \Rightarrow$$

$$u'(c^*) = \eta/p$$

$$\frac{\partial u'(c^*)}{\partial c^*} = \frac{1}{p}\frac{\partial \eta}{\partial c^*}$$

if u is twice differentiable and increasing at c^*;

$$\begin{aligned}\frac{c^*}{u'(c^*)}\frac{\partial u'(c^*)}{\partial c^*} &= \frac{c^*}{u'(c^*)}\frac{1}{p}\frac{\partial \eta}{\partial c^*}\\ &= \frac{c^*}{\eta}\frac{\partial \eta}{\partial c^*}.\end{aligned}$$

For a given income elasticity of utility, then, γ is proportional to the income elasticity of consumption for an optimizing agent. Accordingly we will identify low γ with "luxury" goods and high γ with "necessities".

A second point to note is that the result is driven by the relationship between c and u at low values of c; if the utility of consumption is such that

$$\frac{cu''}{u'}$$

is not constant but is bounded above or below 1 as $c \to 0$ then our results should generalize to that utility function.

2.6 Conclusion

I have noted that, in markets with fixed prices and, where necessary, rationing, there can be shortages even when the quantity supplied would be enough to satisfy contemporaneous demand. I have demonstrated that this holds for storable goods if they are particularly "inelastic" in that consumers' willingness to pay rises very steeply as their consumption goes down, and that some price flexibility can prevent this.

This is useful for private agents wishing to predict which markets are susceptible to such "runs"; it also suggests some policy responses. Most clearly, this suggests an additional cost to any policies that are likely to reduce price flexibility, and an additional benefit to policies that reduce stickiness, especially in the face of shortages. Quantity flexibility could replace price flexibility; this result argues on the side of retaining a national strategic reserve of petroleum and petroleum products and signalling a willingness to release them in response to temporary shortages.

I have noted previously that Cavallo, Cavallo, and Rigobon (2014) find empirical results in the face of actual shortages that are similar to the results of this model. They studied supply disruptions, however; there is no "classical equilibrium" with non-shortage prices in that case. After many disasters, however, the retail-level shortage outlasts the restoration of pre-disaster supplies; in those situations, even where policy-makers might wish to mitigate supply disruptions in other ways, any steps that encourage prices to adjust in response to on-going retail disruptions once the supply disruptions are over would help stabilize the market more quickly.

2.7 Appendix

2.7.1 c and its Derivatives

The quantity purchased at time t has been written as a function of π and c^*; the c^* have been written as functions of μ. Here we place some bounds on the derivative of the q with respect to the π, holding the p constant, by use of the chain rule; first we take the derivative of μ, then of c^*, and then q.

From (2.4),

$$\mu_{t,t-j} = D_t^{-1} D_{t-j} p_{t-j} - \sum_{i=0}^{j-1} \frac{\pi_{t-i}}{(1-\pi_{t-i})} D_t^{-1} D_{t-i} p_{t-i};$$

$$\frac{\partial (D_t^{-1} D_{t-j})}{\partial \pi_{t-k}} = \frac{D_t^{-1} D_{t-j}}{1-\pi_{t-k}} \text{ if } 0 \le k < j$$

$$\frac{\partial (D_t^{-1} D_{t-i} \pi_{t-i}/(1-\pi_{t-i}))}{\partial \pi_{t-i}} = \frac{D_t^{-1} D_{t-i}}{(1-\pi_{t-i})^2}$$

If $0 \le k < j$

$$\begin{aligned} \frac{\partial \mu_{t,t-j}}{\partial \pi_{t-k}} =& \frac{D_t^{-1} D_{t-j}}{1-\pi_{t-k}} p_{t-j} - \frac{D_t^{-1} D_{t-k}}{(1-\pi_{t-k})^2} p_{t-k} \\ & - \sum_{i=k+1}^{j-1} \frac{D_t^{-1} D_{t-i}}{1-\pi_{t-k}} \frac{\pi_{t-i}}{(1-\pi_{t-i})} p_{t-i} \end{aligned}$$

while otherwise it is 0.

Note that if, for some fixed $\bar{\epsilon} > 0$, $1 - \pi_{t+k} \le \bar{\epsilon}$ for $i \le k \le j$, then

$$\frac{\partial \mu_{t,t-j}}{\partial \pi_{t-k}} \le \bar{\epsilon}^{j-1} \delta^j (1-\lambda)^j p_{t-j}.$$

Of particular interest, given μ_{t-1}, for any t, we have

$$\frac{\partial \mu_t}{\partial \pi_t} = \frac{\delta^{-1}(1-\lambda)^{-1} \mu_{t-1} - p_t}{(1-\pi_t)^2}$$

Now we can find the derivative of c^* using the chain rule;

$$\frac{\partial c^*_{t+k,t}}{\partial x} = \frac{\partial \mu_{t+k,t}^{-1/\gamma}}{\partial x} = -\frac{1}{\gamma} \mu_{t+k,t}^{-1/\gamma - 1} \frac{\partial \mu_{t+k,t}}{\partial x}$$

$$= -\frac{1}{\gamma}\frac{c^*_{t+k,t}}{\mu_{t+k,t}}\frac{\partial \mu_{t+k,t}}{\partial x}$$

If, for some fixed $\bar{\epsilon} > 0$, $1 - \pi_{t+k} \leq \bar{\epsilon}$ for $i \leq k \leq j$, then

$$\left|\frac{\partial c^*_{t,t-j}}{\partial \pi_{t-k}}\right| = \left|-\frac{1}{\gamma}\mu_{t,t-j}^{-1/\gamma-1}\frac{\partial \mu_{t,t-j}}{\partial \pi_{t-k}}\right| \leq \frac{1}{\gamma}\mu_{t,t-j}^{-1/\gamma-1}\bar{\epsilon}^{j-1}\delta^j(1-\lambda)^j p_{t-j}. \tag{2.10}$$

Using (2.4) again,

$$\mu_{t,t-j} = D_t^{-1}D_{t-j}p_{t-j} - \frac{\pi_{t-(j-1)}}{(1-\pi_{t-(j-1)})}D_t^{-1}D_{t-(j-1)}p_{t-(j-1)} - \sum_{i=0}^{j-2}\frac{\pi_{t-i}}{(1-\pi_{t-i})}D_t^{-1}D_{t-i}p_{t-i}$$

For $\bar{\epsilon} < 1/2$ and $j \geq 2$, we can write this as

$$D_t^{-1}D_{t-j}\left(p_{t-j} - \delta(1-\lambda)\pi_{t-(j-1)}p_{t-(j-1)} - \sum_{i=0}^{j-2}\frac{\pi_{t-i}}{(1-\pi_{t-i})}D_{t-j}^{-1}D_{t-i}p_{t-i}\right)$$

$$\geq D_t^{-1}D_{t-j}p\left(1 - \delta(1-\lambda)(1-\bar{\epsilon}) - \frac{\bar{\epsilon}(1-\bar{\epsilon})}{1-\delta(1-\lambda)\bar{\epsilon}}\right)$$

when p is constant. Thus for $j \geq 2$, $\left|\frac{\partial c^*_{t,t-j}}{\partial \pi_{t-k}}\right| \leq$

$$\frac{1}{\gamma}\bar{\epsilon}^{j-1}\delta^j(1-\lambda)^j p^{-1/\gamma}\delta^{j/\gamma+j}(1-\lambda)^{j/\gamma+j}\bar{\epsilon}^{j/\gamma+j}\left(1 - \delta(1-\lambda)(1-\bar{\epsilon}) - \frac{\bar{\epsilon}(1-\bar{\epsilon})}{1-\delta(1-\lambda)\bar{\epsilon}}\right)^{-1/\gamma-1}$$

which simplifies to

$$\bar{\epsilon}^{j/\gamma+2j-1}\frac{p^{-1/\gamma}}{\gamma}\left(\delta(1-\lambda)\right)^{j/\gamma+2j}\left(1 - \delta(1-\lambda) + \delta(1-\lambda)\bar{\epsilon} - \frac{\bar{\epsilon}(1-\bar{\epsilon})}{1-\delta(1-\lambda)\bar{\epsilon}}\right)^{-1/\gamma-1} \tag{2.11}$$

Given μ_{t-1},

$$\frac{\partial c_{t,t-k}}{\partial \pi_t} = -\frac{1}{\gamma}\mu_{t,t-k}^{-1/\gamma-1}\frac{\partial \mu_t}{\partial \pi_t} = -\frac{1}{\gamma}\mu_{t,t-k}^{-1/\gamma-1}\frac{\delta^{-1}(1-\lambda)^{-1}\mu_{t-1} - p_t}{(1-\pi_t)^2}$$

$$= -\frac{1}{\gamma}\mu_{t,t-k}^{-1/\gamma-1}\frac{\mu_t - p_t}{1-\pi_t}$$

2.7.2 Constructing a Domain

We have (2.2)

$$\mu_{t+1} = \frac{\delta^{-1}(1-\lambda)^{-1}}{(1-\pi_{t+1})}\mu_t - \frac{\pi_{t+1}}{(1-\pi_{t+1})}p_{t+1};$$

in particular,

$$\mu_{t+1,t} = \frac{\delta^{-1}(1-\lambda)^{-1}}{(1-\pi_{t+1})}p_t - \frac{\pi_{t+1}}{(1-\pi_{t+1})}p_{t+1}$$
$$\mu_{t,t-1} = \frac{\delta^{-1}(1-\lambda)^{-1}}{(1-\pi_t)}p_{t-1} - \frac{\pi_t}{(1-\pi_t)}p_t$$
$$\mu_{t+1,t-1} = \frac{\delta^{-1}(1-\lambda)^{-1}}{(1-\pi_{t+1})}\left(\frac{\delta^{-1}(1-\lambda)^{-1}}{(1-\pi_t)}p_{t-1} - \frac{\pi_t}{(1-\pi_t)}p_t\right) - \frac{\pi_{t+1}}{(1-\pi_{t+1})}p_{t+1}$$

For $p_{t+1} = p_t = p_{t-1} = 1$, $\gamma > 1$, and (given) $\pi_t < 1$ but sufficiently close to 1,

$$\mu_{t,t-1}^{-1/\gamma} + (1-\lambda)^{-1}\mu_{t+1,t-1}^{-1/\gamma} - (1-\lambda)^{-1}\mu_{t+1,t}^{-1/\gamma} = 1/\pi_t - 1$$

has a solution for $\pi_{t+1} \in (1-(1-\lambda)(1-\pi_t), 1)$. Let $\bar{\epsilon}_t = 1 - \pi_t$, then define $\bar{\epsilon}_{t+1}$ as $1-$ that π_{t+1} that solves this equation; we similarly define $\bar{\epsilon}_{t+k+1}$ in terms of $\bar{\epsilon}_{t+k}$, and thus define a sequence of $\bar{\epsilon}$. We will define

$$K = \prod_k [0, \bar{\epsilon}_{t+k}] \subset \ell^2 \tag{2.12}$$

Now, consider

$$1 - c^*_{t,t-1} + (1-\lambda)^{-1}\left(c^*_{t+1,t} - c^*_{t+1,t-1}\right)$$

which $= 1/(1-\bar{\epsilon}_t) - 1$ if $\pi_t = 1 - \bar{\epsilon}_t$ and $\pi_{t+1} = 1 - \bar{\epsilon}_{t+1}$. It is decreasing in π_{t+1} and is increasing in π_t; thus, for any $\pi_t \in [1-\bar{\epsilon}_t, 1)$, if $\pi_{t+1} = 1 - \bar{\epsilon}_{t+1}$,

$$\sum_{j=0}^{\infty}(1-\lambda)^{-j}\left(c^*_{t+j,t} - c^*_{t+j,t-1}\right) \geq 1 - c^*_{t,t-1} + (1-\lambda)^{-1}\left(c^*_{t+1,t} - c^*_{t+1,t-1}\right) \geq 1/(1-\bar{\epsilon}_t) - 1$$

(2.6) gives $q_t =$

$$\sum_{j=0}^{\infty}(1-\lambda)^{-j}c^*_{t+j,t}\pi_t + \sum_{k=1}^{\infty}\sum_{j=0}^{\infty}(1-\lambda)^{-j}\left(-c^*_{t+j,t-k}\right)\pi_{t-k}(1-\pi_{t-k+1})\cdots(1-\pi_{t-1})\pi_t$$

and since

$$\sum_{k=1}^{\infty}\left(c^*_{t+j,t-k}\right)\pi_{t-k}(1-\pi_{t-k+1})\cdots(1-\pi_{t-1}) \leq c^*_{t+j,t-1}$$

we therefore have the result that, if $\pi_t \in [1-\bar{\epsilon}_t, 1]$ and $\pi_{t+1} = 1 - \bar{\epsilon}_{t+1}$, demanded $q_t \geq 1$. On the other hand, if $\pi_{t+1} = 1$, then $c_{t+j,t} = 0$ for $j \geq 1$, and $q_t =$

$$\pi_t - \sum_{k=1}^{\infty} c^*_{t,t-k}\pi_{t-k}(1-\pi_{t-k+1})\cdots(1-\pi_{t-1})\pi_t \leq 1$$

with equality only if $\pi_t = 1$ as well.

Note that K defined in (2.12) is compact in ℓ^2; in particular, it is totally bounded, as, given any $\delta > 0$, $\exists N : \sum_{k=N}^{\infty} \bar{\epsilon}_{t+k}^2 < \delta^2/4$, and $\prod_{k<N}[0, \bar{\epsilon}_{t+k}]$ can be covered with a finite number of balls of radius $\delta\sqrt{3}/2$.

2.7.3 derivatives of q

(2.6) gives $q_t =$

$$\sum_{j=0}^{\infty}(1-\lambda)^{-j}c^*_{t+j,t}\pi_t + \sum_{k=1}^{\infty}\sum_{j=0}^{\infty}(1-\lambda)^{-j}\left(-c^*_{t+j,t-k}\right)\pi_{t-k}(1-\pi_{t-k+1})\cdots(1-\pi_{t-1})\pi_t$$

$\partial q_t/\partial \pi_{t+k} =$

$$1_{t+k=t}\sum_{j=0}^{\infty}(1-\lambda)^{-j}c^*_{t+j,t} + \pi_t\sum_{j=0}^{\infty}(1-\lambda)^{-j}\frac{\partial c^*_{t+j,t}}{\partial \pi_{t+k}}$$
$$-\sum_{l=1}^{\infty}\sum_{j=0}^{\infty}(1-\lambda)^{-j}\frac{\partial c^*_{t+j,t-l}}{\partial \pi_{t+k}}\pi_{t-l}(1-\pi_{t-l+1})\cdots(1-\pi_{t-1})\pi_t$$
$$+\sum_{l=1}^{\infty}\left(\frac{1_{t-l=t+k}}{\pi_{t-l}}+\frac{1_{t=t+k}}{\pi_t}\right)\sum_{j=0}^{\infty}(1-\lambda)^{-j}\left(-c^*_{t+j,t-l}\right)\pi_{t-l}(1-\pi_{t-l+1})\cdots(1-\pi_{t-1})\pi_t$$
$$+\sum_{l=1}^{\infty}1_{t-l+1\le t+k\le t-1}\sum_{j=0}^{\infty}(1-\lambda)^{-j}\left(c^*_{t+j,t-l}\right)\pi_{t-l}(1-\pi_{t-l+1})\cdots(1-\pi_{t-1})\pi_t/(1-\pi_{t+k})$$

$\partial q_t/\partial \pi_{t+k} =$

$$1_{t+k=t}\sum_{j=0}^{\infty}(1-\lambda)^{-j}c^*_{t+j,t} + \pi_t\sum_{j=0}^{\infty}(1-\lambda)^{-j}\frac{\partial c^*_{t+j,t}}{\partial \pi_{t+k}}$$
$$+1_{t=t+k}\sum_{l=1}^{\infty}\sum_{j=0}^{\infty}(1-\lambda)^{-j}\left(-c^*_{t+j,t-l}\right)\pi_{t-l}(1-\pi_{t-l+1})\cdots(1-\pi_{t-1})$$
$$-\sum_{l=\max\{1,1-k\}}^{\infty}\;\sum_{j=\max\{0,k\}}^{\infty}(1-\lambda)^{-j}\frac{\partial c^*_{t+j,t-l}}{\partial \pi_{t+k}}\pi_{t-l}(1-\pi_{t-l+1})\cdots(1-\pi_{t-1})\pi_t$$
$$+1_{t+k\le t-1}\sum_{j=0}^{\infty}(1-\lambda)^{-j}\left(-c^*_{t+j,t+k}\right)(1-\pi_{t+k+1})\cdots(1-\pi_{t-1})\pi_t$$
$$+1_{t+k\le t-1}\sum_{l=1-k}^{\infty}\sum_{j=0}^{\infty}(1-\lambda)^{-j}\left(c^*_{t+j,t-l}\right)\pi_{t-l}(1-\pi_{t-l+1})\cdots(1-\pi_{t-1})\pi_t/(1-\pi_{t+k})$$

For small $\bar{\epsilon}$, if $t = t+k$ this is dominated by the $c^*_{t,t}$ term; the difference between $\partial q_t/\partial \pi_t$ and $c^*_{t,t}$ can be bounded by positive powers of $1-\pi_t$. Similarly, if $t+k >$

t, the dominating term is $\pi_t(1-\lambda)^{-k}\frac{\partial c^*_{t+k,t}}{\partial \pi_{t+k}}$, and for $t+k<t$, we have primarily $-c^*_{t,t+k}(1-\pi_{t+k+1})\cdots(1-\pi_{t-1})\pi_t$. The diagonal elements of

$$\frac{\partial(q_1,\cdots,q_u)}{\partial(\pi_2,\cdots,\pi_{u+1})}$$

are $\pi_t(1-\lambda)^{-1}\frac{\partial c^*_{t+1,t}}{\partial \pi_{t+1}}$, which is bounded above and below by positive multiples of $-(1-\lambda)^{-1}\bar{\epsilon}_{t+1}^{1/\gamma-1}$ (as shown in the previous section); the terms above the diagonal are bounded (in absolute value) by $(1-\lambda)^{-1}\bar{\epsilon}_{t+1}^{1/\gamma}\bar{\epsilon}_{t+2}^{1/\gamma}\cdots\bar{\epsilon}_{k}^{1/\gamma-1}$. The terms immediately below the diagonal correspond to $\partial q_t/\partial\pi_t$, and are bounded above and below by positive constants; the terms below that are bounded in absolute value by a positive multiple of $\bar{\epsilon}_{t-k+1}^{1+1/\gamma}\cdots\bar{\epsilon}_{t-1}^{1+1/\gamma}\bar{\epsilon}_{t}^{1/\gamma}$. A crude sketch is provided in figure 2.2.

For $\gamma>1$, the terms on the diagonal get arbitrarily far from zero for small $\bar{\epsilon}$, (to $-\infty$), while all other terms stay bounded. Consider matrix M in which each row of

$$\frac{\partial(q_1,\cdots,q_u)}{\partial(\pi_2,\cdots,\pi_{u+1})}$$

has been divided by the diagonal term; matrix M has (by construction) 1's along the diagonal and, for sufficiently small starting $\bar{\epsilon}$,[8] the sum of the squares of the off-diagonal terms add up to less than 1. M is therefore positive definite;[9] therefore

$$-\frac{\partial(q_1,\cdots,q_u)}{\partial(\pi_2,\cdots,\pi_{u+1})} \text{ is positive definite for } \gamma>1$$

2.7.4 Synthesis

Let $S=\sum_i(q_i-1)^2$, which we regard as a function of π; note, from arguments in §2.7.2, that the terms in the sum are bounded by multiples of positive powers of $\bar{\epsilon}_i$ and thus S converges on the domain K. K is compact, and S is smooth, and S thus attains it minimum on K.

[8] Because the $\bar{\epsilon}_t$ converge exponentially to 0, a fixed $\bar{\epsilon}$ can be chosen for all u, viz. any size of matrix we consider.

[9] See §2.7.5.

$$\frac{\partial(q_1,\cdots,q_u)}{\partial(\pi_2,\cdots,\pi_{u+1})} = \begin{pmatrix} -\bar{\epsilon}^{1/\gamma-1} & -\bar{\epsilon}^{2/\gamma-1} & -\bar{\epsilon}^{3/\gamma-1} & \cdots \\ \bar{\epsilon}^{0} & -\bar{\epsilon}^{1/\gamma-1} & -\bar{\epsilon}^{2/\gamma-1} & \cdots \\ -\bar{\epsilon}^{1/\gamma} & \bar{\epsilon}^{0} & -\bar{\epsilon}^{1/\gamma-1} & \cdots \\ -\bar{\epsilon}^{1+2/\gamma} & -\bar{\epsilon}^{1/\gamma} & \bar{\epsilon}^{0} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Figure 2.2: A heuristic look at the Jacobian. This is a charicature, but provides a good starting point for intuition. The diagonal term is bounded away from zero; $\bar{\epsilon}_t$ decreases with t, so the power counting should be regarded as a rough guide, even as a bound.

The gradient is $\nabla S = 2\sum_i (q_i - 1)\nabla(q_i)$; because the (negative) Jacobian $\partial q/\partial\pi$ is positive-definite, $\sum_j (q_j - 1)\nabla_j S < 0$ unless $(q_i - 1) = 0$ for all i. Similarly, anywhere on the boundary of K, if $\hat{n}$ is any vector normal to a binding constraint and pointed into K, $\hat{n}\cdot\nabla S < 0$; thus the minimum of S must occur in the interior of K, where $q_t = 1$ for all t.

2.7.5 Positive Definiteness of Almost Diagonal Matrices

Suppose $v \in \mathbb{R}^n$ and i an integer between 1 and n inclusive, and that $\exists j \neq i : v_j \neq 0$. Then

$$\min_{m\in\mathbb{R}^n} m^T m \text{ s.t. } m_i = 1, m^T v = 0$$

is given by a unique m, for which $m_j = -v_i v_j / C$ for $j \neq i$, where $C = \sum_{j\neq i} v_j^2$; the value of the minimum is $m^T m =$

$$1 + \frac{v_i^2}{v^T v - v_i^2}$$

Now consider $v \in \mathbb{R}^n$, $v \neq 0$, and we wish to solve

$$\min_{M_{n\times n}} \sum_{i=1}^{n}\sum_{j=1}^{n} M_{ij}^2 \text{ s.t. } Mv = 0_{n\times 1}, M_{ii} = 1\forall i$$

If only one element of v is non-zero, the constraints cannot be satisfied; otherwise, the problem separates along rows of M, and we can apply the previous result to each row of M, giving

$$\sum_i \left(1 + \frac{v_i^2}{v^T v - v_i^2}\right) > n + \sum_i \frac{v_i^2}{v^T v} = n + 1.$$

The result that interests us is the contrapositive: given $M_{n\times n}$ such that $M_{ii} = 1 \forall i$ and $\sum_{ij:i\neq j} M_{ij}^2 \leq 1$, $\nexists v \in \mathbb{R}^n, v \neq 0 : Mv = 0_{n\times 1}$. The immediate corollary is

Proposition 3. *Given $M_{n\times n}$ such that $M_{ii} = 1 \forall i$ and $\sum_{ij:i\neq j} M_{ij}^2 \leq 1$, M is non-singular.*

Note that if a matrix satisfies the conditions of the proposition, then so does its symmetric part; furthermore, the set $\mathcal{S}$ of symmetric matrices satisfying those conditions is connected in the usual Euclidean topology in $\mathbb{R}^{n^2}$. Finally, $1_n \in \mathcal{S}$, thus the eigenvalues of any element of $\mathcal{S}$ are positive, and thus

Proposition 4. *Given $M_{n\times n}$ such that $M_{ii} = 1 \forall i$ and $\sum_{ij:i\neq j} M_{ij}^2 \leq 1$, M is positive definite.*

Chapter 3

Risk Dominance beyond Equilibrium

3.1 Introduction

In “A Discourse on the Origin of Inequality”, Rousseau asserts that primitive man is incapable of sustained cooperation (Rousseau, 1950, p 238):

> If a deer was to be taken, every one saw that, in order to succeed, he must abide faithfully by his post: but if a hare happened to come within the reach of any one of them, it is not to be doubted that he pursued it without scruple, and, having seized his prey, cared very little, if by so doing he caused his companions to miss theirs.

Simplifying to two players and translating into the language of modern game theory, each player would be best off by cooperating to catch and halve a stag, but would be worst off *trying* to catch a stag while the other player chased a hare, e.g. with payoffs as in figure 3.1. If either player assesses the probability that the other player will chase the hare at more than 20%, then that player's own better strategy is to chase the hare. (Stag, Stag) and (Hare, Hare) are both Nash equilibria, and (Stag, Stag) is Pareto dominant, but each player will only choose it if they are pretty confident that it will work out — which requires, in particular, that each player is pretty confident that the other player is pretty confident, too.

Harsanyi and Selten discuss a similar situation in (Harsanyi and Selten, 1988, §3.7). “Suppose that the players are in a state of mind where they think that either U or V must be the solution of the game. What is the risk of deciding one way or the other?”

	Stag	Hare
Stag	5,5	-10,2
Hare	2,-10	2,2

Figure 3.1: The Stag Hunt

In any finite game with strict Nash equilibria, there will be, for each player and for each such equilibrium, a smallest probability p_{ij} such that if player i believes the other player(s) will choose equilibrium j with probability at least p_{ij}, then doing so will be a best response for player i as well. (Harsanyi and Selten, 1988, §3.7) suggest verbally that if either p_{ij} is very small for one of the j, that equilibrium is particularly risky. For 2×2 games with multiple strict equilibria, there is generically one equilibrium for which $\sum_i p_{ij}$ is less than 1 and another for which it is greater. Harsanyi and Selten (1988) identify the former as "risk-dominant".

Morris, Rob, and Shin (1995) would later define a notion of "p-dominance" in a way related to the above construction, and would identify equilibria (in larger finite games) as "risk-dominant" if they obey a similar relationship; in these larger games, such an equilibrium frequently fails to exist. Other authors, in other contexts, have characterized risk-dominance in other ways. At times, the term is used without any clear definition at all. This essay represents a new attempt to make the notion of "risk-dominance" more precise, and to highlight its relationship to other phenomena. It will be shown that the idea of risk-dominance can apply to non-equilibrium outcomes, and may suggest them as more likely or supportable than equilibrium outcomes.

Equilibrium, however, is where it is usually applied, and we will start with the verbal definition

Definition 1. *An equilibrium is* ***risk dominant*** *if coordination on it is easier than on other equilibria for agents who are worried about failing to coordinate.*

In section 3.5, we will provide a more mathematically precise definition based on

	1	2
1	a^1_{11}, a^2_{11}	a^1_{12}, a^2_{21}
2	a^1_{21}, a^2_{12}	a^1_{22}, a^2_{22}

Figure 3.2: A general 2×2 game.

this verbal definition, but first we will review some additional literature extending the definition of risk-dominance given by Harsanyi and Selten (1988), providing other characeterizations of their notions, or using the term without a clear definition. We will then introduce the Maximum Entropy idea of Jaynes (1957), which we illustrate while providing a new characterization of "logistic quantal response" (McKelvey and Palfrey (1995); cf. Blume (1993)), giving a new look at experimental data from Nagel and Tang (1998); some other relationships between LQR equilibrium and risk-dominance in 2×2 games will be noted. We then apply Jaynes's idea to finite two-player games and demonstrate its relationship to other notions of risk-dominance, and also how it favors non-equilibrium profiles under certain circumstances (as in the results of Nagel and Tang (1998)).

3.2 Risk-dominance Review

3.2.1 Usage

There are a few contexts in which the term "risk-dominance" is used consistently, and a few interesting characterizations of risk-dominance in some of those contexts. An aspiring definition of risk-dominance in general should be consistent with these.

Figure 3.2 gives a general 2×2 game — note that the notation is such that $a^1_{ij} = a^2_{ij}$ if it is symmetric — and those situations in which we are interested are those in which there are two strict Nash equilibria; without loss of generality, $a^k_{11} > a^k_{21}$ and $a^k_{22} > a^k_{12}$ for $k \in \{1, 2\}$. There will then be a mixed equilibrium in which action 1 is played by player i with probability σ_i given by[1]

$$\sigma_i = \frac{a_{22}^{-i} - a_{12}^{-i}}{a_{11}^{-i} - a_{12}^{-i} + a_{22}^{-i} - a_{21}^{-i}},$$

which is to say that if player i plays action 1 more often than action 1 will be player $-i$'s best response, and vice versa. Three useful and algebraically equivalent characterizations of Harsanyi and Selten (1988)'s definition of risk-dominance in this setting are

$$\begin{matrix} (a_{11}^1 - a_{21}^1)(a_{11}^2 - a_{21}^2) \\ (1-\sigma_1) + (1-\sigma_2) \\ (1-\sigma_1)(1-\sigma_2) \end{matrix} \left\{ \begin{matrix} < \\ \\ > \end{matrix} \right\} \begin{matrix} (a_{22}^1 - a_{12}^1)(a_{22}^2 - a_{12}^2) \\ \sigma_1 + \sigma_2 \\ \sigma_1\sigma_2 \end{matrix} \tag{3.1}$$

where (1,1) is risk-dominant if the quantity on the left is higher than the corresponding quantity on the right, while (2,2) is risk-dominant if the quantity on the right is higher. The second characterization is the notion of p-dominance from Morris, Rob, and Shin (1995) as applied to this class of games.

Another special class of games in which the term "risk-dominance" is used consistently is the "minimum games" of Van Huyck, Battalio, and Beil (1990) and subsequent papers. These games have a finite number n of players each of which has a finite number of actions and are symmetric among all players; actions are conveniently denoted by positive integers, e.g. $A_i = \{1, 2, 3, 4, 5, 6, 7\}$ with payoffs of $u_i = \alpha \min_j\{a_j\} - \beta a_i$ for $\beta > 0$ and $\alpha n > \beta$, such that each player's best-response is the minimum of other players' actions and the outcome in which each player chooses the action denoted with the greatest integer is Pareto-dominant. Any outcome in which all agents play the same action is a Nash equilibrium, and when αn is much bigger than β, that in which players choose the action with the least integer is "risk-dominant" — if players are not terribly confident in the others' choices, each player is likely to choose that action.

Some proposed extensions of the notion of "risk-dominance" in two-player games with more than two actions focus only on two actions for each player at a time. Most

[1] $-i$ is, in this sort of context, $3 - i$ arithmetically.

		Player 2		
		S	H	C
Player 1	S	5,5	0,3	0,0
	H	3,0	3,3	0,0
	C	0,0	0,0	4,4

Figure 3.3: A symmetric 3×3 game with actions "Stag", "Hare", and "Coordinate", with a strict equilibrium whenever the two players play the same action. Restricting the game to only two of the three, Hare risk-dominates Stag, Stag risk-dominates Coordinate, and Coordinate risk-dominates Hare.

straightforwardly, one pure-strategy equilibrium might be said to "risk-dominate" another if that relationship holds in the sense of Harsanyi and Selten (1988) in the game in which each player is restricted to only the two actions that are played in the equilibria under consideration. It should be noted that this relationship is not transitive; see figure 3.3.[2]

3.2.2 Characterizations

(Harsanyi and Selten, 1988, §3.9) shows that there is only one mapping from the space of 2×2 games with two strict equilibria to the space of orderings on those equilibria such that

- any two games with the same best-response correspondences result in the same ordering
- any games that differ only by relabeling actions carry that relabeling over to the ordering
- if $A \succeq B$ for game G, then for a game H that differs from G only by giving a higher payoff to one player for A than G does, $A \succ B$.

[2]Peski (2010) looks at pairwise relationships, declaring that no profile is risk-dominant in a game such as this.

Any 2×2 game with multiple strict equilibria has the same best-response correspondences as a potential game (Monderer and Shapley (1996)), and in any such potential game the global maximum of potential corresponds to the risk-dominant equilibrium.

Carlsson and van Damme (1993) prove an interesting result about games of almost-complete information. A 2×2 game of complete information can be identified with a point in $\mathbb{R}^8$; an interesting class of games of incomplete information stipulates common knowledge of the action spaces and a common prior on $\mathbb{R}^8$ from which the payoffs of the game are drawn. If

- players have a smooth common prior
- with support on a smooth, connected manifold that includes the possibility that action H is dominant for either player
- and players have sufficiently precise but independent signals of the actual payoffs,
- and H is risk-dominant on the entire manifold,[3]

then in any Bayes-Nash equilibrium of the full game, Hare is always played.

Ui (2001); Frankel, Morris, and Pauzner (2003) note that similar properties hold for the global maximum of potential in "potential games" and also for p-dominant equilibria with $p < 1$.

Some simple models of bounded rationality suggest that risk-dominant equilibria of 2×2 games should be played. Stahl and Wilson (1994, 1995); Nagel (1995) introduce "level-k reasoning"; a level-0 type makes a naive choice, typically uniform mixing, and a level-k agent for $k > 0$ best-responds to level-$k-1$ play. If level-1 play only includes one action, and it is a strict Nash equilibrium for both players to engage in level-1 play, as it will be in symmetric 2×2 games, then level-k players for any $k > 0$ will all play that equilibrium; in such 2×2 games with multiple strict equilibria, that equilibrium is

[3] For purposes of exposition, I've narrowed their result somewhat.

the risk-dominant one. Camerer, Ho, and Chong (2004) proposed a related "cognitive hierarchy" model in which the level-k agents (for $k > 0$) are best-responding to mixtures of level-m players for $m < k$; under the same circumstances as for the simpler level-k model, this, too, will result in play of the risk-dominant equilibrium.

In learning models, players are often most likely to end up in or near risk-dominant equilibria. Blume (1993) puts players on a network and has them play games with their neighbors; starting from any action profile and repeatedly selecting a player at random to update to a best-response with a vanishing amount of noise, players converge with arbitrarily high probability to the maximum of any potential game and to the risk-dominant equilibrium of a 2×2 game with multiple strict equilibria.[4] Young (1993) produces a similar result based on agents who play the same game repeatedly and simultaneously, but with a finite memory for what others have played. He calls outcomes that are reached with positive probability as the probability of mistakes goes to 0 "stochastically stable"; in 2×2 games in which one equilibrium risk-dominates another, the risk-dominant equilibrium is uniquely stochastically stable in his sense. Note in figure 3.4 that the addition of a third action for each player can affect which outcome is stochastically stable, even if the added actions are not played in any stochastically stable equlibrium of the augmented game. Binmore, Samuelson, and Gale (1995) gives a model in which agents learn, but with a small but positive amount of forgetting (or experimentation or replacement with ignorant new players) that keeps the stable outcome far from Nash equilibrium.

The term "stability set" is introduced in Harsanyi and Selten (1988) to refer to the inverse of the best-response correspondence; for a player with a possible (pure) strategy j, the stability set of j is the set of the player's first-order beliefs to which j would be a best response. Harsanyi (1995) constructs a measure on first-order belief spaces (over

---

[4]See Alós-Ferrer and Netzer (2010), though, on the potential sensitivity of this to the updating process.

		Player 2		
		L	C	R
Player 1	T	5,60	0,0	0,0
	M	0,0	7,40	0,0
	D	0,0	0,0	100,1

Figure 3.4: (Young, 1998, p. 104). (T,L) (strictly) risk-dominates both other equilibria in the respective 2×2 games, but (M,C) is (uniquely) "stochastically stable" in Young's sense; in particular, it is easier to slip from T,L to D,R to M,C than back.

finite state spaces) and calls the measure of a strategy's stability set the "structural incentive" to play that strategy. His measure is equivalent to the Lebesgue measure when there are only two possible states, but in general is different. He suggests that, after some initial reduction of a game, that players should coordinate on the Nash equilibrium that maximizes the product of these structural incentives.

3.3 Maximum Entropy Principle

> The process of induction is the process of assuming the simplest law that can be made to harmonize with our experience. —Ludwig Wittgenstein (1922), quoted in Aragones, Gilboa, Postlewaite, and Schmeidler (2005)

Jaynes (1957) considers the existence of a system known to have n states, and asks whether there is a single "correct" probability distribution to place on the system if one is also given the expected value of some real-valued function of the state.

> Just as in applied statistics the crux of a problem is often the devising of some method fo sampling that avoids bias, our problem is that of finding a probability assignment which avoids bias, while agreeing with whatever information is given. The great advance provided by information theory lies in the discovery that there is a unique, unambiguous criterion for the 'amount of uncertainty' represented by a discrete probability distribution, which agrees with our intuitive notions that a broad distribution represents more uncertainty than does a sharply peaked one.

In particular, Shannon showed that any quantity that "is positive, which increases with increasing uncertatinty, and is additive for independent sources of uncertainty" is his measure of "information". More precisely: if we want a real-valued function S that can be defined on the union of sets of probability distributions on finite spaces to be continuous, to give an "information" in a uniform prior that is increasing in n, and for equalities of the form $S(p_1, p_2, p_3, p_4) = S(p_1 + p_2, p_3 + p_4) + (p_1 + p_2)S(p_1, p_2) + (p_3 + p_4)S(p_3, p_4)$ to hold, then S has to be a multiple of

$$S = -\sum_i p_i \log(p_i)^5 \tag{3.2}$$

subject to the linear constraint. Jaynes asserts that imputing the probability that maximizes entropy subject to known constraints is the least presumptuous; most importantly, perhaps, it gives positive support to any possibility that is not ruled out by the constraints. (Jaynes, 2003, §11.1–3) offers a more verbose presentation of these ideas. (Jaynes, 2003, §11.4) relates another observation (of which I'll make independent use): suppose you generated prior distributions by taking N tokens, distributing each uniformly over the n possibilities, and calling the resulting distribution a prior; if you throw away those priors you get that violate the constraints, the modal and mean distribution, as N gets larger, both approach the maximum entropy distribution. (The Shannon-McMillan-Breiman theorem in fact assures that, because the entropy maximum is unique and entropy and the constraints are smooth, for large N almost all realizations, conditional on the constraint, are arbitrarily close to the maximum entropy distribution.)

Shore and Johnson (1980) axiomatize this differently, and somewhat more generally. Let an "information operator" take a prior (with full support) and a set of linear constraints and yield a posterior distribution; if it can be expressed as the operation of minimizing a functional $F(q, p)$ of the prior and the posterior over the set of posteriors that satisfy the constraints, and if the information operator is well-defined and is invariant under linear transformations of the system and under aggregation and separation of independent subsystems, then it is uniquely determined as the minimization of the

[5] Here, and throughout, any term in such a sum for which $p_i = 0$ is taken to be equal to 0, which is the limit as x approaches 0 from above of $x \log(x)$.

cross-entropy,

$$\sum_i q_i \log\left(\frac{q_i}{p_i}\right) \tag{3.3}$$

If the "information operator" takes only a set of linear constraints *without* a prior, and obeys the analogous axioms, it gives as the posterior those beliefs that maximize entropy subject to the constraint. (This is equivalent to minimize the relative entropy relative to a uniform prior.)

3.4 Logistic Quantal Response

3.4.1 Logistic Quantal Response

McKelvey and Palfrey (1995) define "quantal response equilibrium" as "a statistical version of Nash equilibrium where each player's utility for each action is subject to random error"; Hofbauer and Sandholm (2002) later note that any finite game in which each player's payoffs are some commonly known value plus a somewhat regular "random error", and in which each player knows its own payoffs, is equivalent in a straightforward sense to a deterministic game whose action space is that of the mixed-extension of the complete information game, with payoffs modified by a strictly concave function on the action space. Of special interest is the logistic quantal response equilibrium, in which a player chooses $\sigma \in \Delta(A_i)$ to maximize

$$u^\gamma(\sigma) = \sum_j \sigma_j u_j - \gamma \sum_j \sigma_j \log(\sigma_j)$$

given an expected payoff of u_j for choosing action j. This corresponds to an agent's payoffs' having independently and identically distributed log-Weibull noise; the first-order condition is equivalent to

$$\sigma_j = e^{\frac{u_j - \bar{u}}{\gamma}} \tag{3.4}$$

where $\bar{u}$ is a Lagrange multiplier on the constraint that $\sum_j \sigma_j = 1$. Note that as $\gamma \to 0$ (from above) the payoff function returns to that of the mixed-extension of the finite game, and the best-response approaches a best-response in that game; as $\gamma \to \infty$, the player's best-response approaches a uniform mixture. We will let $\gamma \in [0, \infty]$ without confusion. McKelvey and Palfrey (1995) demonstrate that a number of human subject experiments with suitable games could be well modeled by LQRE, with statistical estimation of γ often excluding both endpoints.

(3.4) (or something like it) is most often how LQRE is defined, but the formulation as a solution to a maximization problem will be more useful for us; for $\gamma \in (0, \infty)$ we can solve the maximization problem already given or, equivalently, for $\gamma = 1/\lambda$,[6] the agent can maximize

$$S + \lambda u$$

with S as in (3.2). Logistic quantal response is equivalent to maximizing S for a fixed value of u, with λ as a Lagrangian multiplier. In particular,

Proposition 5. *Given a finite decision problem, with*

- *known expected payoffs to each possible action, and*
- *a given threshold "satisficing" payoff less than the highest possible payoff but greater than the average payoff over all possible actions, and*
- *given only the information that the agent seeks to attain at least the given expected payoff level,*

then, as per Jaynes (1957), "the least biased estimate possible on the given information" is a belief that places Logistic Quantal Response probabilities on the player's choice of actions.

Alternatively, in light of the previously mentioned results of (Jaynes, 2003, §11.4) and the Shannon-McMillan-Breiman theorem, we can view LQR as the result of a decision process in which an agent experiments with a high-entropy distribution, then

[6] Going forward, $\lambda \in [0, \infty]$, $\lambda = 0 \Leftrightarrow \gamma = \infty$, $\lambda = \infty \Leftrightarrow \gamma = 0$.

	2	4	6	8	10	12	14
1	4,1	4, 1	4, 1	4, 1	4, 1	4, 1	4, 1
3	2,5	8, 2	8, 2	8, 2	8, 2	8, 2	8, 2
5	2,5	3,11	16, 4	16, 4	16, 4	16, 4	16, 4
7	2,5	3,11	6,22	32, 8	32, 8	32, 8	32, 8
9	2,5	3,11	6,22	11,45	64,16	64, 16	64,16
11	2,5	3,11	6,22	11,45	22,90	128, 32	128,32
13	2,5	3,11	6,22	11,45	22,90	44,180	256,64

Figure 3.5: The centipede game in normal form used in Nagel and Tang (1998).

periodically looks back at the most recent N periods and locks in the observed frequency of play of each action if the average payoff meets the required level.

Because, for a finite non-degenerate decision problem, S is strictly concave on $\Delta(A)$ and u is linear on $\Delta(A)$, $\lambda \leftrightarrow \gamma \leftrightarrow \langle u \rangle$ are continuous[7] and invertible on their ranges.[8] The set of logistic quantal responses and solutions to the satisficing problems are the same, though for certain applications (e.g. deciding whether two players in a game should use the same λ) the parameterization might affect the natural implications of the concept. In §3.5 we will encounter a similar set of parameters that pose a bit more of a conundrum.

3.4.2 Centipede Game

Let's look at an illustration of these ideas, using data from Nagel and Tang (1998). They study the game in figure 3.5, which is the reduced normal form of a game in which the odd player (with actions 1, 3, etc.) may choose to "take" (ending the game immediately, with payoffs of 4 to the odd player and 1 to the even player) or to "pass" (after which the even player has a similar choice). If the odd player passes seven times, the even player is required to take in the next stage. The more times players pass, the

---

[7]Using the Alexandroff (one-point) compactification topology for λ and γ.

[8]The range of $\langle u \rangle$ is the closed interval from the mean value of u to the maximum value of u over available actions.

action	preponderance	own value	std dev
2	0.022	5.0	0.3
4	0.027	10.9	0.9
6	0.140	20.8	4.5
8	0.313	30.4	18.4
10	0.312	38.8	38.1
12	0.143	42.3	61.2
14	0.043	23.5	19.9

Figure 3.6: Observed context and behavior of the even-type players in session 1 in Nagel and Tang (1998). For example, the behavior of odd-type players in that session was such that an even-type player who plays 12 gets an expected payoff of 42.3 with a standard deviation of 61.2.

larger the total payoff get, but a player who takes gets a substantially higher payoff than the player who passes, such that at any decision node at which the player deciding knows that the other player would take at the next stage, it is that player's best response to immediately take instead.

By backward induction, the only subgame perfect equilibrium is that in which each player chooses to take at every decision node. Relatedly, in normal form, in any correlated equilibrium, the odd player chooses 1 with certainty and the even player plays 2 with a probability of at least 2/3. However, high-numbered actions tend to be best-responses to "most"[9] mixed strategies, as well as to high actions; in a cognitive hierarchy model, for example, level 1 players play the highest undominated action, with level 2 players playing their second highest possible action.

3.4.3 Derivatives of LQR

Let's work from

$$\sigma_{ij} = \frac{1}{m} e^{(u_{ij} - \bar{u}_i)\lambda_i}$$

[9] There will be some discussion of measures in section 3.5.

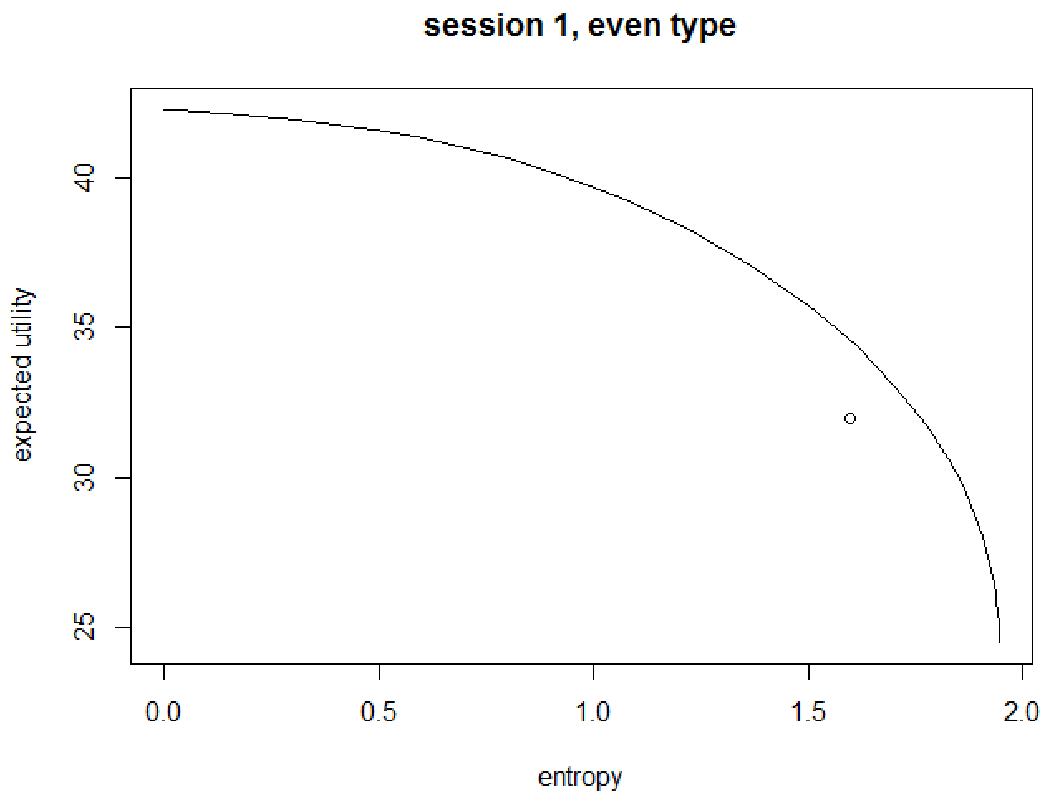


Figure 3.7: The entropy and expected payoff for the distribution played by even players in figure 3.6, plotted against the frontier. (This frontier depends on odd player behavior.)

		payoff	entropy	max pay	max ent
1	1	30.34	1.535	31.79	1.665
1	2	31.99	1.597	34.61	1.764
2	1	60.12	1.405	62.89	1.524
2	2	48.60	1.495	51.92	1.649
3	1	26.08	1.505	26.54	1.563
3	2	23.27	1.322	23.34	1.339
4	1	45.68	1.191	46.92	1.335
4	2	39.85	1.381	40.88	1.468
5	1	22.96	1.371	23.57	1.503
5	2	20.79	1.430	20.84	1.442

Figure 3.8: Each maximum is conditional on the other variable. For LQR the maxima are achieved. Uniform mixing would give an entropy of 1.946.

with special interest in the case where $\lambda_i \approx 0$, the "noisy limit" in which each player is expected to randomize fairly liberally.[10] For finite games the LQRE is unique in the neighborhood of $\lambda = 0$; let's consider the LQRE as a function of λ. We'll find that the derivatives of the LQR with respect to λ represent orders of beliefs, such that, insofar as risk-dominance is primarily a second-order idea — I worry that you worry that I'm going to chase the hare — it corresponds to the second derivatives of LQR.[11]

Note that the σ_{ij} are all smooth, even for $\lambda < 0$; in particular, all derivatives are bounded as the $\lambda \to 0$. For players i and k_1, k_2, $\ldots k_d$ (for $d \geq 1$),[12]

$$\left(\prod_h \frac{\partial}{\partial \lambda_{k_h}}\right) \log \sigma_{ij} = \sum_{g:k_g=i} \left(\prod_{h \neq g} \frac{\partial}{\partial \lambda_{k_h}}\right)(u_{ij} - \bar{u}_i) + \lambda_i \left(\prod_h \frac{\partial}{\partial \lambda_{k_h}}\right)(u_{ij} - \bar{u}_i)$$

and for $\lambda_i = 0$,

$$\left(\prod_h \frac{\partial}{\partial \lambda_{k_h}}\right) \log \sigma_{ij} = \sum_{g:k_g=i} \left(\prod_{h \neq g} \frac{\partial}{\partial \lambda_{k_h}}\right)(u_{ij} - \bar{u}_i)$$

Each term in the sum contains one fewer derivative than we started with. $\frac{\partial}{\partial \lambda_i} \log \sigma_{ij}$ is a positive affine transformation of the payoffs the player would get against a player (or players) who is uniformly randomizing, i.e. against level-0 play.

$u_{ij} - \bar{u}_i$ will depend on others' strategies and therefore, in this context, on their λ. In a two-player game with payoff matrix M_{jl} such that, with other's (mixed) strategy of σ', the expected payoff to action j is $u_j = \sum_l M_{jl}\sigma'_l$. For example, a second derivative of $\log \sigma_{ij}$ — with respect to λ_i and λ_k — is a positive affine transformation of

$$\frac{\partial u_j}{\partial \lambda_k} = \sum_l M_{jl} \frac{\partial}{\partial \lambda_k} \sigma'_l;$$

[10] I have redefined $\bar{u}$, which varies with λ_i, subtracting off $\log(n)/\lambda_i$ from it so that it does not go to infinity as $\lambda_i \to 0$.

[11] To some extent "game theory" is second-order; zero and first-order are decision theory.

[12] It is hoped that the notation for "products" of differential operators is clear.

if the other player has a payoff matrix of N, this is a positive affine transform of the row sum of MN, favoring actions that get high payoffs against actions by the other player that get that other player high payoffs. Thus different levels of derivatives of $\log \sigma_{ij}$ at $\lambda = 0$ generate a "soft" version of the level-k model.

For example, consider a general 2×2 game in which the payoffs to the different actions are

$$u_1 = a^1_{11}\eta_1 + a^1_{12}\eta_2 \quad u_2 = a^1_{21}\eta_1 + a^1_{22}\eta_2$$
$$v_1 = a^2_{11}\sigma_1 + a^2_{12}\sigma_2 \quad v_2 = a^2_{21}\sigma_1 + a^2_{22}\sigma_2$$

Defining

$$m_1 = a^1_{21} - a^1_{11} - a^1_{22} + a^1_{12} \tag{3.5a}$$

$$b_1 = a^1_{22} - a^1_{12} \tag{3.5b}$$

$$m_2 = a^2_{21} - a^2_{11} - a^2_{22} + a^2_{12} \tag{3.5c}$$

$$b_2 = a^2_{22} - a^2_{12} \tag{3.5d}$$

the (non-zero) first two derivatives (at $\lambda_1 = \lambda_2 = 0$) are

$$\frac{\partial \sigma_1}{\partial \lambda_1} = (-1/4)(m_1/2 + b_1)$$
$$\frac{\partial \eta_1}{\partial \lambda_2} = (-1/4)(m_2/2 + b_2)$$
$$\frac{\partial}{\partial \lambda_1}\frac{\partial \sigma_1}{\partial \lambda_2} = (-1/4)m_1\frac{\partial \eta_1}{\partial \lambda_2}$$
$$\frac{\partial}{\partial \lambda_2}\frac{\partial \eta_1}{\partial \lambda_1} = (-1/4)m_2\frac{\partial \sigma_1}{\partial \lambda_1}$$

Note that (1,1) and (2,2) are strict Nash equilibria if and only if $b_1 > 0$, $b_2 > 0$, $m_1 + b_1 < 0$, and $m_2 + b_2 < 0$, in which case (1,1) is risk-dominant if

$$(m_1 + b_1)(m_2 + b_2) > b_1 b_2$$

and (2,2) is risk-dominant if the inequality is reversed. The second derivative of $\sigma_1 + \eta_1$ can be found to be

$$\frac{\partial}{\partial \lambda_1}\frac{\partial(\sigma_1 + \eta_1)}{\partial \lambda_2} = (1/16)\left[(m_1 + b_1)(m_2 + b_2) - b_1 b_2\right];$$

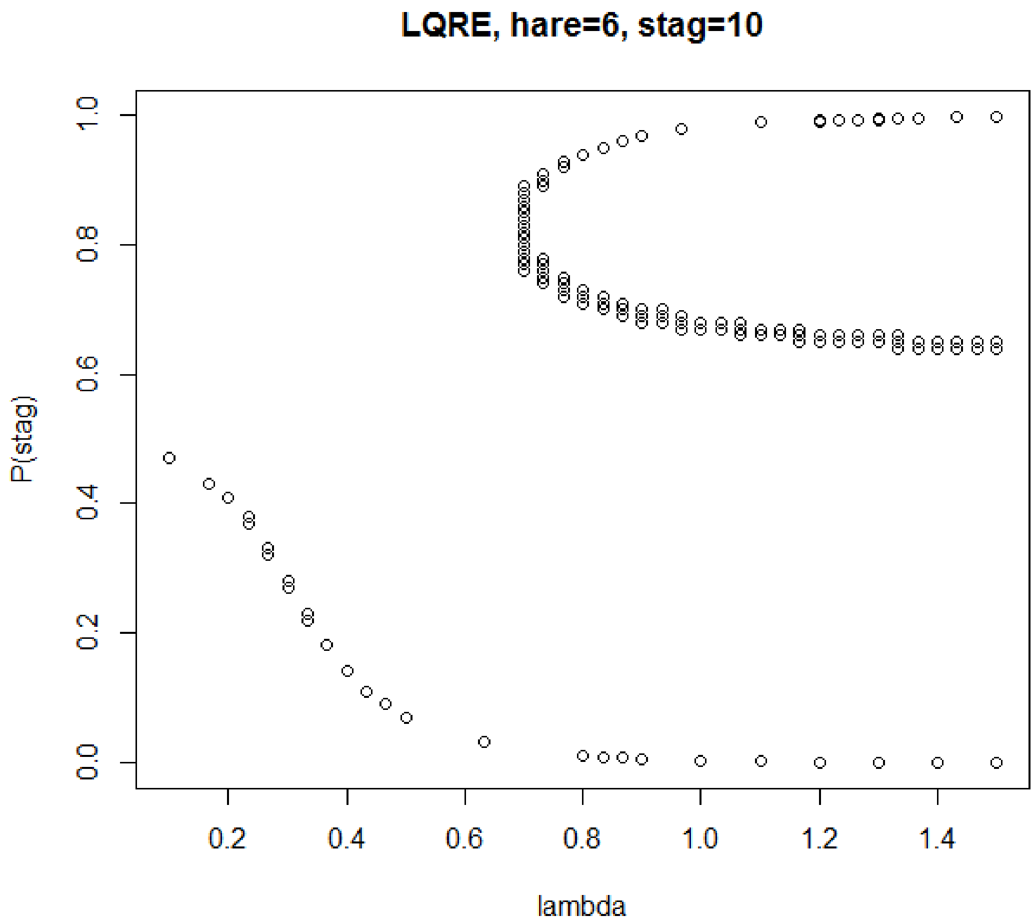


Figure 3.9: For a 2×2 game, the graph of logistic quantal response equilibria against λ will include a connected component that includes $(0, (1/2, 1/2))$ and at least one point with $\lambda = \infty$ and a Nash equilibrium. McKelvey and Palfrey (1995) show that if one equilibrium risk-dominates another, only the risk-dominant equilibrium is in this connected component.

if one of the equilibria risk-dominates the other, then the players' total *second-order* tendency will be to favor that equilibrium.

3.5 Relaxed equilibrium

3.5.1 Relaxed NE

Consider a finite two-player game. A profile of mixed strategies is a Nash equilibrium if

1. player 1 is best-responding to player 1's beliefs
2. player 2 is best-responding to player 2's beliefs
3. the players have the correct first-order beliefs.

The last item in the list is the coordination problem — that the players agree, with certainty, as to what strategy profile they intend to be playing.

Suppose we are given a mixed strategy $\bar{\sigma}$ by player 1, and we want to examine under what conditions player 1 might play according to that rule; in particular, we want to look at the first-order and second-order beliefs that player 1 might hold such that

1. player 1 is almost best-responding to player 1's first-order beliefs, and
2. player 1 believes that player 2 is almost best-responding to player 2's first-order beliefs, and
3. player 2 has some idea that player 1 is playing $\bar{\sigma}$.

The notion of risk-dominance will then be connected to how hard it is to construct such beliefs.

How well player 2 knows $\bar{\sigma}$ will be measured in terms of the "relative entropy", sometimes called "Kullback-Leibler divergence"[13], as defined in equation (3.3);

$$\sum_i \bar{\sigma}_i \log\left(\frac{\bar{\sigma}_i}{\sigma_i}\right) \tag{3.6}$$

is 0 if $\sigma = \bar{\sigma}$ and is positive otherwise; in information theoretic terms, it represents how much information would be required to correct player 2 as to the correct first-order beliefs if player 2 started with a prior of σ.

As suggested by §3.3, "how hard it is to construct such beliefs" will be measured in terms of the amount of entropy those beliefs can have. The rationality constraints (that each player be sufficiently-well responding) look like the constraints that we relaxed to get LQR, but will be somewhat different; instead of formulating beliefs conditional on well-responding to a fixed decision problem, we will here infer the beliefs of a player that would lead to certain behavior.

Let $P_j(\sigma)$ denote player 1's second-order beliefs; this is the probability density that player 1 places on player 2's beliefs conditional on player 2's playing action j. (P_j is thus a density on $\Delta(A_1)$.) Taking these, for the moment, as given, the total entropy of the joint distribution $P(j\&\sigma) = P(j)P_j(\sigma)$ is the entropy of the discrete distribution $P(j)$ plus $\sum_j P(j)S_j$, where S_j is the entropy of $P_j(\sigma)$. Maximizing the entropy of the joint distribution requires choosing P_i to maximize $\sum_j P_j(-\log(P_j) + S_j)$ subject to the normalization constraint; including a set of linear constraints as well, the first-order condition for

$$\sum_j P(j)(-\log(P(j)) + S_j) + \nu P(j) - \sum_k \lambda_k P(j) s_{kj}$$

gives

$$\log(P(j)) = S_j + 1 - \nu - \sum_k \lambda_k s_{kj}$$

i.e.

$$P(j) = Ce^{S_j - \sum_k \lambda_k s_{kj}}$$

[13]Or, sometimes, "Kullback-Leibler distance", in spite of the fact that it isn't symmetric and is therefore not a metric.

where C is a normalization constant.

The set of constraints in which we're interested are constraints of the form

$$\sum_{ij} \bar{\sigma}_i (M_1)_{ij} P(j) - \sum_j P(j)(M_1)_{kj} \geq -\bar{\epsilon},$$

i.e. from the fact that player 1 is playing $\bar{\sigma}$ we infer that there is no strategy k that player 1 expects to be more than $\bar{\epsilon}$ better. We thus have

$$s_{kj} = \sum_i \bar{\sigma}_i (M_1)_{ij} - (M_1)_{kj}$$

with $\lambda_k > 0$ only if k is a best response to mixture P and the constraint binds. The achieved entropy is $S =$

$$\sum_j P(j)(-\log(P(j)) + S_j) = -\log(C) + \sum_j C e^{S_j - \sum_k \lambda_k s_{kj}} \left(\sum_k \lambda_k s_{kj} \right)$$

which can be written in a form similar to that of equation 3.11,

$$C^{-1}(\{\lambda_i\}) = \sum_j e^{S_j - \sum_k \lambda_k s_{kj}}$$

$$S = \log(C^{-1}) - C \sum_k \lambda_k \frac{\partial}{\partial \lambda_k} C^{-1} = \log(C^{-1}) - \sum_k \lambda_k \frac{\partial}{\partial \lambda_k} \log(C^{-1}). \tag{3.7}$$

For $\sum_k \lambda_k > 0$, we can write

$$\sum_k \lambda_k s_{kj} = \left(\sum_k \lambda_k \right) \sum_i \left(\bar{\sigma}_i - \frac{\lambda_i}{\sum_k \lambda_k} \right) (M_1)_{ij}$$

and note that $\lambda_i / (\sum_k \lambda_k)$ is a strategy for player 1 that is a best-response to P; the sum is proportional to how much better a response to j it is than $\bar{\sigma}$ is.

As noted in section 3.4, in LQR we can parameterize the set of satisficing equilibria in terms of $\lambda \in [0, \infty]$ or in terms of a target payoff level or shortfall from optimum, but in this context that doesn't work; if λ gets too large, we can "overshoot".[14]

[14] To clarify the distinction: in a game setting, LQR takes the agent's beliefs as given and supposes the agent plays a distribution that maximizes entropy subject to a given payoff shortfall conditional on those beliefs. In this section we're taking the agent's play as fixed, and asking for the highest-entropy beliefs that would make that a sufficiently good response.

	L	R
T	3,	0,
B	2,	2,

Figure 3.10: "Overshooting".

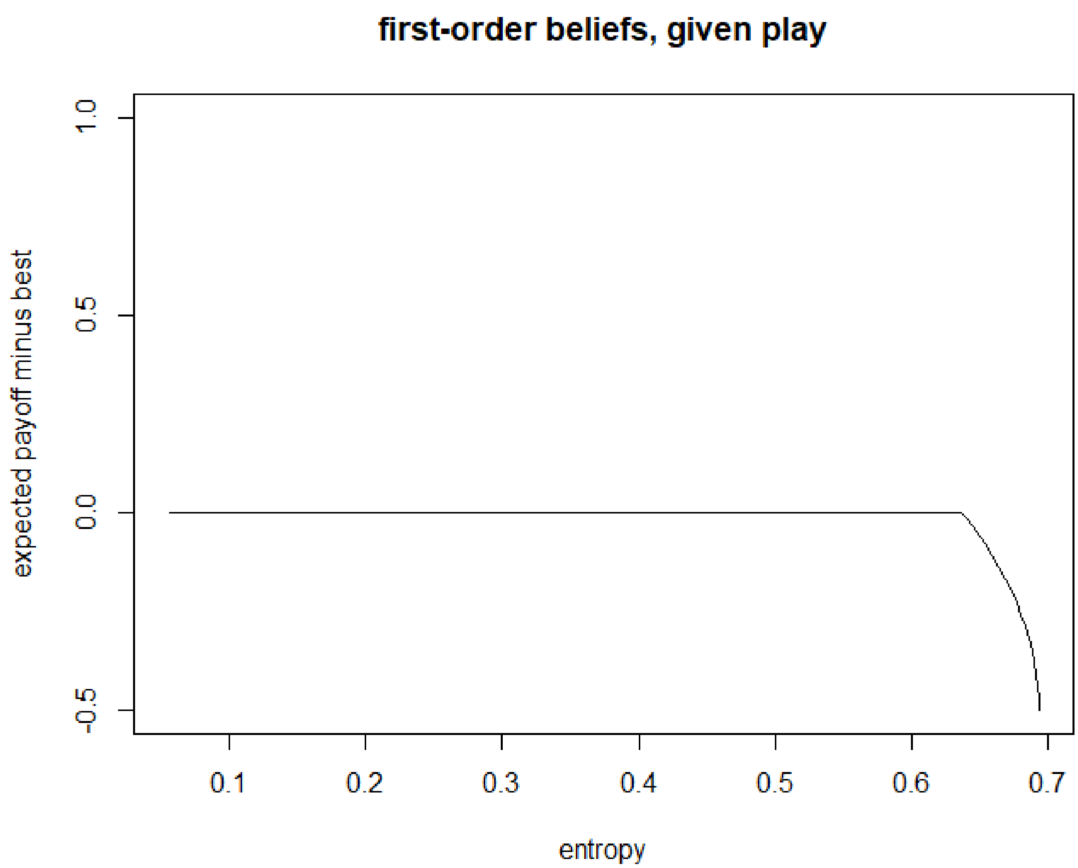


Figure 3.11: The locus of points that optimize expected payoff shortfall conditional on entropy of first-order beliefs, or vice versa. Unlike the LQR setting (figure 3.9), some values of S and λ do not correspond to maximizing entropy conditional on a shortfall.

Consider the game in figure 3.10 with $\bar{\sigma}$ placing full weight on T and with L and R of equal S, thus equally likely if $\lambda = 0$. If $\eta = 0$, then for $\sum_k \lambda_k = \ln(2)/3$ T and B are both best responses, and for greater λ T is the unique best-response. For $\sum_k \lambda_k \in [0, \ln(2)/3]$, the result corresponds to a maximum of entropy conditional on some shortfall constraint, but for larger λ it does not; the resulting beliefs have lower entropy even than the level required for T to be a best-response.

Now we seek $P_j(\sigma)$. Conditional on j being played, we have constraints of

$$\int P_j(\sigma) s_{\sigma j l} \geq -\bar{\epsilon}$$

for each l, where

$$s_{\sigma j l} = \min\left\{0, \sum_i \sigma_i \left((M_2)_{ij} - (M_2)_{il}\right)\right\}, \tag{3.8}$$

so that as a function of σ the constraints are piecewise linear, and player 1's expectation of player 2's expectation of the shortfall from playing j instead of l is bounded. The maximum entropy beliefs[15] that satisfy the other two conditions conditional on player 2's choosing j will satisfy first-order conditions for

$$\int_\sigma \left[-P_j(\sigma)\log P_j(\sigma) + \eta P_j(\sigma)\sum_i \bar{\sigma}_i \log\sigma_i + \sum_l \lambda_{jl} P_j(\sigma) s_{\sigma j l} + \kappa P_j(\sigma)\right]$$

which is to say

$$\log P_j(\sigma) = \eta \sum_i \bar{\sigma}_i \log \sigma_i + \sum_l \lambda_{jl} s_{\sigma j l} + \kappa - 1$$

Again, if (for a given j) not all λ_{jl} (for all l) are 0, then $\lambda_{jl}/\sum_l \lambda_{jl}$ is a best-response to $\int \sigma P_j(\sigma)$,[16] with

$$P_j(\sigma) = C_j \prod_i \left(e^{\sum_l \lambda_{jl} s_{\sigma j l}}\right)\left(\sigma_i^{\eta\bar{\sigma}_i}\right) \tag{3.9}$$

where C_j normalizes the integral to 1; for $\lambda_{jl} = 0$ this is a Dirichlet distribution, which is a beta distribution when there are two possible actions.

---

[15] In what follows, the integral will be done using the Lebesgue measure, out of convenience more than conviction; the choice of measure will correspond to a choice of what is truly "neutral" in the sense Jaynes intends.

[16] i.e. the expected payoff of l, given P_j, is maximal.

$\lambda_{jl} \geq 0$, and is equal to 0 if the well-enough-responding constraint isn't binding at maximum entropy.[17] η will be higher the better player 2 is expected to know $\bar{\sigma}$, with $\eta = 0$ corresponding to complete ignorance; if the agent is supposed to know less, entropy can be made larger, so a positive value of η supposes that we are excluding greater levels of ignorance.

This distribution has an entropy of

$$S_j = -\log(C_j) - \int \mathrm{d}\sigma P_j(\sigma) \left(\sum_l \lambda_{jl} s_{\sigma jl} + \eta \sum_i \bar{\sigma}_i \log(\sigma_i) \right) \tag{3.10}$$

It may be useful to note that, since

$$C_j^{-1} = \int \mathrm{d}\sigma \left(e^{\sum_l \lambda_{jl} s_{\sigma jl}} \right) \prod_i \left(\sigma_i^{\eta \bar{\sigma}_i} \right),$$

$$S_j = -\log(C_j) - C_j \left(\sum_l \lambda_{jl} \frac{\partial C_j^{-1}}{\partial \lambda_{jl}} + \eta \frac{\partial C_j^{-1}}{\partial \eta} \right) \tag{3.11}$$

The same conundrum associated with parameterization of the player 1 rationality constraint/tradeoff applies here as well; in addition, it's not obvious that players should be seeking the same level of payoff or shortfall regardless of beliefs.[18] In the special cases that I will consider below, this doesn't matter; in 2×2 games, the rationality constraint can bind for no more than one action for each player in these second-order beliefs, and where we consider larger games, in the next subsection, we only consider the contexts in which agents are taken to be best-responding to beliefs. Another related issue is that the maximization problem may lack any solution; in particular, if j is a strictly dominated strategy and player 1 believes that player 2 is best-responding, there is no set of beliefs $P_j(\sigma)$ that is consistent with that. In those situations we take $S_j = -\infty$ and $e^{x+S_j} = 0$ for any real number x; player 1 will conclude that player 2 must not be playing j, and will assign 0 probability to that possibility.

---

[17] In particular, for at least one action j, $\lambda_{jl} = 0$ for all l; viz. for j the best-response to the relevant Dirichlet distribution (see below).

[18] An analysis of the data from Nagel and Tang (1998) suggests that shortfalls from the best-response tend to be about 1/5 times the standard deviation in payoffs due to the other player's mixing; a noisier environment will result in agents "satisficing" at lower levels relative to the optimum.

	A	B
A	7,4	0,0
B	0,0	3,6

Figure 3.12: A 2×2 game with two strict equilibria, neither Pareto dominant.

3.5.2 Ignorant but Rational

An interesting and tractable limiting case is that in which agents are mutually known to be best-responding but have no information about the other's intentions. The conditional second-order beliefs take a simple form: because the player knows the other player is rational, conditional on the other player's playing j, the second-order beliefs put full support on the stability set of j; because those beliefs are otherwise unconstrained, the maximum entropy beliefs are those that are uniform on that set. The conditional entropy S_j is thus the logarithm of the measure of the stability set for j, and e^{S_j} is what Harsanyi (1995) calls the "structural incentive" for the other player to play j.[19]

To make discussion of second-order beliefs (and ignorance) more concrete, consider the 2×2 game in figure 3.12, and assume (for this section) mutual knowledge of rationality. Player 2's first-order beliefs are parameterized by the probability that player 1 will choose B, $P_B \in [0, 1]$. Player 1's second-order beliefs are a distribution on $[0, 1]$. Rational play by player 1 will depend only on whether player 2 is choosing B with a probability more than, less than, or exactly .7; rational play by player 2 entails choosing B if $P_B > .4$ and not if $P_B < .4$.

Consider a "level-1" player 1 who just supposes that player 2 plays each action equally; $P_B = 1/2 < .7$, so player 1 chooses A. A "level-2" player 1, who believes that player 2 believes that player 1 chooses each action equally, expects player 2 to choose B, and thus player 1 will choose B as well. Similarly, a level-k model for player 2 suggests

[19] As previously noted, we will use the Lebesgue measure, and the measure he uses is equivalent for 2×2 games, but not in general.

that player 2 plays B when $k = 1$ and A when $k = 2$; if both players have the same k, they will fail to coordinate.

Now consider a different conception of second-order naiveté: instead of believing with probability 1 that player 2 believes that player 1 will play each action with probability 1/2, player 1 believes with probability $y - x$ that player 2 believes that $x \leq P_B \leq y$ for all $0 \leq x \leq y \leq 1$. Letting $\bar{P}_1$ denote the probability of player 1's playing A that results in player 2's being indifferent between the two actions, and $\bar{P}_2$ the probability of player 2's playing A that results in player 1's being indifferent between the two actions, player 1 with these beliefs expects player 2 to play A with a probability of $1 - \bar{P}_1$; player 1 therefore plays A if $1 - \bar{P}_1 > \bar{P}_2$ and B if $1 - \bar{P}_1 < \bar{P}_2$. Thus

Proposition 6. *In a 2×2 game with two strict Nash equilibria, one of which risk-dominates the other, a player with uniform (under the Lebesgue measure) second-order beliefs will play the action associated with the risk-dominant equilibrium under mutual knowledge of rationality.*

Note that this is stronger than the result in Harsanyi (1995), which required that each player calculate every player's structural incentive and combine them in a symmetric way. In our setting, each player solves only that player's own problem, but in 2×2 games with two strict equilibria, it is generically the case that the players will end up in an equilibrium.

Before we move to other games, let's briefly consider the situation in which players have more information, and the second-order beliefs are β-distributions; in particular, let's focus on beliefs that player 2 has some reason to believe that player 1 will play B 80% of the time, viz. the second-order beliefs are $\propto P_B^{(.8)\eta}(1 - P_B)^{(.2)\eta}$. If η is sufficiently large, player 1 expects rational player 2 to play B most of the time, so player 1 plays B. Because (.8,.8) isn't an equilibrium, if they know too well that they will play it, they won't. For η small enough, as before, player 1 plays A. At some point in between player 1 may mix; with mutual knowledge of rationality this is only true for a single value of η.

Consider now the minimum games mentioned in §3.2.1; in the literature, values of $\alpha = 5$ and $\beta = 2$ are typical, with (say) 7 players. Second-order beliefs in this context involve distributions over all other players' beliefs about what all other players are doing, but loosely, we note that if everyone else randomizes with a cdf of $F(a)$ then the minimum of everyone else's action has a cdf of

$$1 - (1 - F(x))^{n-1}$$

and since

$$1 - \left(\frac{2}{5}\right)^{\frac{1}{7-1}} < 1/7,$$

if a player believes that everyone else is randomizing uniformly, that player's best response is 1 (the lowest action in the action space). Similarly, if any player believes that any player has uniform beliefs about what other players will do and optimizes in response to those beliefs, then that player will choose 1.

Finally, in the centipede game of figure 3.5, a player expecting the other to uniformly randomize will play the highest undominated action, and a player expecting the other to do so will play one less, thus 11 and 12 are selected. In this case it is of particular importance that I'm emphasizing second-order beliefs; cutting off the belief hierarchy at a different level, with agents perfectly rational up to that point, would select different actions.[20]

3.5.3 Risk Dominance in 2 by 2 Games

Proposition 7. *Let $\mathcal{G}$ be a 2×2 game with payoffs as in figure 3.2 and $a^k_{11} > a^k_{21}$ and $a^k_{22} > a^k_{12}$ for $k \in \{1, 2\}$. Given any value of $\eta \in \mathbb{R}^+$, and rationality constraints that obey*

$$\frac{\epsilon_1}{a^1_{11} - a^1_{21} + a^1_{12} - a^1_{22}} = \frac{\epsilon_2}{a^2_{11} - a^2_{21} + a^2_{12} - a^2_{22}} > 0$$

[20] As noted before, logistic quantal response equilibria, which is to say agents in this maximum-entropy framework who are not quite best-responding, will also tend to favor actions in this range.

	L	R
T	0,0	10,10
B	10,10	x, y

Figure 3.13: If each player recieves independent signals $\sim N((x, y), \sigma^2 I_2)$, for sufficiently diffuse common prior and sufficiently small σ, a Bayes-Nash equilibrium will result in B,R with an arbitrarily high probability < 1 when $x = y = 9$; neither equilibrium is risk-dominant, and players fail to coordinate on either.

	L	C	R
T	2,2	2,0	2,0
M	0,2	3,0	0,3
B	0,2	0,3	3,0

Figure 3.14: (T,L) is the unique Nash equilibrium, but all strategies are rationalizable; non-Nash actions are consistent with common knowledge that agents are best-responding, but require beliefs that players believe that the other player has beliefs that are wrong; for high values of λ (low values of ϵ), these actions become much harder to support for high values of η than low values.

for any binding constraint. Then the sum of the maximum entropy for each player associated with the risk-dominant equilibrium is greater than or equal to that in the risk-dominated equilibrium.

Let's fix $\eta > 0$, and also $\delta_1 > 0$ and $\delta_2 > 0$, and set

$$\epsilon_{11} = \delta_1 m^1 \qquad m^1 = (a^1_{11} - a^1_{21} + a^1_{22} - a^1_{12})$$

$$\epsilon_{12} = \delta_1 m^2 \qquad m^2 = (a^2_{11} - a^2_{21} + a^2_{22} - a^2_{12})$$

$$\epsilon_{21} = \delta_2 m^2$$

$$\epsilon_{22} = \delta_2 m^1$$

(so that ϵ_{ij} is the maximum amount by which player j may expect the expected payoff to fall short of its optimum at ith order of beliefs) and define

$$p_1 = \frac{a^2_{22} - a^2_{12}}{a^2_{11} - a^2_{12} + a^2_{22} - a^2_{21}}$$

$$p_2 = \frac{a^1_{22} - a^1_{12}}{a^1_{11} - a^1_{12} + a^1_{22} - a^1_{21}}$$

as earlier; p_j is the probability with which player j plays strategy 1 in the mixed-strategy equilibrium. We will similarly use normalized Lagrange multipliers $\mu = \lambda m$.

The constraints can be specified entirely in terms of the p and δ; in particular, if player 1 is choosing ($\bar{\sigma}$) row 1, the first-order (self) rationality constraint binds if

$$P(1) = p_2 - \delta_1$$

while for row 2 it binds if

$$P(1) = p_2 + \delta_1.$$

Similarly, second-order belief constraints are a function of δ_2 and (for player 1) p_1 alone; the constraints can be written as

$$\int P_j(\sigma) s_{\sigma jl}/m^2 \geq -\bar{\epsilon}_{21}/m^2 = -\delta_2$$

where

$$s_{\sigma 12}/m^2 = \min\{0, \sigma_1 - p_1\} \qquad s_{\sigma 21}/m^2 = \min\{0, p_1 - \sigma_1\} \tag{3.8$'$}$$

The entropies of the conditional second-order beliefs are

$$S_j = \log(C_j^{-1}) - \left(\sum_l \lambda_{jl} \frac{\partial \log(C_j^{-1})}{\partial \lambda_{jl}} + \eta \frac{\partial \log(C_j^{-1})}{\partial \eta}\right) \tag{3.11$'$}$$

where the relevant C_j are of the form

$$C_-^{-1}(\eta, \mu, p) = \int_0^p x^\eta e^{\mu(x-p)} \mathrm{d}x + \int_p^1 x^\eta \mathrm{d}x$$

or

$$C_+^{-1}(\eta, \mu, p) = \int_0^p x^\eta \mathrm{d}x + \int_p^1 x^\eta e^{\mu(p-x)} \mathrm{d}x$$

Because these solve constrained maximization problems for which higher p tightens the constraint for C_-^{-1} and loosens the constraint for C_+^{-1}, S_-^{-1} must be non-increasing in p (with δ_2 held fixed, not just with μ held fixed) and S_+^{-1} must be similarly non-decreasing in p.

In equilibrium j the first-order belief of player i places a probability of x_{ij} on the other player's playing 1, where x_{ij} maximizes

$$S_{ij} = x_{ij}(S_1 - \log(x_{ij})) + (1 - x_{ij})(S_2 - \log(1 - x_{ij}))$$

subject to $x_{ij} \in [0, 1]$ and

$$\begin{array}{ccc} 1 - x_{11} \in [0, 1 - p_2 + \delta_2] & S_1 = S_-(p_1) & S_2 = S_+(p_1) \\ x_{22} \in [0, p_1 + \delta_2] & S_2 = S_-(1 - p_2) & S_1 = S_+(1 - p_2) \\ x_{12} \in [0, p_2 + \delta_2] & S_2 = S_-(1 - p_1) & S_1 = S_+(1 - p_1) \\ 1 - x_{21} \in [0, 1 - p_1 + \delta_2] & S_1 = S_-(p_2) & S_2 = S_+(p_2) \end{array}$$

where S_{ij} is the maximum entropy of first-and-second order beliefs for which player i plays j.

The maximization problem for player 1 in equilibrium 1 and that for player 2 in equilibrium 2 are seen to be equivalent if $p_1 + p_2 = 1$; otherwise, the problem in the risk-dominated equilibrium maximizes the same function on a strict subset of the domain available in the risk-dominant equilibrium. This is also true of the comparison of the maximization problem for player 1 in equilibrium 2 and that for player 2 in equilibrium 1. Thus the sum of the achievable entropies is equal if $p_1 + p_2 = 1$, and if one of the sums is larger than the other, it must be that corresponding to the risk-dominant equilibrium.

3.6 Conclusion

The term "risk-dominance" has long been used in a loose way to suggest when we players in a game with multiple equilibria might be expected to find it easier to coordinate on one than on one or more others. In this essay I have noted how we might make this more precise: using an information theoretic measure of ignorance that players have of each others' plans, and using Jaynes's maximum entropy idea, there is a mathematically precise sense in which agents who are sufficiently constrained in their prior information about each others' plans should be expected to hold beliefs that should lead them to play a strategy profile that, in the case of a 2×2 game, is the risk-dominant equilibrium.

The information theoretic tools that have been used can be understood to provide something of a reduced-form model of context that has been abstracted away from the strategic-form game itself; in principle we might, as in the "global game" idea, embed our finite game in a fully-specified game of incomplete information. Sometimes the results can be very sensitive to the details of the additional structure (Weinstein and Yildiz (2011)); instead we have adopted a much more parsimonious model.

Harsanyi (1995) finishes (and Harsanyi and Selten (1988) starts) with an argument that there should be a precise theory of equilibrium selection. I disagree. Harsanyi and Selten (1988) notes that the epistemic case for Nash equilibrium relies on common knowledge of *which* equilibrium will obtain, and suggests that a rule should be commonly agreed upon that would then become self-fulfilling. Harsanyi (1995) notes that non-cooperative models of situations that seem like cooperative games — e.g. of bargaining — often give results that are sensitive to the strategic dynamics of the non-cooperative game. This feels to me to go too far toward assuming away coordination problems that I feel should be modeled instead.

Savage (as reported in Binmore (2015)) insisted that a positive model of rationality be restricted to "small worlds"; in practice, this would mean a setting in which a "small" strategic situation can be abstracted from the big world without affecting its analysis. A good theory should predict its own domain of applicability, in the sense that it should provide guidance as to when it can't make a good prediction; we shouldn't expect to be able to impose a single outcome on a simple strategic situation that is likely to be highly sensitive to the context from which it has been deracinated.

Bibliography

ALLAIS, M. (1953): "Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américane," *Econometrica*, 21(4), 503–546.

ALÓS-FERRER, C., AND N. NETZER (2010): "The logit-response dynamics," *Games and Economic Behavior*, 68(2), 413–427.

ARAGONES, E., I. GILBOA, A. POSTLEWAITE, AND D. SCHMEIDLER (2005): "Fact-Free Learning," *American Economic Review*, 95(5), 1355–1368.

BAYER, R.-C. (2010): "Intertemporal price discrimination and competition," *Journal of Economic Behavior & Organization*, 73(2), 273–293.

BINMORE, K. (2015): "Rationality," in *Handbook of Game Theory with Economic Applications*, vol. 4 of *Handbook of Game Theory with Economic Applications*, chap. 1, pp. 1–26. Elsevier.

BINMORE, K., L. SAMUELSON, AND J. GALE (1995): "Learning to be Imperfect: The Ultimatum Game," *Games and Economic Behavior*, 8, 56–90.

BLUME, L. E. (1993): "The Statistical Mechanics of Strategic Interaction," *Games and Economic Behavior*, 5(3), 387–424.

CAMERER, C., AND T.-H. HO (1999): "Experience-weighted Attraction Learning in Normal Form Games," *Econometrica*, 67(4), 827–874.

CAMERER, C. F., T.-H. HO, AND J.-K. CHONG (2004): "A Cognitive Hierarchy Model of Games," *The Quarterly Journal of Economics*, 119(3), 861–898.

Carlsson, H., and E. van Damme (1993): "Global Games and Equilibrium Selection," *Econometrica*, 61(5), 989–1018.

Cavallo, A., E. Cavallo, and R. Rigobon (2014): "Prices and Supply Disruptions during Natural Disasters," *Review of Income and Wealth*, 60(S2), S449–S471.

Cooper, R. W., D. V. DeJong, R. Forsythe, and T. W. Ross (1990): "Selection Criteria in Coordination Games: Some Experimental Results," *American Economic Review*, 80(1), 218–233.

Diamond, D. W., and P. H. Dybvig (1983): "Bank runs, deposit insurance, and liquidity," *Journal of Political Economy*, 91(3), 401–419.

Duffie, D., N. Gârleanu, and L. H. Pedersen (2005): "Over-the-counter Markets," *Econometrica*, 73(6), 1815–1847.

——— (2007): "Valuation in Over-the-Counter Markets," *Review of Financial Studies*, 20(6), 1865–1900.

Ellsberg, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, 75(4), 643–669.

Fischbacher, U. (2007): "z-Tree: Zurich Toolbox for Ready-made Economic Experiments," *Experimental Economics*, 10(2), 171–178.

Frankel, D. M., S. Morris, and A. Pauzner (2003): "Equilibrium selection in global games with strategic complementarities," *Journal of Economic Theory*, 108(1), 1–44.

Gul, F., and W. Pesendorfer (2001): "Temptation and Self-Control," *Econometrica*, 69(6), 1403–1435.

Harrison, J. M., and D. M. Kreps (1978): "Speculative Investor Behavior in a Stock

Market with Heterogeneous Expectations," *The Quarterly Journal of Economics*, 92(2), 323–36.

Harsanyi, J. C. (1995): "A New Theory of Equilibrium Selection for Games with Complete Information," *Games and Economic Behavior*, 8(1), 91–122.

Harsanyi, J. C., and R. Selten (1988): *A general theory of equilibrium selection in games*. MIT Press, Cambridge, Mass.

Hendel, I., P. Dudine, and A. Lizzeri (2006): "Storable Good Monopoly: The Role of Commitment," *American Economic Review*, 96(5), 1706–1719.

Hillas, J., and E. Kohlberg (2002): "Foundations of strategic equilibrium," in *Handbook of Game Theory with Economic Applications*, ed. by R. J. Aumann, and S. Hart, vol. 3, chap. 42, pp. 1597–1663. Elsevier, 1 edn.

Hofbauer, J., and W. H. Sandholm (2002): "On the Global Convergence of Stochastic Fictitious Play," *Econometrica*, 70(6), 2265–2294.

Jaynes, E. T. (1957): "Information Theory and Statistical Mechanics I," *Physical Review*, 106, 620–630.

Jaynes, E. T. (2003): *Probability Theory: The Logic of Science*. Cambridge University Academic Press, Inc.

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263–291.

Kiyotaki, N., and R. Wright (1989): "On Money as a Medium of Exchange," *Journal of Political Economy*, 97(4), 927–54.

——— (1993): "A Search-Theoretic Approach to Monetary Economics," *American Economic Review*, 83(1), 63–77.

McKelvey, R. D., and T. R. Palfrey (1995): "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10(1), 6–38.

Mertens, J.-F. (2002): "Stochastic games," in *Handbook of Game Theory with Economic Applications*, ed. by R. J. Aumann, and S. Hart, vol. 3, chap. 47, pp. 1809–1832. Elsevier, 1 edn.

Mitraille, S., and H. Thille (2014): "Speculative storage in imperfectly competitive markets," *International Journal of Industrial Organization*, 35(C), 44–59.

Monderer, D., and L. S. Shapley (1996): "Potential Games," *Games and Economic Behavior*, 14(1), 124–143.

Mookherjee, D., and B. Sopher (1994): "Learning Behavior in an Experimental Matching Pennies Game," *Games and Economic Behavior*, 7(1), 62–91.

——— (1997): "Learning and Decision Costs in Experimental Constant Sum Games," *Games and Economic Behavior*, 19(1), 97–132.

Morris, S. (1996): "Speculative Investor Behavior and Learning," *The Quarterly Journal of Economics*, 111(4), 1111–1133.

Morris, S., R. Rob, and H. S. Shin (1995): "p-Dominance and Belief Potential," *Econometrica*, 63(1), 145–57.

Nagel, R. (1995): "Unraveling in Guessing Games: An Experimental Study," *American Economic Review*, 85(5), 1313–26.

Nagel, R., and F. F. Tang (1998): "Experimental Results on the Centipede Game in Normal Form: An Investigation on Learning," *Journal of Mathematical Psychology*, 42, 356–384.

Nichols, A. L., and R. J. Zeckhauser (1977): "Stockpiling Strategies and Cartel Prices," *Bell Journal of Economics*, 8(1), 66–96.

PESKI, M. (2010): "Generalized risk-dominance and asymmetric dynamics," *Journal of Economic Theory*, 145(1), 216–248.

ROCHETEAU, G., AND P.-O. WEILL (2011): "Liquidity in Frictional Asset Markets," *Journal of Money, Credit and Banking*, 43, 261–282.

ROCKAFELLAR, R. T. (1970): *Convex Analysis*. Princeton University Press.

ROTH, A. E. (2008): "What Have We Learned from Market Design?," *Economic Journal*, 118(527), 285–310.

ROUSSEAU, J. J. (1950): *The Social Contract and Discourses*. E.P. Dutton and Company, New York, translation by G.D.H. Cole.

SALANT, Y. (2011): "Procedural Analysis of Choice Rules with Applications to Bounded Rationality," *American Economic Review*, 101(2), 724–748.

SHAPLEY, L. S. (1953): "Stochastic games," *PNAS*, 39(10), 1095–1100.

SHORE, J. E., AND R. W. JOHNSON (1980): "Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy," *IEEE Transactions on Information Theory*, 26(1), 26–37.

STAHL, D. I., AND P. W. WILSON (1994): "Experimental evidence on players' models of other players," *Journal of Economic Behavior & Organization*, 25(3), 309–327.

STAHL, D. O., AND P. W. WILSON (1995): "On Players' Models of Other Players: Theory and Experimental Evidence," *Games and Economic Behavior*, 10(1), 218–254.

UI, T. (2001): "Robust Equilibria of Potential Games," *Econometrica*, 69(5), 1373–1380.

Van Damme, E. (2002): "Strategic equilibrium," in *Handbook of Game Theory with Economic Applications*, ed. by R. J. Aumann, and S. Hart, vol. 3 of *Handbook of Game Theory with Economic Applications*, chap. 41, pp. 1521–1596. Elsevier.

Van Huyck, J. B., R. C. Battalio, and R. O. Beil (1990): "Tacit Coordination Games, Strategic Uncertainty, and Coordination Failure," *American Economic Review, American Economic Association*, 80(1), 234–48.

Vayanos, D., and J. Wang (2011): "Theories of Liquidity," *Foundations and Trends in Finance*, 6(4), 221–317.

Vieille, N. (2002): "Stochastic games: Recent results," in *Handbook of Game Theory with Economic Applications*, ed. by R. J. Aumann, and S. Hart, vol. 3 of *Handbook of Game Theory with Economic Applications*, chap. 48, pp. 1833–1850. Elsevier.

Wallace, N. (1988): "Another attempt to explain an illiquid banking system: the Diamond and Dybvig model with sequential service taken seriously," *Quarterly Review*, (Fall), 3–16.

Weinstein, J., and M. Yildiz (2011): "Sensitivity of equilibrium behavior to higher-order beliefs in nice games," *Games and Economic Behavior*, 72(1), 288–300.

Yaari, M. (1987): "The Dual Theory of Choice under Risk," *Econometrica*, 55(1), 95–115.

Young, H. P. (1993): "The Evolution of Conventions," *Econometrica*, 61(1), 57–84.

Young, H. P. (1998): *Individual Strategy and Social Structure*. Princeton University Press.