Aggregating Probabilistic Beliefs: Market Mechanisms and Graphical Representations

by

David M. Pennock

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Computer Science and Engineering)
in The University of Michigan
1999

Doctoral Committee:

Associate Professor Michael P. Wellman, Chair
Associate Professor Edmund H. Durfee
Professor Daniel E. Koditschek
Professor Stephen M. Pollock

ABSTRACT

Aggregating Probabilistic Beliefs: Market Mechanisms and Graphical Representations

by

David M. Pennock

Chair: Michael P. Wellman

A long-standing question in statistics is how best to aggregate the probabilistic beliefs of multiple agents. Related is the practical question of how to represent the combined beliefs efficiently. This dissertation reports contributions on both fronts.

First, I formulate and analyze a securities market mechanism for aggregating beliefs. Equilibrium prices in the market are interpreted as consensus beliefs. Under homogeneity conditions regarding agents' utilities, the market mechanism corresponds with standard aggregation functions, and the market's outward behavior is indistinguishable from that of an individual. I also explore extensions to the model in which agents learn from prices and the market as a whole adapts over time. In certain circumstances, price fluctuations can be viewed as the Bayesian updates of a rational individual.

Second, I investigate the use of graphical models, and in particular Bayesian networks, for representing aggregate beliefs. I derive two impossibility theorems which contradict widely held intuitions about how Bayesian networks should be combined. The so-called logarithmic opinion pool is shown to admit relatively concise encodings. I describe the nature of graphical structures consistent with this pooling function, and give algorithms for computing the logarithmic and linear opinion pools with, in some cases, exponential speedups over standard methods.

Finally, I apply and extend the graphical modeling results to the market framework, deriving sufficient conditions for compact markets to be operationally complete. Such markets still induce a complete consensus distribution and support Pareto optimal allocations of risk, but with exponentially fewer securities than required for traditional completeness.

To my parents ...

for unsurpassed support and encouragement.

Acknowledgements

I would like to thank, first and foremost, my advisor, Mike Wellman. His vision, ideas, technical insight, guidance, support, encouragement, and editorial voice underlie any and all of the positive contributions of this dissertation, from motivations to specific results to writing style. While I can't match his intelligence, I do aspire to emulate his diligence, and his dedication to students. No one can persuade me that I did not have his complete and undivided attention throughout my time at UM—someone else must have tended to his other students and pursued his other research and teaching duties. His feedback on my ideas and writings are always thorough, cogent, and timely. By devoting obvious time and effort to my cause, he inspires me to respond in kind and to produce my best work. Finally, I simply don't know when he sleeps: he answers all email immediately and completely at any time, day or night. If you don't believe me, see for yourself: wellman@umich.edu (sorry, Mike).

I also owe a debt of gratitude to the other members of my dissertation committee. Stephen Pollock was as conscientious and responsive a committee member as I could imagine, and his grasp of decision-theoretic concepts helped me immeasurably. From the beginning, he took great interest in this work, and throughout the process he remained a completely committed cognate constituent of my committee. Dan Koditschek offered a fresh perspective through the lens of dynamical systems. Among other contributions, he was a great proponent for mathematical rigor. Ed Durfee brought to the table a deep knowledge of, and experience with, group coordination problems; he also encouraged me to step back to examine the big picture and consider real-world implications. All of the committee members' enthusiasm was infectious, especially at the proposal and defense, and their inputs influenced the final product greatly.

Thanks are due to Eric Horvitz for guidance, ideas, support, and inspiration during my internship at Microsoft Research (MSR). Many of the results in Chapters 5 and 7 began to take shape while at MSR with Eric's help. In addition, the survey of research on group coordination in Section 2.1.6 was conducted under his watch. Many other people at MSR deserve acknowledgement; I would especially like to thank Jack Breese and Chris Meek.

I'd like to thank the entire Decision Machine Research Group (DMG)—Pete Wurman, David Pynadath, Chao-Lin Lui, Junling Hu, Bill Walsh, Terence Kelly, Tracy Mullen, John Cheng, and Dan Reeves—for invaluable feedback on my presentations and writings. I enjoyed our spirited discussions on wide ranging topics, and learned at least as much from the members of DMG as from any courses or readings.

Thanks to Frank Yates for welcoming me into the Michigan Decision Behavior Consortium, thereby exposing me to a world of relevant interdisciplinary research and a broad array of colleagues, to whom I owe thanks as well. Bob Savit, Rick Riolo, Franco Nori, Kathy Richards, and other faculty, staff, and students in the Program for the Study of Complex Systems played a large role in my education, as did the reading group on complex systems—including Ted Belding, Leeann Fu, and Corinne Coen—and friends at the 1996 Santa Fe Institute Complex Systems Summer School.

Jeff MacKie-Mason provided valuable feedback at my dissertation proposal and in subsequent meetings. Quentin Stout got me started on the right foot at UM, directing my oral qualification research project. Sang Lee and Bill Rounds offered constructive advice and ideas at my oral qualification exam, and in chance meetings in the halls and rooms of the AI Lab. I am grateful to Atul Prakash and Matthew Jones for their guidance during my stint as a teaching assistant. Pete, Karsten, Peter, Glen, James, and Fritz helped me to hammer home hard AI concepts in preparation for the written qualification exam.

The entire AI Lab staff has been incredibly supportive (as well and friendly and fun). Without their help, I would still be lost in bureaucracy and paperwork. Through their diligence, I was entirely shielded from such headaches. I'd especially like to thank Kathy, Cindy, Deb, Jeff, Karen, Dolores, and Rob.

Throughout my education, many teachers and professors have smoothed the path toward this dissertation. Alan Biermann was an exceptional teacher and mentor during my time in graduate school at Duke. His engaging style fueled my interest in AI research and software for computer science education. John Reif helped stir up inventive juices, and greatly influenced my decision to pursue a doctorate in computer science in the process. Thomas Anewalt enriched my sense of the wonder and beauty of math and science.

This dissertation incorporates insights, feedback, and relevant pointers from Didier Dubois, Robin Hanson, Chris Meek, Robert Nau, Charles Plott, Ross Shachter, James

Smith, Mike West, and Fredrik Ygge.

For making UM a fun place, and for keeping me sane, I'd like to thank my friends in Ann Arbor, including Karsten, Denise, Diana, Roger, Pete, Tamara, Sam, Andrea, Ann, Crystal, Mehmet, David, the Rackham social group, and the floorball-ers. I am grateful for Rachel's constant and unconditional support, and her encouragement, care, time, energy, patience, and selflessness. Kerry's kindhearted soul is a source of strength; her silly and satirical sensibility a steady source of satisfaction and smiles. Anne was a good friend and companion; her drive to succeed couldn't help but rub off on me.

Thanks, Mom and Dad. Thanks for supporting me without exception, throughout the development of this dissertation, and indeed throughout my life. Thanks for encouraging me without pressuring. Thanks for guiding me without directing. For suggesting without imposing. For always expecting the best while accepting the rest. For incubating in me the desire to learn, explore, and create. The fingerprints of your wonderfully expert nurturing—and Nature-ing!—permeate this research and this dissertation.

A dusting for other fingerprints would surely implicate the rest of my family, including Jen, Moss, Todd, Danny, Sheryl, Aunt Marie, Aunt Penny, Aunt Margie, Momom, Nana, Popop, and David Pennock—The Original.

Many more people played pivotal roles in making this dissertation possible, and I apologize for not listing everyone here. My sincerest thanks go out to all of them.

Table of Contents

List of Figures

Figure

Chapter 1

Introduction

Were it not for uncertainty, foresight would be 20/20 too. All decisions would be just as good as those deemed best in hindsight.

Alas, foresight is rather more nearsighted, and the (*ex post*) quality of decisions often hinges greatly on the outcomes of uncertain events. The decision to carry an umbrella, or wear a seatbelt, or purchase flood insurance, depends on the probabilities of rain, car accidents, and floods. Decisions also rest on one's evaluation of the relative utilities (benefits modulo costs) of the possible outcomes—the discomfort of being rained on versus the inconvenience of carrying an umbrella; the pain of injury versus the time and energy cost of buckling up; the devastation incurred by flood versus the price of the insurance contract. Decision theory posits that, without the benefit of hindsight, one should choose actions that maximize *expected* utility, with respect to one's current subjective assessments of the probabilities.

Accurate assessments of probabilities are often crucial for good decision making. How does one arrive at these assessments? Take, for example, the probability of rain tomorrow. One could observe today's weather and extrapolate to tomorrow. One could look up the average weather conditions in an almanac or pour over satellite imagery. Or, as is usually the case for weather, one could base one's own opinion on the reported forecasts of one or more experts (i.e., meteorologists). But what should be made of several experts' differing assessments? How can one arrive at a *consensus* probability? For decades, statisticians have grappled with this problem of *aggregating beliefs*, or combining a group of agents' subjective probability distributions to form a single, representative distribution. Many solutions have been proposed over many years in statistics and the decision sciences. The simplest aggregation functions are the *linear opinion pool* (LinOP), which computes a weighted average of the agents beliefs, and the *logarithmic opinion pool* (LogOP), which

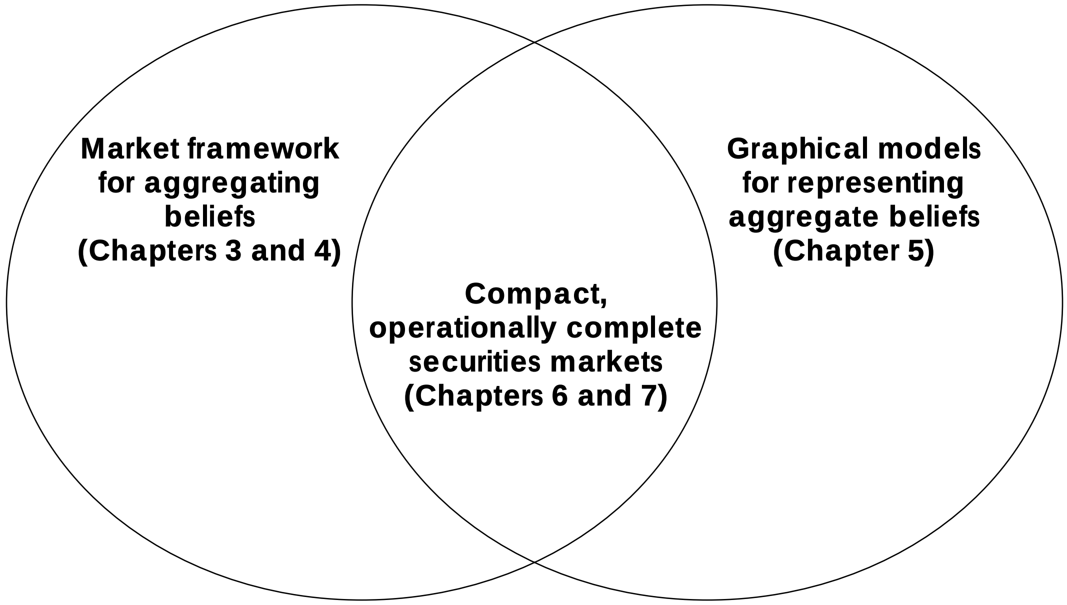


Figure 1.1: Venn diagram portraying the organization of this dissertation, and the relationship between the three main categories of contributions.

computes a weighted geometric mean. This dissertation develops a *market* framework for pooling opinion and analyzes its merits and its reasonableness in a variety of ways, including how it compares to these standard methods.

Tangential to the particular mathematical form of a belief aggregation function are purely computational concerns. Can we design algorithms for computing the aggregate distribution efficiently? Can we design *decentralized* algorithms that respect the separation, privacy, and self-interest of the individuals in the group? Once computed, can we represent that distribution compactly, for example using a Bayesian network (BN)? In particular, if each individual distribution is represented as a structured BN, can the aggregate distribution retain some of that structure? This dissertation investigates each of these questions.

Chapter 2 motivates this research and provides background material on decision theory, belief aggregation, markets, and graphical models. Subsequent chapters present research contributions, organized into three main categories. The relationship between these three categories is pictured in Figure 1.1. First, Chapters 3 and 4 develop a market mechanism for aggregating beliefs in a decentralized manner. Chapter 3 derives and critiques various properties of the market mechanism through formal analysis, and Chapter 4 through empirical simulation. Second, Chapter 5 describes algorithms and data structures for combining probabilities efficiently and representing them compactly using graphical models like BNs. That chapter also derives two impossibility results which circumscribe further efficiency gains. Third, Chapters 6 and 7 present results at the intersection of the first two sets of results, describing when it possible to build *structured* markets, with topologies and semantics similar to BNs, that support both decentralized aggregation and compact representations. Sections 1.1, 1.2, and 1.3 further preview each of these three categories of results.

1.1 Market-Based Belief Aggregation (Chapters 3 and 4)

Consider again the decision to purchase flood insurance. Presumably the cost is greater for a home on the bank of the Mississippi than for a similar home in Santa Fe, New Mexico. That is, the price reflects to some degree the probability of a flood. Purchasing flood insurance is essentially like buying a lottery ticket, or a security, that pays off contingent on the occurrence of a flood.

Now consider another security, this one paying off \$1 if and only if Al Gore wins the 2000 US Presidential election. If you are reading this after November of 2000, the security is worth either \$1 or nothing at all, depending on the outcome of the election. In the meantime, its value to a potential buyer is that person's assessment of the probability that Gore wins.[1] If the security is publicly traded, its current market price can be viewed as an amalgamation of the subjective probabilities of all participants in the market. Notice that the aggregation is naturally decentralized and there are monetary incentives for the agents to bet according to their true beliefs. Such securities are actually available for exchange, at the Iowa Electronic Market (IEM)[2] (Forsythe et al., 1992), and *Slate* magazine plans to use IEM as its main index to the 2000 election "horse race".

Chapter 3 formalizes this notion of equilibrium prices as consensus probabilities. I develop a straightforward model of a securities market and characterize how prices depend on the participating agents' beliefs and utilities. The LinOP and LogOP are shown to arise as special cases of the market mechanism, where agents' risk tolerances in the market play

[1]This assumes that the buyer is risk-neutral for small dollar amounts, and has no other significant stakes in the election outcome. For example, Gore himself, who presumably has a large stake in the outcome, might actually want to *sell* large quantities of the security, even if the price is less than his subjective probability, in order to hedge against the emotional toll of losing the election. Never mind that the action of betting against oneself might not be well received, or might even depress the price and adversely affect public perception of his chances. Or that, if he sold enough, he might prefer to throw the election.

[2]`http://www.biz.uiowa.edu/iem/`

the role of expert weights in the opinion pools. For a fixed configuration of securities, the collective behavior of all agents can be rationalized as that of a *composite* agent whose beliefs equal the equilibrium prices. On the other hand, if an external observer can create and cancel markets, I show how (s)he might distinguish between a group and an individual. Section 4.1 extends the basic model to include agents that learn from prices. In this case, expert weights are shown to additionally depend on self-assessed confidence. Section 4.2 examines a multiperiod market via simulation and analysis. For some agent constituencies, changes in price over time exactly mimic the belief updates of a Bayesian-rational individual observing the event outcomes. For these agents, the market's reward structure facilitates adaptation—the process of transferring wealth from inaccurate agents to accurate agents causes prices to converge toward the observed frequencies of events.

1.2 Graphical Representations of Aggregate Belief (Chapter 5)

Return to the task of assigning probabilities to RAIN, CAR-ACCIDENT, and FLOOD. Assume that the three events are well-defined binary propositions (e.g., the spatio-temporal definition of the occurrence of rain is precise). A full joint probability distribution describing this situation would assign a probability to any seven of the eight possible combinations of outcomes (the eighth follows from normalization). As the number of events under consideration increases, the number of probabilities required to fully specify a joint distribution grows exponentially. Yet often, probabilistic relationships can be specified more naturally and compactly in terms of local probabilistic dependencies among events. For example, if FLOOD and CAR-ACCIDENT are conditionally independent, given knowledge of the outcome of RAIN, then only five probability assignments are required rather than seven. Graphical models, and in particular BNs, specify joint distributions in terms of events and the conditional dependencies among them (Jensen, 1996; Pearl, 1988; Whittaker, 1990). Adopting such a representation can effect an exponential reduction in the space required to encode an otherwise unwieldy joint distribution.

Even when all agents agree on a particular independence, none of the standard opinion pools generically maintain that independence within the aggregate distribution. In fact, researchers have shown that very weak constraints completely rule out such independence preservation (Genest and Wagner, 1987; Lehrer and Wagner, 1983; Wagner, 1984). In Chapter 5, I show that, by restricting attention to the subset of independencies representable as graphical models, new possibilities emerge. An additional, though trivial, assumption is required to rule out independence preservation in this context. Moreover, a significant subclass of independencies, called Markov independencies, are maintained by the logarithmic opinion pool (LogOP). I give procedures for constructing Markov networks (MNs) and decomposable BNs that represent the aggregate beliefs implied by the LogOP. In some cases exponential improvements are possible in both the size of the representation and the speed of inference. Chapter 5 also characterizes the computational complexity of the linear opinion pool (LinOP), when input distributions are given as BNs.

1.3 Structured Securities Markets (Chapters 6 and 7)

As mentioned, securities serve both to allocate risk (e.g., transfer risk from an insurance purchaser to an insurance company, in return for a sure payment) and to disseminate probabilistic information. *Complete* securities markets, which contain one security for every possible state of nature, support Pareto optimal allocations of risk and define probabilities across all states of nature.

Complete markets suffer from the same exponential dependence on the number of underlying events as do joint distributions. Chapters 6 and 7 examine whether markets can be structured and "compacted" in the same manner as BN representations of joint distributions. Chapter 6 takes a top-down approach, where given independencies are explicitly enforced by introducing appropriate arbitrageurs into the economy. On the other hand, Chapter 7 derives structured markets that conform to the belief structures of given agents. I show that, if all agents' risk-neutral independencies agree with the independencies encoded in the market structure, then the market is *operationally complete*. In such a market, risk is still Pareto optimally allocated, and securities prices still define a complete consensus. I also show that, for collections of agents of a certain type, agreement on Markov independencies is sufficient to admit compact and operationally complete markets.

Chapter 2

Background

2.1 Decision-Theoretic Foundations

2.1.1 States and Events

The *state* of the world is, at the most fine-grained level, the position and velocity of every particle in the universe (or the appropriate quantum-mechanical analog). In practice, state is of course more narrowly defined: for IBM's computer chess program *Deep Blue*, it is the position of every piece on the board. In most settings, the appropriate level of modeled detail is an important and difficult design consideration. For an office robot, the state might include the locations of walls, power sources, people, itself, other robots, the intentions of other robots, etc. Denote the set of all possible modeled states of the world (for *Deep Blue*, every legal chess board combination) as $\Omega = \{\omega_1, \omega_2, \ldots\}$. The ω are mutually exclusive and exhaustive.

State is often more concisely and naturally characterized as the set of outcomes of *events*. Events are logical true/false propositions; for example, "It rains in Ann Arbor on 9/23/99", "I am involved in a car accident on 9/23/99" and "My apartment floods on 9/23/99". Denote the set of modeled events as $Z = \{A_1, A_2, \ldots, A_M\}$. Underlying M arbitrary events is a state space Ω of size $|\Omega| = 2^M$, consisting of all possible combinations of event outcomes (in the example, $|\Omega| = 2^3 = 8$). Conversely, any set of states can be factored into a set of $M = \lceil \lg |\Omega| \rceil$ events. Without further assumption, the two representations are equivalent in both expressivity and size, although the event factorization may be more natural: if I am purchasing stock, planning a picnic, or buying insurance, I am more directly concerned with the realizations of the example events rather than the underlying state space. Additionally, if independencies are present among events, a graphical representation may yield models that are both concise and insightful, as discussed in Section 2.3.4.

In most of what follows, the events A_j are the focus of attention, with Ω the implied joint outcome space. The A_j are sometimes referred to as the *primary* events, so as to distinguish them from the other $2^{2^M} - M$ possible sets of states, each of which is also an event.

2.1.2 Uncertainty and Belief

Uncertainty is ubiquitous, ... almost surely. Seatbelts, insurance, spare tires, and other instruments for coping with risk are manifestations of this reality. Even ignoring Heisenberg uncertainty, real-world agents cannot fully observe the current state nor predict the future state. One reason is *laziness*, or the inability to even record the attributes of state. A second reason is *ignorance*, brought about by imperfect and partial knowledge, and noisy sensors (Russell and Norvig, 1995).

Variable prices for insurance contracts and the common practice of carrying one but *not* two spare tires are evidence of relative *degrees of belief*. Under different axiomatizations, each persuasive in its own right, it has been shown that *subjective probability theory* is the *only* consistent framework for reasoning about degrees of belief (Cox, 1946; de Finetti, 1974; Savage, 1954; von Neumann and Morgenstern, 1953). It provides the foundation for handling uncertainty adopted by most decision theorists (Raiffa, 1968), economists (Mas-Colell et al., 1995), and statisticians (Berger, 1985), and is enjoying growing popularity and success within AI (Pearl, 1988; Jensen, 1996). Alternative formalisms exist (e.g., possibility theory (Zadeh, 1979; Dubois and Prade, 1988), Dempster-Shafer belief functions (Dempster, 1967; Shafer, 1976), epistemic beliefs (Spohn, 1988), and the transferable belief model (Smets and Kennes, 1994)), and some scientists oppose the subjective interpretation. Further philosophical discussion is beyond the scope of this dissertation; see Chapter 2 of Neapolitan's text (1990) for an introduction to some of the issues.

One of the most convincing justifications for subjective probabilities is de Finetti's famous *no Dutch book* argument (de Finetti, 1974). Imagine that a bettor B must decide at what price to buy and sell each of two lottery tickets: one pays \$1 if an event A_1 occurs, and nothing otherwise; the other pays \$1 if A_2. If A_1 and A_2 are disjoint events, then B's total price for both tickets should never exceed \$1—otherwise, B is exposed to a guaranteed loss: a shrewd opponent O will sell both tickets to B for more than \$1, paying out at most \$1 in all outcomes. More generally, if B's internal prices for a set of such lottery tickets are not legal probabilities, then B is subject to certain losses—also called *Dutch books*, or *arbitrage* (Varian, 1987). Behavior reflective of some coherent probability distribution is both necessary and sufficient for the avoidance of arbitrage (de Finetti, 1974).

2.1.3 Aggregate Belief

Now imagine that a *group* of bettors, $B_1, B_2, \ldots$ are locked in a room. O is outside, and slides offers to buy and sell lottery tickets under the door. To be concrete, imagine that B_1 will sell "\$1 if A_1" at \$0.5, and B_2 will buy the same ticket at \$0.7. O could potentially buy a ticket from B_1 and sell it to B_2 for a riskless profit of \$0.2. O does not care who accepts the offers: if opportunities for risk-free profit exist, O will take advantage of them. From O's perspective, such opportunities imply that the group or "room" is irrational. To *collectively* avoid sure losses, the group's outward behavior must appear to follow a single probability distribution, called variously a *consensus*, *aggregate*, or group distribution, and denoted Pr_0.

The argument requires of the group only consistency, and places no constraints on the mechanism to reach consensus or the relationship between individual and group beliefs. Other motivations for studying aggregate belief appear in Section 2.2; here I simply describe the generic formulation. Given N agents, indexed $i = 1, \ldots, N$, each with a subjective probability distribution Pr_i over Ω, an *opinion pool* is any function that aggregates many beliefs into a single belief,

$$\mathrm{Pr}_0 \equiv f(\mathrm{Pr}_1, \mathrm{Pr}_2, \ldots, \mathrm{Pr}_N). \tag{2.1}$$

In Section 2.3.1, we review various proposed forms for f, and some of the arguments for and against each.

Return to our hypothetical room of bettors, where O can presumably extract a risk-free profit from the group by buying a lottery ticket from B_1 and selling it to B_2. If, however, the group were to first exhaust all opportunities for within-group trades, then B_1 would sell instead to B_2. Each bettor's transactions are exactly the same and the sum of all bettors' coffers is larger, by \$0.2. Only if there is *excess demand* within the group for some lottery ticket at some price (unmatched offers to buy or sell), should they accept a countering outside offer. If the group is in an *equilibrium* state—that is, all possible within-group trades have been exhausted—then arbitrage opportunities cannot exist (Nau and McCardle, 1991). From O's perspective, the group appears rational. In other words, by de Finetti's same argument, the group behaves as if guided by a *representative* probability distribution. In this example, the group essentially created a market in lottery tickets, and the equilibrium prices might be interpreted as the group's consensus probability distribution. In Chapters 3 and 4, I examine the reasonableness of this interpretation under various market modeling assumptions.

2.1.4 Preference, Utility, and Decision

Preferences refer to ordinal rankings of outcomes. For example, B might prefer sunny days (sd) to cloudy days (cd), and cloudy days to rainy days (rd). Utilities, on the other hand, are numeric expressions. B's utilities $\mathbf{u}$ for the outcomes sd, cd, and rd might be $u(sd) = 10$, $u(cd) = 4$, and $u(rd) = 2$, respectively. If B's utilities are such that $u(sd) > u(cd)$, then B prefers sd to cd. Axiomatizations by Savage (1954) and von Neumann and Morgenstern (1953) provide persuasive postulates implying the existence of utilities such that maximization of expected utility (where expectation is taken with respect to the agent's subjective probabilities) is the optimal decision procedure. If two utility functions $\mathbf{u}$ and $\mathbf{u}'$ are positive linear transformations of one another, then they are considered strategically equivalent, since maximizing expected utility leads to the same choice in both cases.

In general, utility is defined over the cross product of available actions and possible states. In most of the dissertation, utility is restricted to be *state-independent* and arise from an underlying utility for *money*. If B's utility for μ dollars is $u(\mu)$, its utility U for a particular action a is its expected utility for money,

$$U(a) = E\left[u\left(\Upsilon^{\langle a,\omega\rangle}\right)\right] = \sum_{\omega\in\Omega} \Pr(\omega) u\left(\Upsilon^{\langle a,\omega\rangle}\right), \tag{2.2}$$

where $\Upsilon^{\langle a,\omega\rangle}$ is B's wealth when action a is taken in state ω. B's *decisions* are made by maximizing expected utility, or choosing the action a that maximizes (2.2).

I assume throughout that utility increases monotonically with wealth. *Local risk aversion* at μ, denoted $r(\mu)$, is defined as $r(\mu) \equiv -u''(\mu)/u'(\mu)$. B is *risk-averse* if $r(\mu) > 0$ for all μ, or, equivalently, if u is everywhere concave. Under this condition, B always prefers a guaranteed payment equal to the expected value of a lottery rather than the lottery itself, thus exhibiting an "aversion" to gambling. B is risk-neutral if $r(\mu) = 0$ for all μ, or u is linear; in this case, maximizing (2.2) coincides with maximizing expected payoff.

Computer programs face decisions under uncertainty as well—if not usually couched in such anthropomorphic terms—whenever limited resources (e.g., CPU time or network bandwidth) must be allocated and system state is not fully observable or not completely predictable. For example, a browser might prefetch web pages that are deemed likely to be requested, or an operating system might choose how to allot otherwise idle CPU time (Horvitz, 1997; Horvitz, 1999). In other cases, the decision-making analogy is quite literal. For example, medical expert systems can diagnose the probabilities of diseases and recommend treatments (Heckerman, 1991; Ng and Abramson, 1994). Researchers have developed compact data structures to record subjective probabilities and utilities, and efficient algorithms to compute probabilistic queries and return optimal decisions. A key enabling technology is the Bayesian network (BN), surveyed in Section 2.3.4 below.

2.1.5 Observability and Risk-Neutral Probability

Notice that an outside observer O, privy only to B's chosen actions, cannot uniquely discern either B's belief or its utility; the two quantities are inextricably linked (Kadane and Winkler, 1988). Any one of a continuous family of belief–utility pairs offers an equally valid rationalization for B's actions. That is, for any function $f(\omega)$, subjective probabilities proportional to $\Pr(\omega)f(\omega)$ matched with utilities $u\left(\Upsilon^{\langle a,\omega\rangle}\right)/f(\omega)$ result in strategically equivalent utilities for actions $U(a)$.

Risk-neutral probabilities are defined as

$$\Pr^{\mathrm{RN}}(\omega) \propto \Pr(\omega)u'\left(\Upsilon^{\langle a,\omega\rangle}\right), \tag{2.3}$$

where u' is the derivative of utility (Nau, 1995). B's observable behavior, manifested as actions, is indistinguishable from that of a hypothetical agent with transformed probabilities $\Pr^{\mathrm{RN}}(\omega)$ and reciprocally transformed utility $u^{\mathrm{RN}}(\mu) \equiv u(\mu)/u'\left(\Upsilon^{\langle a,\omega\rangle}\right)$. The observer *can* uniquely assess B's risk-neutral probability. In fact, all standard elicitation procedures designed to reveal B's beliefs based on monetary incentives (de Finetti, 1974; Winkler, 1968)—for example, querying the prices at which B would buy or sell various lottery tickets—essentially reveal $\Pr^{\mathrm{RN}}$, and *not* Pr (Kadane and Winkler, 1988). The agent's *observable* beliefs are in effect its risk neutral probabilities, not its true probabilities.

2.1.6 Group Coordination: A History of Paradox and Impossibility

I use the term *group coordination* to encompass all manner of multiagent scenarios, from cooperative to competitive, involving belief and/or decision, and as examined across many disciplines, including social choice, game theory, auction theory, statistics, decision analysis, and AI. Common to all formal treatments is a hobbling and unsettling number of paradoxes and impossibility theorems.

Much of social choice theory is concerned with aggregating individual preferences into a societal preference order, usually by voting. In the late 1700s, Borda formalized a so-called *voting paradox*, later generalized by Saari (1995). Suppose that a set of n candidates are ranked by *plurality* vote (the standard one-person, one-vote method, where the candidate with the most votes wins). Now consider whether the outcome might have changed if the least popular candidate had dropped out before the election. Saari shows that, depending on the agents' preferences, when the bottom candidate is removed any one of the others, including the second from bottom, might win in an $(n-1)$-way race. In fact, the ranking of the $n-1$ remaining candidates might follow *any* of the 2^{n-1} possible permutations. Also in the late 1700s, Condorcet illustrated another seeming paradox. It is possible that a group of agents elects a particular candidate by plurality vote, even though the same candidate loses to *every* other candidate in pairwise comparisons assessed by majority vote.

In his startling and influential *Impossibility Theorem*, Arrow (1963, 1967) generalized these examples into a sweeping, mostly negative result. He proved that aggregating preferences is in fact *impossible*, if the combination procedure is to satisfy a few compelling and rather innocuous-looking properties. Sen (1986) provides an excellent survey of the massive literature spun from Arrow's seminal work. Researchers have since extended Arrow's theorem to the case of combining utilities. In general, economists argue that the absolute magnitude of utilities are not comparable between individuals, since (among other reasons) utilities are invariant under positive affine transformations. In this context, Arrow's theorem on preference aggregation applies to the case of combining utilities as well (Fishburn, 1987; Sen, 1986). Countless attempts at weakening one or another of Arrow's axioms in pursuit of a satisfying back door—two notable examples being Fishburn's *infinite societies* (1970) and Black's *single-peakedness* (1987)—all corroborate a conclusion that Arrow's result is remarkably stable. Sen (1970) details another forceful impossibility theorem based on an alternative axiomatization.

A voting scheme is *strategy-proof* if no agent has incentive to misrepresent its preferences, regardless of how others vote. The *only* strategy-proof voting schemes are: 1. always adopt the same agent's preferences (a *dictatorship*); 2. choose one agent at random and adopt its preferences; 3. choose two candidates at random and decide between them by majority vote; and 4. some probabilistic mixture of 2 and 3 (e.g., flip a coin and implement 2 if heads and 3 if tails) (Gibbard, 1973; Gibbard, 1977; Satterthwaite, 1975). There is a one-to-one correspondence between strategy-proof voting schemes and Arrow-consistent preference aggregation functions (Satterthwaite, 1975). Under certain voting schemes, although agents may have an incentive to strategize, actually *computing* a beneficial strategy is NP-hard (Bartholdi et al., 1989b; Bartholdi et al., 1989a; Bartholdi and Orlin, 1991; Bartholdi et al., 1992).

Hylland and Zeckhauser (1979) prove an impossibility theorem regarding *group decision making*. The authors derive strong limitations on any procedure that separately aggregates individual beliefs into a group belief and utilities into a group utility, and chooses the action that maximizes group expected utility. The *only* such procedure that simultaneously maintains unanimous agreement on beliefs and unanimous agreement on decisions is a dictatorship. Other results evince an equally grim outlook for the prospects of a satisfactory theory for group decisions along the lines of that for individuals (Mongin, 1995; Schervish et al., 1991; Seidenfeld et al., 1989; Seidenfeld and Schervish, 1991). Even ignoring utilities and decisions, the aggregation of belief is itself associated with severe impossibility theorems, surveyed below in Section 2.3.1.

The occurrence of event A is *common knowledge* among a group if every member knows that A has occurred, knows that everyone else knows that A has occurred, knows that everyone else knows that everyone else knows that A has occurred, ad infinitum. Common knowledge is instrumental for ensuring coordination of actions (Fagin et al., 1996) and is analogous to a fundamental concept termed *consensus* in the distributed computing literature (Turek and Shasha, 1992). If communication is at all unreliable, then a group can *never* transition to a state of common knowledge (Fagin et al., 1996; Turek and Shasha, 1992).

Other formalisms of multiagent interaction are littered with counterintuitive and disturbingly restrictive results, including auction theory (Myerson and Satterthwaite, 1983). The implications of such impossibility theorems extend to domains directly relevant to computer science and AI, including distributed computing (Turek and Shasha, 1992), engineering design (Hazelrigg, 1996), default logic (Doyle and Wellman, 1991), multiagent bargaining (Sandholm and Vulkan, 1999), combining Bayesian networks (Chapter 5) and

(Pennock and Wellman, 1999), and collaborative filtering (Pennock and Horvitz, 1999).

The ubiquity of problems involving group coordination crosses disciplinary boundaries and underlies a massive body of literature devoted to the topic; this review hardly scratches the surface. See Chapter 8 of Raiffa's superb textbook (1968) for an enlightening and accessible examination of some of the models and the agonizing tradeoffs required therein.

2.2 Motivations

2.2.1 Why aggregate beliefs?

If we are convinced by any of the formal justifications of either Savage (1954), von Neumann and Morgenstern (1953), de Finetti (1974), or Cox (1946), and if we desire to build rational agents, then those agents should possess—or at least act as if they possess—a subjective probability distribution.

A group held to the same standard demands the existence of a consensus probability distribution. The axioms, difficult to refute for individuals, are certainly not infallible in the context of a group. At bottom, most agree that it is individuals only that can hold beliefs and make decisions. And perhaps those individuals are unconcerned if the groups in which they are members cannot manage to avoid sure losses. Yet, for agents with even the slightest cohesion, it seems eminently reasonable to favor an increase in one's utility, provided it means no detriment to any other. Conjured over all agents, this is precisely the criterion of *Pareto optimality*, and is satisfied only if arbitrage is not possible. A decidedly secondary motivation is theoretical elegance. A unified theory of decision, applicable to both individuals and groups, would admit treatment of institutions (committees, companies, nations) as individuals, or individuals as a collection of disagreeing "inner-selves", as in Minsky's *society of mind* (1986).

Even disregarding any axiomatic justification, motivation arises as well out of practical necessity. An agent cannot realistically gather all available evidence bearing on its decision; the cost is simply prohibitive. More reasonable might be to learn from the opinions of others, either by explicit elicitation or by observation. Such is the standard methodology of the so-called *knowledge engineer* engaged in the design of *expert systems* (Russell and Norvig, 1995), including those based on probabilistic models (Neapolitan, 1990). For this designer, decisions about how to aggregate beliefs are almost unavoidable. In many situations—for example, when the modeled events encompass a domain broader than any one expert's specialty—more than one source is consulted for probabilities. In practice, a simple average might be employed to coalesce assessments, though the implications of this choice can be surprisingly severe, as we see in Section 2.3.1 and Chapter 5. Expert systems are notorious for their narrow scope and ungraceful degradation as queries diverge from the anticipated domain. Potential solutions include embarking on much more ambitious and costly projects (Lenat and Guha, 1990) and/or *merging* several existing systems into one with wider coverage. Beyond these obvious occasions for employing aggregation, every designer must face a more subtle consideration. Even when consulting only one expert to construct only one model, his own beliefs inevitably become entangled with those of the expert. No designer can act as a completely transparent, bias-free conduit to transcribe knowledge from an expert to a model. For example, the designer may choose to correct for typical biases of those unfamiliar with probability theory—in fact, choosing *not* to correct for bias itself may distort the expert's true beliefs. It seems then that only a system built solely to represent one's own belief is immune from some attention toward methods of aggregation.

Rather than interrogating experts, an agent might *induce* external opinions from some publicly observable quantity—for example, the price of an insurance contract. In Section 4.1, we explore the similarity between learning from the opinions of others and learning from prices.

2.2.2 Why a market approach?

Arguing the value of *some* procedure for aggregating beliefs is not difficult; isolating the *best* method has proved much more elusive. Mirroring much work in group coordination, various impossibility theorems (Dalkey, 1972; Dalkey, 1975; Genest, 1984b; Genest and Wagner, 1987; Lehrer and Wagner, 1983; Wagner, 1984) and definitional controversies (French, 1985) shade prospects for an entirely agreeable solution.

Typical justifications for opinion pools are axiomatic or *top-down*, deriving the aggregation function entailed by a set of desirable criteria. Little attention is paid to the motivating incentives of the agents, and whether they assent to the chosen criteria or procedure. The conceptual framework is that of a central coordinator responsible for gathering, computing, and reporting the consensus. However, among self-motivated and decentralized agents, no coordinator would generally be trusted, and any potential coordinator may not desire to expend the resources necessary to perform what could be an expensive calculation. Moreover, the cost of transmitting joint distributions to a central location may be prohibitive.

In contrast, the market approach developed in this thesis is largely *bottom-up.* The consensual beliefs arise from the interaction of inherently self-interested agents, and incentives are purposefully aligned to reward participation and honesty. Agents are expected only to solve their own decision problem, and are not reliant upon, nor responsible for, the decisions of any others. By not requiring explicit cooperation, the market is likely to provide a more acceptable setting for aggregation in a wider variety of multiagent scenarios. Auction sites for each security can be distributed, and agent-to-agent communication is not required. Agents need only send messages to auctions of interest, and the timing can be asynchronous. Trading protocols are well documented, in current widespread use, and easily transferable to computer implementation. Economic theory provides a library of mathematical tools that are specifically tailored for multiagent settings that can contribute to our understanding of the aggregation process.

2.3 Background

This section covers prerequisite background material for subsequent chapters. All chapters depend to some degree on the basic concepts of belief aggregation surveyed in Section 2.3.1. The material on general equilibrium economics in Section 2.3.2 is directly required only for Chapter 6. Chapters 3, 4, and 7 all rest on the foundations of the theory of securities markets presented in Section 2.3.3. Graphical models for representing probability distributions, including Bayesian networks (BNs), are reviewed in Section 2.3.4, and revisited later in Chapters 5, 6, and 7. In addition, extensions in Chapter 4 build upon the foundation of Chapter 3, and results in Chapter 7 rest on derivations from both Chapters 3 and 5.

2.3.1 Opinion Pools

A variety of authors have proposed or advocated a corresponding variety of aggregation functions of the form (2.1); Genest and Zidek (1986) and French (1985) provide comprehensive surveys. Three approaches are generally distinguished. The first assumes a single Bayesian decision maker h (real or fictitious, within or outside the group), called the *supra Bayesian,* with a joint distribution over all events and all participants' beliefs. The supra Bayesian updates its beliefs via Bayes's rule, given the "evidence" of everyone else's beliefs. The resulting posterior is taken to be the consensus belief,

$$\Pr_0(\omega|\Pr_1,\ldots,\Pr_N) \propto \Pr_h(\Pr_1,\ldots,\Pr_N|\omega)\Pr_h(\omega). \tag{2.4}$$

The second approach is to apply a prespecified function, usually some form of weighted average, that maps any set of probability distributions to a singleton. Note that some such functions can be interpreted as the updating procedure of a supra Bayesian. Each pooling function is usually justified axiomatically, by assuming a "reasonable" set of properties of the aggregate distribution. The two most common and well-studied aggregation functions are the linear and logarithmic opinion pools (LinOP, LogOP). The LinOP is a weighted arithmetic mean of the members' probabilities,

$$\boxed{\Pr_0(\omega) = \sum_{i=1}^{N} \alpha_i \Pr_i(\omega),} \tag{2.5}$$

and the LogOP is a normalized, weighted geometric mean,

$$\boxed{\Pr_0(\omega) \propto \prod_{i=1}^{N} [\Pr_i(\omega)]^{\alpha_i},} \tag{2.6}$$

where the α_i are called *expert weights*, usually nonnegative numbers that sum to one. The LinOP and LogOP can actually be characterized as two instances of a parameterized family of weighted aggregation functions (Cooke, 1991). A third, relatively new, approach to pooling opinions is based on maximum entropy inference. The consensus is the unique probability distribution that maximizes Shannon entropy, chosen from among the distributions that are consistent with all available information, including the experts' beliefs, their past performance, and/or dependencies among experts (Levy and Delic, 1994; Myung et al., 1996).

The debate over which aggregation method is best continues to rage (Benediktsson and Swain, 1992; Cooke, 1991; Jacobs, 1995; Ng and Abramson, 1992; Winkler, 1986). Several authors (most emphatically Lindley (1985, 1988)) argue that the supra Bayesian approach is superior, as it is grounded in standard normative Bayesian theory (Clemen and Winkler, 1993; Morris, 1974; Morris, 1977; Rosenblueth and Ordaz, 1992; West and Crosse, 1992; Winkler, 1981). However, several conceptual questions arise in characterizing the *single* supra Bayesian responsible for the consensus of a (possibly leaderless) group (Genest and Zidek, 1986). If the supra Bayesian is real, is it a member of the group or an outsider? If an outsider, will the group accept its authority to generate the consensus? If a member, which one? If it is fictitious, where do its prior probabilities come from? These must somehow be agreed upon in advance, but then how is *that* consensus reached? A second problem is analytic tractability—the supra Bayesian must have a joint distribution over all members' beliefs and the events in question. In almost every treatment of this method, strong assumptions are made for the supra Bayesian's priors and likelihood function—for example, by assuming that the members' log-odds are normally distributed and independent—in order to yield tractable results (Clemen and Winkler, 1993; Jacobs, 1995; Lindley, 1988; Rosenblueth and Ordaz, 1992; Winkler, 1981).

Attempts to justify more symmetric opinion pools often proceed by posing axioms on the combination function, and arguing that they represent desirable properties (Dalkey, 1975; Genest, 1984c; Genest, 1984b; Genest, 1984a; Genest and Zidek, 1986; Genest et al., 1986; Genest and Wagner, 1987; Wagner, 1984). Many of these properties seem reasonable, but disagreement persists on which are essential.[1] Researchers have proved that certain pooling formulae are implied by certain sets of properties. We begin with two seemingly incontrovertible assumptions.

Property 2.1 (Unanimity (UNAM)) *If* $\Pr_h(\omega) = \Pr_i(\omega)$ *for all agents h and i, and for all states* $\omega \in \Omega$, *then* $\Pr_0(\omega) = \Pr_1(\omega)$.

Property 2.2 (Nondictatorship (ND)) *There is no single agent i such that* $\Pr_0(\omega) = \Pr_i(\omega)$ *for all* $\omega \in \Omega$, *and regardless of the agents' beliefs.*

UNAM states that if everyone's assessments are in complete agreement, then the consensus agrees as well. ND simply ensures that what is inherently a multiagent problem is not reduced to the single-agent case.

Property 2.3 (Marginalization property (MP)) *Let E be an* arbitrary *event, that is, any subset of* Ω. *Then,*

$$f(\Pr_1, \Pr_2, \ldots, \Pr_n)(E) = \\ f(\Pr_1(E), \Pr_2(E), \ldots, \Pr_n(E)).$$

[1] For example, Lindley (1985) regards the so-called *marginalization property* as an "adhockery" while Cooke (1991) characterizes any consensus function that does *not* respect it as "downright queer".

Property 2.4 (External Bayesianity (EB)) *Let E and F be arbitrary events. Then,*

$$f(\mathrm{Pr}_1, \mathrm{Pr}_2, \ldots, \mathrm{Pr}_n)(E|F) = \\ f(\mathrm{Pr}_1|F, \mathrm{Pr}_2|F, \ldots, \mathrm{Pr}_n|F)(E).$$

MP and EB require consistency for probabilistic operations performed before and after pooling. MP states that we obtain the same probability for an event E whether we pool the opinions first, and then compute $\mathrm{Pr}_0(E) = \sum_{\omega \in E} \mathrm{Pr}_0(\omega)$, or if we first compute $\mathrm{Pr}_i(E) = \sum_{\omega \in E} \mathrm{Pr}_i(\omega)$ for each agent i, and then pool their opinions only over E. Similarly, EB holds that we obtain the same $\mathrm{Pr}_0(E|F)$ whether we combine opinions first and condition on F second, or condition on F first and combine opinions second. It has been shown that any f satisfying both MP and UNAM is a LinOP (Genest, 1984c), and any satisfying EB and UNAM is a LogOP (Genest, 1984a). Genest (1984b) also shows that f cannot simultaneously satisfy MP, EB, UNAM, and ND.

Property 2.5 (Proportional dependence on states (PDS))

$$\mathrm{Pr}_0(\omega) \propto f(\mathrm{Pr}_1(\omega), \mathrm{Pr}_2(\omega), \ldots, \mathrm{Pr}_n(\omega)).$$

PDS is sometimes called *independence of irrelevant states*, or termed a *likelihood principle*. It assures that the consensus likelihood ratio between two states does not depend on the agents' assessments of any other "irrelevant" state. The LinOP, LogOP, and most other proposed opinion pools satisfy PDS.

Property 2.6 (Independence preservation property (IPP)) *Let E and F be* arbitrary *events. If $\mathrm{Pr}_i(E|F) = \mathrm{Pr}_i(E)$ for all agents i, then $\mathrm{Pr}_0(E|F) = \mathrm{Pr}_0(E)$.*

IPP requires that *all* unanimously held independencies are preserved in the consensus. Advocates of IPP reason that identifying the independencies in a model is central to understanding the underlying phenomena, and that complete agreement on this dimension should be embraced. On the other hand, Genest and Wagner (1987) make a compelling case *against* the use of IPP by proving that *no* aggregation function whatsoever can satisfy it along with PDS and ND, when $|\Omega| \geq 5$.

Another point of contention is how best to determine the expert weights. In most cases they are chosen in an ad hoc manner to encode some measure of confidence, reliability, or importance (Benediktsson and Swain, 1992; French, 1985; Winkler, 1968). Some more formal methods to derive weights have been proposed, by making assumptions concerning the form of, or interdependence among, participants' beliefs (Cooke, 1991; Degroot and Mortera, 1991; Jacobs, 1995; Morris, 1977), or through iterative self-weighting procedures (Degroot, 1974). Little work, in any of the traditional categories of opinion pools, explicitly addresses truth incentives for reporting either self-assessed weights or the probabilities themselves.

2.3.2 General Equilibrium Economics Under Certainty

A *market price system* is defined by a set of S *goods* and a *price vector* $\mathbf{p} = \langle p^{\langle 1 \rangle}, \ldots, p^{\langle S \rangle} \rangle$ that associates a price with each good. The system requires that all goods be exchanged in proportion to their relative prices.

Goods are exchanged by two types of agents—consumers and producers. *Consumer* agents receive value from direct consumption of the goods (metaphorically, they "eat" the goods). Let $\mathbf{x} = \langle x^{\langle 1 \rangle}, \ldots, x^{\langle S \rangle} \rangle$ denote a *consumption bundle* where each $x^{\langle j \rangle} \in \Re_+$ specifies the quantity of good j consumed. The consumption bundles are ranked according to preference by the consumer's utility function $U(\mathbf{x}) : \Re_+^S \to \Re$. Consumers also start with an initial allocation of the goods, termed their *e*ndowment and denoted by $\mathbf{e} = \langle e^{\langle 1 \rangle}, \ldots, e^{\langle S \rangle} \rangle$. The consumer's objective is to choose an affordable bundle of goods, $\mathbf{x}$, so as to maximize its utility. A bundle is affordable if its total cost at the going prices does not exceed the value of the consumer's endowment at the same prices. The consumer's choice can thus be expressed as the following constrained optimization problem:

$$\max_{\mathbf{x}} U(\mathbf{x}) \text{ s.t. } \mathbf{p} \cdot \mathbf{x} \leq \mathbf{p} \cdot \mathbf{e}. \tag{2.7}$$

Agents of the second type, *producers*, extract value from goods by transforming them into other goods, and selling their product in the market. A producer's ability to transform goods is defined by its *technology*, $Y \subset \Re^S$, which specifies the set of feasible production vectors. If $\mathbf{y} = \langle y^{\langle 1 \rangle}, \ldots, y^{\langle S \rangle} \rangle$ is feasible, then the producer is capable of transforming bundles of *input* goods (goods j for which $y^{\langle j \rangle} < 0$) into bundles of *output* goods ($y^{\langle j \rangle} > 0$), in respective amounts $|y^{\langle j \rangle}|$.

Unlike the consumer, a producer has no preferences in the sense of agent-specific desires. Rather, we assume that a producer selects its production activity solely according to *profit*—the difference between the value of its output and the cost of its input, evaluated at a given set of prices. The producer's constrained optimization problem can be expressed succinctly as

$$\max_{\mathbf{y}} \mathbf{p} \cdot \mathbf{y} \text{ s.t. } \mathbf{y} \in Y. \tag{2.8}$$

An agent is *competitive* if it takes prices as given, ignoring the potential effect of its own choices on resulting prices. The agent definitions above assume competitive behavior, in that the prices are treated as parameters of the respective optimization problems. Note that only relative prices matter; behavior is unchanged if all prices are multiplied by a positive constant. We typically scale prices by designating one good (the first, without loss of generality) as *numeraire*, with a fixed price, $p^{\langle 1 \rangle} = 1$.

Consider an economy with consumers indexed $1, \ldots, \eta$ and producers indexed $\eta + 1, \ldots, N$. A *competitive equilibrium* for this economy is a set of prices, $\mathbf{p}$, such that all of the goods are in material balance,

$$\sum_{i=1}^{\eta} \mathbf{x}_i(\mathbf{p}) \leq \sum_{i=\eta+1}^{N} \mathbf{y}_i(\mathbf{p}) + \sum_{i=1}^{\eta} \mathbf{e}_i,$$

where $\mathbf{x}_i(\mathbf{p})$ and $\mathbf{y}_i(\mathbf{p})$ denote the solutions of consumer or producer i's respective optimization problem, as defined above, at prices $\mathbf{p}$.

2.3.3 General Equilibrium Economics under Uncertainty: Securities Markets and Rational Expectations

Under uncertainty, endowments and utilities are in general state-dependent, and thus so are the spot prices of goods. Agents will generally exhibit varying marginal value of wealth in the different states. Consequently, if they are risk-averse, they will desire to *hedge* or *insure* against risk by distributing wealth across states.[2] The *Arrow-Debreu securities market* is the fundamental theoretical framework in economics and finance for resource allocation under uncertainty (Arrow, 1964; Drèze, 1987; Mas-Colell et al., 1995). A *security*, denominated in money or other exchangeable good, pays off variously contingent upon the realization of an uncertain state. Let $\langle A \rangle$ denote a security that pays off if and only if the event A occurs. In the standard model, agents trade securities prior to

[2] For example, insuring a package delivery transfers wealth from the *package-received* state to the *package-lost* state.

revelation of the world state, and then trade real goods once the state is known. A market is termed *complete* if it contains at least $|\Omega| - 1$ *linearly independent* securities, plus spot markets in the original goods. Such a market guarantees, under classical assumptions, that equilibrium entails a *Pareto optimal*, or *efficient*, allocation of both goods and risk.

A *conditional security* $\langle A_1|A_2\rangle$ pays off *contingent* on A_1 and *conditional* on A_2. That is, if A_2 occurs, then it pays out exactly as $\langle A_1\rangle$; on the other hand, if $\bar{A}_2$ occurs, then the bet is called off and any price paid for the security is refunded (de Finetti, 1974). The canonical complete market consists of one security paying out in each state of nature. In general, though, any set of securities—included those composed of conditional securities—whose payoff-by-state matrix has rank $|\Omega| - 1$ is complete.

Prices as probabilities

When one unit of each security pays out one dollar, it follows, from de Finetti's argument, that the equilibrium prices in a securities market form a coherent probability distribution.[3] In fact, these prices coincide with the agents' risk-neutral probabilities (2.3) for the available securities, which are in agreement at equilibrium (Drèze, 1987; Nau and McCardle, 1991). Derived formally in Section 7.1.1, I simply sketch the intuition here. Since a risk-neutral agent buys $\langle A_j\rangle$ if $p^{\langle A_j\rangle} < \Pr(A_j)$ (it simply maximizes expected payoff), then *any* agent buys $\langle A_j\rangle$ if $p^{\langle A_j\rangle} < \Pr^{\mathrm{RN}}(A_j)$. Similarly, the agent sells if $p^{\langle A_j\rangle} > \Pr^{\mathrm{RN}}(A_j)$. If two agents h and i have differing risk neutral probabilities—that is, $\Pr_h^{\mathrm{RN}}(A_j) \neq \Pr_i^{\mathrm{RN}}(A_j)$—then there is an intermediate price at which they both desire to trade. It follows that, at equilibrium, when by definition opportunities for exchange have been exhausted, all agents' risk neutral probabilities agree across available securities. Furthermore, since offers to buy and sell must match, the equilibrium prices equal these consensus probabilities.

There are two, largely inseparable, reasons for agents to trade in securities: to insure against risk ("hedge") and to profit from perceived mispricings ("speculate"). The more averse to risk, the more the former consideration dominates an agent's decision making. On the other hand, risk-neutrality—the limit of diminishing risk aversion—is synonymous with pure speculation. These two behaviors are aligned with the two central roles of securities markets in the theory of economics under uncertainty. The first, as mentioned, is to support the reallocation of risk. The second is to *aggregate* and *disseminate* information. Agents that disagree on the likelihood of states may seek to exchange securities at prices that yield, according to each's subjective viewpoint, an increase in expected re-

[3] For example, $p^{\langle A_1\rangle} = p^{\langle A_1 A_2\rangle} + p^{\langle A_1 \bar{A}_2\rangle}$ or $p^{\langle A_1 A_2\rangle} = p^{\langle A_1|A_2\rangle} p^{\langle A_2\rangle}$.

turns. Moreover, each agent is privy, albeit implicitly, to the evidence gathered by other agents (perhaps at great cost) via fluctuations in price.

Indeed, a common interpretation of the prices of securities, or similar lottery instruments, is as an amalgamation of opinions. For example, the price of a stock represents the "market evaluation" of the expected present value of future dividends, and odds in a horse race summarize the bettors' beliefs about the winning horse's identity. Under certain regularity conditions, the exact dependence of prices on beliefs can be characterized (Huang and Litzenberger, 1988; Rubinstein, 1974; Rubinstein, 1976; Wilson, 1968), and compared with the opinion pools of Section 2.3.1. In contrast to the Bayesian or axiomatic approaches, economic models are typically built upon behavioral assumptions and descriptions of the mechanism for exchange.

Early derivations of pricing formulae centered on the *prior information* (PI) equilibrium concept, where agents' beliefs are presumed to remain fixed during price formation (Huang and Litzenberger, 1988; Rubinstein, 1974; Rubinstein, 1976). As discussed in Section 3.3, prices can in this case be characterized as a combination of the agents' beliefs, but cannot reasonably be regarded as an accumulation of diverse *information*. The latter, perhaps more satisfying, interpretation does become admissible under the *rational expectations* (RE) equilibrium assumption (Grossman, 1981; Lucas, 1972). This theory asserts that prices actually reflect all of the information available to all agents. Even when some agents have exclusive access to inside information, prices equilibrate exactly as if everyone had access to all information. The procedural explanation is that prices *reveal* to the ignorant agents any initially private information; that is, agents *learn* from prices. Several authors show that, if agents begin with identical priors and disparate evidence, repeated observation of some aggregate statistic (e.g., price) will converge to a consensus on posteriors, for various sufficient statistics (Hanson, 1998a; McKelvey and Page, 1986; McKelvey and Page, 1990; Nielsen et al., 1990). I describe a supra Bayesian model of RE in Section 4.1.

Both Rubinstein (1974) and Wilson (1968) develop formal conditions, examined further in Section 3.3.3, under which agents in a securities market behave collectively in a manner indistinguishable from a single, *composite* agent whose beliefs equal the equilibrium prices (Huang and Litzenberger, 1988). Raiffa (1968) (pp. 211–216) recounts an example, due to Pratt, where the optimal group action requires randomization, and thus a composite agent *cannot* exist. Nevertheless, Kasanen and Trigeorgis (1994) demonstrate empirically that the analogy of a market as a representative investor can sometimes apply, even outside the theoretical bounds. The composite-agent construction constitutes perhaps the most compelling interpretation for prices as consensus beliefs.

2.3.4 Graphical Models

Graphical models (Darroch et al., 1980) have proved invaluable as a language for encoding probability distributions, due primarily to their facility in expressing *independencies* among variables. Sufficient independencies yield not only parsimonious models, but insights into the underlying system being represented; conversely, the lack of any independence structure renders a graphical description, per se, essentially devoid of content. Let $\mathrm{CI}[A_j, W, X]$ be shorthand for $\Pr(A_j|WX) = \Pr(A_j|W)$, indicating that A_j is conditionally independent of the set of events X, given another set W.

In a *Bayesian network* (BN), conciseness is achieved by exploiting conditional independence among the primary events. Consider the event $A_k \in Z$, with predecessors $\mathbf{pred}(A_j) \equiv \{A_1, A_2, \ldots, A_{k-1}\}$. Suppose that, given the outcomes of a subset $\mathbf{pa}(A_k) \subseteq \mathbf{pred}(A_k)$ of its predecessors—called A_k's *parents*—the event A_k is conditionally independent of all other preceding events, or $\mathrm{CI}[A_k, \mathbf{pa}(A_k), \mathbf{pred}(A_k) - \mathbf{pa}(A_k)]$. This structure can be depicted graphically as a *directed acyclic graph* (DAG): each event is a node in the graph, and there is a directed edge from node A_j to node A_k if and only if A_j is a parent of A_k. We also refer to A_k as the *child* of A_j, and $A_k \cup \mathbf{pa}(A_k)$ as the *family* of A_k. A DAG has no directed cycles and thus defines a partial order over its vertices. We assume without loss of generality that the event indecies are consistent with this partial ordering; in other words, if A_j is a predecessor of A_k then $j < k$. We can write the joint probability distribution in a (usually) more compact form:

$$\Pr(A_1 A_2 \cdots A_M) = \prod_{k=1}^{S} \Pr(A_k|\mathbf{pa}(A_k)).$$

For each event A_k, we record a *conditional probability table* (CPT), which contains probabilities $\Pr(A_k|\mathbf{pa}(A_k))$ for all possible combinations of outcomes of events in $\mathbf{pa}(A_k)$. Thus it is possible to implicitly represent the full joint with $O\left(M \cdot 2^{\max\{q(k)\}}\right)$ probabilities, instead of $2^M - 1$, where $q(k) = |\mathbf{pa}(A_k)|$ is number of parents of A_k.

A *polytree* is a DAG with no undirected cycles. A *tree* is a polytree where every node has at most one parent.

A *Markov network* (MN) is another graphical language for modeling conditional independence and for implicitly describing a joint distribution (Whittaker, 1990; Darroch et al., 1980). Events are again associated with nodes in a graph, and edges encode probabilistic dependencies. The underlying structure of a MN is an *undirected* graph. Given the outcomes of its direct neighbors, an event A_j is conditionally independent of *every* other event in the network, not just preceding events. The neighbors of an event form a *Markov blanket* around it, "shielding" it from direct influence from the rest of the events (Pearl, 1988). We call the node A_j and the set of nodes $X \subseteq Z - A_j$ *Markov independent*, given another set $W \subseteq Z - X - A_j$, if $\mathrm{CI}[A_j, W, X]$ and $A_j \cup W \cup X = Z$. Thus a node is Markov independent of all other nodes, given its blanket. Encoding the joint distribution implicit in a MN involves assigning a *potential* probability to each clique (Neapolitan, 1990; Pearl, 1988).

The Markov blanket of a node in a BN consists of its direct parents, its direct children, and its children's direct parents (Pearl, 1988). Therefore a BN can be converted into a MN by *moralizing* the network, or fully connecting ("marrying") each node's parents, and dropping edge directionality (Lauritzen and Spiegelhalter, 1988; Neapolitan, 1990). A MN can be converted into a BN by *filling in* or *triangulating* (Kloks, 1994) the graph, and adding directionality according to the fill-in ordering (Jensen, 1996; Lauritzen and Spiegelhalter, 1988; Neapolitan, 1990; Pearl, 1988). Both transformations are sound with respect to independence, but neither is complete.

A graph (either directed or undirected) is an *independency map*, or an *I-map*, of a probability distribution Pr if every independency implicit in the graph (according to BN or MN semantics, respectively) holds within Pr. Note that a complete graph is a trivial I-map of any distribution over Ω.

BN *inference algorithms* compute arbitrary probabilistic queries, for example $\Pr(A_1|A_3\bar{A}_5)$. BN inference in polytree networks is polynomial-time. For general topologies, though, inference is #P-complete (Cooper, 1990), and even approximate inference is NP-hard (Dagum and Luby, 1993). Nevertheless, practical algorithms have been developed that take advantage of some of the network's independence structure. The most popular inference algorithms first convert the BN into a *join tree*, by adding arcs that moralize the network and triangulate the graph according to a *fill-in ordering* θ of the events (Lauritzen and Spiegelhalter, 1988; Jensen et al., 1990; Spiegelhalter et al., 1993; Neapolitan, 1990). Each added arc essentially doubles the size of a CPT. Determining the optimal ordering associated with the minimum triangulation is NP-complete; in practice, hueristics are employed. Though the join tree representation can be exponentially larger than the original BN, it can still be exponentially more compact than the full joint distribution.

Typical algorithms runs with time and space complexity $O\left(M \cdot 2^{\max\{q(j)\}}\right)$, where $q(j)$ is the number of neighbors of node A_j in the join tree.

A BN is *decomposable* (Chyu, 1991b; Darroch et al., 1980; Pearl, 1988; Shachter et al., 1991) if there is an edge between every two nodes that share a common child. Trees are a subset of decomposable DAGs, since every node has at most one parent. Complete graphs are also decomposable since *every* two nodes are connected. A BN can be made decomposable with respect to any ordering θ. First, *arc reversals* (Shachter, 1990; Shachter, 1988; Shachter, 1986) orient the DAG in a manner consistent with θ. Then, in reverse θ order, each node's parents are fully connected. The result is structurally equivalent to a join tree constructed by moralizing and triangulating according to θ, and supports inference with precisely the same time and space complexity (Chyu, 1991b; Chyu, 1991a; Pennock, 1997; Shachter et al., 1991).

Chapter 3

A Market Framework for Aggregating Beliefs

This chapter examines the use of a securities market as a device for pooling opinion. Agents bid on securities to maximize their own expected utility. The equilibrium prices form a probability distribution, interpretable as the group's consensus distribution. Section 3.1 introduces the model and Section 3.2 derives some general properties. Section 3.3 demonstrates that under each of two homogeneity assumptions regarding agents' utility types, equilibrium prices have the same functional form as a standard opinion pool. The so-called *expert weights*, notoriously hard to define in the context of opinion pools (French, 1985), correspond to parameters of the agents' utility functions. In both cases, the group's outward behavior is indistinguishable from that of a single, *composite* agent. Even when such a representative agent exists in the context of a *fixed* market configuration, Section 3.5 describes how an outside observer might differentiate between a group and an individual by introducing and withdrawing securities in a systematic way. The litmus test for a group is whether prices fail to satisfy either MP or EB, as defined in Section 2.3.1.

3.1 Market Setup

Section 2.3.3 reviewed the role of securities markets in classical economic theory to allocate risk and aggregate information and belief. In what follows, I develop a simplified market model tailored toward the central objective of aggregating beliefs.

The market consists of S securities "\$1 if A_j", abbreviated $\langle A_j \rangle$, for every event of interest A_j.[1] One unit of the security pays off one dollar if the event occurs, and nothing

[1]Throughout this chapter and the next, the number of securities S equals the number of events M. Separate notation is required in Chapters 6 and 7, whence the correspondence between securities and

otherwise. A quantity x of security $\langle A\rangle$ is a lottery resulting in \$$x$ if and only if A occurs. If an agent must pay a price p per unit of security $\langle A\rangle$, then purchasing x units is equivalent to accepting a lottery with payoff $(1-p)x$ if A occurs, and $-px$ otherwise. Denote this lottery as $L(x) = [\Pr(A), (1-p)x; \Pr(\bar{A}), -px]$. Positive x indicates a quantity to buy, and negative x a quantity to sell. *Either* action incurs risk: buying $\langle A\rangle$ at price p yields exactly the same lottery as selling $\langle \bar{A}\rangle$ at price $1-p$.

Agents are *competitive*; that is, they maximize expected utility at a given set of prices, or equivalently, ignore any effect their own demand may have on prices. Utility for money is independent of state and monotonically increasing. For most results we assume that agents are *risk averse* (Keeney and Raiffa, 1976; Pratt, 1964), that is, have concave u, always preferring the expected payoff of a lottery to the lottery itself. We also restrict attention to linearly independent securities (i.e., none can be replicated as a combination of others). Without this, agents could buy one security and sell the equivalent portfolio without risk, and demand may be nonunique or unbounded.[2] We consider the PI equilibrium condition in Sections 3.3 and 3.4, and a model of RE equilibirum in Section 4.1, respectively. Given the assumptions of competitive behavior and state-independent utility, we can define the agent's utility, U, for purchasing a quantity x of the security as its expected u,

$$U(x) = E[u(L(x))] = \Pr(A)u((1-p)x) + \Pr(\bar{A})u(-px).$$

In a market of S securities, let $x^{\langle j\rangle}$ represent a quantity of security $\langle A_j\rangle$, and $p^{\langle j\rangle}$ the security's price. The agent's utility for a bundle $\mathbf{x} = \langle x^{\langle 1\rangle}, x^{\langle 2\rangle}, \ldots, x^{\langle S\rangle}\rangle$ of securities can also be written as its expected u:

$$U(\mathbf{x}) = \sum_{\omega\in\Omega} \Pr(\omega) u\left(\Upsilon^{\langle\omega\rangle}\right), \tag{3.1}$$

where

$$\Upsilon^{\langle\omega\rangle} = \sum_{j=1}^{S} \left(1_{\omega\in A_j} - p^{\langle j\rangle}\right) x^{\langle j\rangle}$$

is the payoff in state ω, with $1_{\omega\in A_j}$ the indicator function that equals one if $\omega \in A_j$, zero otherwise.

The agent's *demand function*, $\mathbf{x}(\mathbf{p}) = \langle x^{\langle 1\rangle}(\mathbf{p}), x^{\langle 2\rangle}(\mathbf{p}), \ldots, x^{\langle S\rangle}(\mathbf{p})\rangle$, represents the quantities of securities that maximize utility (3.1) at prices $\mathbf{p} = \langle p^{\langle 1\rangle}, p^{\langle 2\rangle}, \ldots, p^{\langle S\rangle}\rangle$. Con-

events is no longer one-to-one.

[2]This assumption implies that each event is strictly uncertain ($0 < \Pr(A) < 1$).

$X = \{$ ... i's amount of a security − i's amount of a security in best response $\}$

cavity of u (risk aversion) entails concavity of U, and hence any critical point of (3.1) is a global maximum.

Equation 3.1 captures the optimization problem of a single agent faced with the decision to invest in securities. In an economy of N agents, each continually maximizing (3.1), prices adjust until all buy orders (positive $x^{\langle j \rangle}$) match with sell orders (negative $x^{\langle j \rangle}$) for each security $\langle A_j \rangle$. A market is in *competitive equilibrium* at prices $\mathbf{p}$ if and only if

$$\sum_{i=1}^{N} \mathbf{x}_i(\mathbf{p}) = \mathbf{0}. \tag{3.2}$$

The classical price formation mechanism is an iterative process called the *Walrasian auction* or *tatonnement* (Mas-Colell et al., 1995). Current prices $\mathbf{p}$ are publicly announced, and agents submit their demand functions $\mathbf{x}(\mathbf{p})$. Each price $p^{\langle j \rangle}$ is set to satisfy $\sum_i x_i^{\langle j \rangle} = 0$ and subsequently reannounced. The agents respond by submitting revised demand functions, and prices are again computed to locally balance supply and demand for each security. The process continues until a *fixed point* or equilibrium is reached; that is, all agents are optimizing and prices are stable. Under restrictions on the form of agent demand functions, the Walrasian mechanism is guaranteed to converge (Mas-Colell et al., 1995). [3]

Most opinion pools require that agents directly, and truthfully, reveal their probabilities to a central location. Scoring rules (Winkler, 1968), promissory notes (de Finetti, 1974), or other incentive devices could be leveraged to elicit honest assessments, though these generally require outside subsidies. Moreover, French (1985) points out that, when incented via scoring rules, a group of agents may obtain a greater total reward by systematically misrepresenting their beliefs; consequently, across-the-board truthfulness is *not* a Pareto optimal response for the group. Kadane and Winkler (1988) demonstrate that the probabilities revealed by monetary elicitation aids are inextricably confounded with utilities, whenever agents are risk averse and can own stakes in the events. The securities market framework provides built-in rewards for participation, honesty,[4] and the collection

[3] My model assumes that exchange occurs only after equilibrium is reached. In an alternative formulation, agents make incremental trades for securities, at mutually beneficial prices, until no such opportunities exist. Under my assumptions, such a process would also reach an equilibrium, where agents' risk-neutral probabilities are equal (and thus represent clearing prices). Solutions for equilibrium prices for the case of negative exponential utility (derived below) would remain the same, though my other analytic results are specific to the exchange-at-equilibrium formulation. The dynamic model of Section 4.2 reflects some aspects of the incremental approach.

[4] Incentives exist, though the market is not fully *incentive compatible* (Mas-Colell et al., 1995).

of cost-effective evidence. Agent beliefs are revealed only implicitly through their trades. In fact, if measures are taken to ensure the privacy of bids, the *only* personal information directly revealed is an agent's risk-neutral probability distribution at equilibrium, which equals the consensus anyway. Since supply and demand are balanced in equilibrium, the mechanism requires no subsidy, and so the only cost involved is that of organizing and running the market.

3.2 General Properties

There are two categories of properties to consider: those that deal with a single agent's demand for securities, and those that examine the equilibrium prices arising from group interaction. We begin with several propositions concerning the former.

Proposition 3.1 (Incentive to participate) *Any agent with* $\Pr(A_j) \neq p^{\langle j \rangle}$ *for some* j *has nonzero demand.*

Proof. Without loss of generality, assume $\Pr(A_1) \neq p^{\langle 1 \rangle}$. We show that the agent prefers $\mathbf{x} = \langle \epsilon, 0, \ldots, 0 \rangle$ to $\mathbf{x} = \mathbf{0}$, for some nonzero ϵ. For small enough ϵ, utility $U(\epsilon, 0, \ldots, 0)$ (3.1) can be approximated around $\mathbf{0}$ by a first-order Taylor expansion,

$$\begin{aligned} U(\epsilon, 0, \ldots, 0) &\approx U(\mathbf{0}) + \epsilon \left. \frac{\partial U}{\partial x^{\langle 1 \rangle}} \right|_{\mathbf{0}} \\ &= U(\mathbf{0}) + \epsilon \sum_{\omega \in \Omega} \Pr(\omega)(1_{\omega \in A_1} - p^{\langle 1 \rangle}) u'(0) \\ &= U(\mathbf{0}) + \epsilon u'(0)(\Pr(A_1) - p^{\langle 1 \rangle}). \end{aligned}$$

Utility for money is monotonically increasing, so $u'(0) > 0$. If $\Pr(A_1) > p^{\langle 1 \rangle}$, then $U(\epsilon, 0, \ldots, 0) > U(\mathbf{0})$ for positive ϵ. Similarly, if $\Pr(A_1) < p^{\langle 1 \rangle}$, then the agent prefers $\epsilon < 0$ to $\epsilon = 0$. □

Note that Proposition 3.1 does *not* assert that optimal demand includes positive or negative $x^{\langle j \rangle}$, only that $\mathbf{x} = \mathbf{0}$ is strictly dominated.

Proposition 3.2 (Bounded demand) *Any agent with utility for money such that either*

1. $u(\mu) = o(\mu)$ *and* $-u(-\mu) = \Omega(\mu)$, **or**

2. $u(\mu) = O(\mu)$ *and* $-u(-\mu) = \omega(\mu)$

has bounded demand if prices form coherent probabilities.[5]

Proof. Consider a nonzero, finite portfolio of securities $\delta = \langle \delta^{\langle 1 \rangle}, \delta^{\langle 2 \rangle}, \ldots, \delta^{\langle S \rangle} \rangle$. Let $\Delta_\omega \equiv \sum_{j=1}^{S} (1_{\omega \in A_j} - p^{\langle j \rangle}) \delta^{\langle j \rangle}$ be the payoff of this portfolio in outcome ω. Let Ω_+ be the set of outcomes with nonnegative payoff ($\Delta_\omega \geq 0$ for all $\omega \in \Omega_+$), and Ω_- the set of outcomes with strictly negative payoff. Prices are coherent (do not admit arbitrage) if and only if, for every portfolio δ, there is at least one negative-payoff outcome $\omega \in \Omega_-$ (Varian, 1987). The agent invests in $z > 0$ "shares" of this portfolio so that $\mathbf{x} = z\delta$. Utility is given by (3.1):

$$\begin{aligned} U(z\delta) &= \sum_{\omega \in \Omega} \Pr(\omega) u(z\Delta_\omega) \\ &= \sum_{\omega \in \Omega_+} \Pr(\omega) u(z\Delta_\omega) + \sum_{\omega \in \Omega_-} \Pr(\omega) u(z\Delta_\omega) \end{aligned}$$

If $-u(-\mu) = \Omega(\mu)$, then, from the definition of Ω-notation (Cormen et al., 1990), there exist positive constants l and $z*$ such that $-\Pr(\omega) u(z\Delta_{\omega_-}) \geq l\mu$ for all $z > z*$ and all $\omega_- \in \Omega_-$. If $u(\mu) = o(\mu)$, then $\lim_{z \to \infty} \Pr(\omega) u(z\Delta_{\omega_+}) / l\mu = 0$ for all $\omega_+ \in \Omega_+$. Combining these two results, we find that $\lim_{z \to \infty} u(z\Delta_{\omega_+}) / u(z\Delta_{\omega_-}) = 0$ for all $\omega_- \in \Omega_-$ and $\omega_+ \in \Omega_+$. Let $\omega_- \in \Omega_-$ be any negative-payoff outcome. Then,

$$\begin{aligned} \lim_{z \to \infty} U(z\delta) &= \lim_{z \to \infty} \sum_{\omega \in \Omega_+} \Pr(\omega) u(z\Delta_\omega) + \sum_{\omega \in \Omega_-} \Pr(\omega) u(z\Delta_\omega) \\ \lim_{z \to \infty} \frac{U(z\delta)}{u(z\Delta_{\omega_-})} &= \lim_{z \to \infty} \sum_{\omega \in \Omega_+} \Pr(\omega) \frac{u(z\Delta_\omega)}{u(z\Delta_{\omega_-})} + \sum_{\omega \in \Omega_-} \Pr(\omega) \frac{u(z\Delta_\omega)}{u(z\Delta_{\omega_-})} \\ \lim_{z \to \infty} U(z\delta) &= \lim_{z \to \infty} \sum_{\omega \in \Omega_-} \Pr(\omega) u(z\Delta_\omega). \end{aligned}$$

That is, in the limit as $z \to \infty$, the negative-payoff terms dominate U. Thus, regardless of the choice for δ, optimal z is bounded. An analogous argument applies if $u(\mu) = O(\mu)$ and $-u(-\mu) = \omega(\mu)$. □

Note that risk aversion implies both $u(\mu) = O(\mu)$ and $-u(-\mu) = \Omega(\mu)$, providing half of each sufficient condition, 1 and 2. Also note that both $u(\mu) = -e^{-c\mu}$ and $u(\mu) = \ln(\mu + b)$, considered in detail in Sections 3.3.1 and 3.3.2 below, satisfy the requisite

[5] Essentially, $o(\mu)$ means "asymptotically grows strictly slower than μ", $O(\mu)$ means "asymptotically grows slower than or proportional to μ", $\omega(\mu)$ (not to be confused with our notation for states) means "asymptotically grows strictly faster than μ", and $\Omega(\mu)$ means "asymptotically grows faster than or proportional to μ" (Cormen et al., 1990).

condition for this proposition.[6]

Lemma 3.3 (No negative expected payoffs) *A risk neutral or risk averse agent chooses demand* $\mathbf{x}$ *to yield nonnegative expected payoff; moreover, if* $\Pr(A_j) \neq p^{\langle j \rangle}$ *for some* j*, then he chooses demand with strictly positive expected payoff.*

Proof. A risk neutral agent, by definition, maximizes expected payoff. A risk averse agent would prefer zero demand (a lottery with zero certainty equivalent), to any nonzero (risky) demand that yields negative or zero expected payoff (a lottery with negative certainty equivalent). If $\Pr(A_j) \neq p^{\langle j \rangle}$ for some j, then Proposition 3.1 implies nonzero demand and thus strictly positive expected payoff. □

Proposition 3.4 (Qualitative partial demand) *A risk neutral or risk averse agent with* $\Pr(A_j) > (<)\ p^{\langle j \rangle}$ *for some* j*, and* $\Pr(A_k) = p^{\langle k \rangle}$ *for all* $k \neq j$ *has positive (negative) demand* $x^{\langle j \rangle}$ *for security* A_j*. In the boundary case, a risk averse agent with* $\Pr(A_j) = p^{\langle j \rangle}$ *has optimal demand* $\mathbf{x} = \mathbf{0}$*.*

Proof. Note that, regardless of his holdings of other securities, the agent's expected payoff is $(\Pr(A_j) - p^{\langle j \rangle})x^{\langle j \rangle}$. Then, by Lemma 3.3, any demand $\mathbf{x}$ must include $x^{\langle j \rangle} > (<)\ 0$ to ensure strictly positive expected payoff. If $\Pr(A_j) = p^{\langle j \rangle}$, *every* demand yields zero expected payoff; thus any nonzero demand for a risk averse agent is precluded by Lemma 3.3. □

Propositions 3.1 and 3.4, at first glance, may seem to violate the intuition of risk aversion. After all, a risk averse agent always values a lottery at a price strictly less than its expected payoff. Then why will he buy a security at any price arbitrarily smaller than his belief? The key point here is that demand is continuous. The participant may not wish to buy a fixed amount (say, one unit) of the security at a price slightly below his belief, but he will always be willing to buy some (possibly very small) amount of it.

We turn now to analysis of equilibrium prices. The next proposition establishes that the market has one of the properties often considered desirable for opinion pools.

Proposition 3.5 (Unanimity) *If all agents are risk averse or risk neutral with equal beliefs across the events of interest (*$\Pr_h(A_j) = \Pr_i(A_j)$ *for all* $h, i = 1 \ldots N$ *and* $j = 1 \ldots S$*), then the unique competitive equilibrium prices equal these beliefs (*$p^{\langle j \rangle} = \Pr(A_j)$ *for all* j*).*

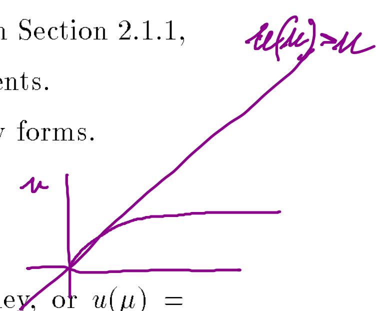
[6] Actually $\ln(\mu + b)$ is imaginary for $\mu < -b$. But since utility is increasing, and is $-\infty$ at $\mu = -b$, we extrapolate that it is $-\infty$ for $\mu < -b$.

Proof. At prices $p^{\langle j \rangle} = \Pr(A_j)$, $\mathbf{x} = \mathbf{0}$ is optimal for risk averse agents (by Proposition 3.4) and *any* demand, including $\mathbf{x} = \mathbf{0}$, is optimal for risk neutral agents. Thus excess demand can be zero, (3.2) is satisfied, and these prices form a competitive equilibrium. We prove by contradiction that, at any other prices $p^{\langle j \rangle} \neq \Pr(A_j)$, the market cannot be in equilibrium. By Lemma 3.3, each agent i has demand such that $E[L(\mathbf{x}_i)] > 0$. The equilibrium condition (3.2) and identical beliefs imply that $E[\sum_{i=1}^{N} L(\mathbf{x}_i)] = \sum_{i=1}^{N} E[L(\mathbf{x}_i)] = 0$. But this implies that $E[L(\mathbf{x}_i)] \leq 0$ for some i. $\square$

3.3 Market Pooling Functions with Disjoint Securities

In this section, we characterize the dependence of prices on beliefs within our model, when all events are mutually exclusive; more general event structures are considered in Section 3.4. To retain linear independence, events are not exhaustive. For notational purposes, define the event $A_0 \equiv \bar{A}_1 \bar{A}_2 \cdots \bar{A}_S$, though it does not have a corresponding security.[7] Disjoint events are certainly a special case, however, as discussed in Section 2.1.1, any event space can be partitioned into securities representing disjoint events.

We consider, in turn, two specific assumptions regarding agents' utility forms.



3.3.1 Exponential Utility

First, suppose that all agents have negative exponential utility for money, or $u(\mu) = -e^{-c\mu}$. This utility form is synonymous with *constant absolute risk aversion* (CARA), where c is the coefficient of risk aversion, or $1/c$ the risk tolerance.

We break up the discussion into two parts: (1) each agent's optimization problem, and (2) their collective effect on equilibrium prices.

Individual Demand

To facilitate discussion, first consider the simplest case of a market in only one security $\langle A \rangle$ priced at p. Each participant must decide what quantity, x, to buy ($x > 0$) or sell ($x < 0$) given its belief, $\Pr(A)$, and utility for money, $u(\mu) = -e^{-c\mu}$. The decision to purchase x units is equivalent to the acceptance of a lottery $L(x)$ with payoff $(1-p)x$ if

[7] $\langle A_0 \rangle$ can be replicated by selling equal amounts of all securities.

$U(x) = \Pr(A)\cdot(1-p)x + \Pr(\bar{A})\cdot -px = \Pr(A)\cdot(1-p)x - (1-\Pr(A))px$

$= x\cdot\left(\Pr(A)(1-p) - (1-\Pr(A))-p)\right)$

A occurs, and $-px$ otherwise. In this case, utility (3.1) becomes

$$U(x) = E[u(L(x))] = -\Pr(A)e^{-c(1-p)x} - \Pr(\bar{A})e^{cpx}.$$

Solving for optimal demand yields a closed form,

$$x(p) = \frac{1}{c}\ln\left(\frac{\Pr(A)}{p}\cdot\frac{(1-p)}{\Pr(\bar{A})}\right). \tag{3.3}$$

Consistent with Proposition 3.4, the agent's demand is directly related to its belief in the probability of A, and inversely related to the price of the security. At $p = \Pr(A)$, demand is zero. As risk aversion approaches zero (approximating risk neutrality), the agent is willing to buy or sell increasing amounts of the good, assuming that price does not equal belief.

In a market of S securities, each agent's unique optimal demand is:

$$x^{\langle j\rangle}(\mathbf{p}) = \frac{1}{c}\ln\left(\frac{\Pr(A_j)}{p^{\langle j\rangle}}\cdot\frac{1-p^{\langle 1\rangle}-p^{\langle 2\rangle}-\cdots-p^{\langle S\rangle}}{\Pr(A_0)}\right). \tag{3.4}$$

Equilibrium Prices

We next examine how agent decisions affect the market's equilibrium prices, which, by construction, form the group's consensus probabilities. The condition for competitive equilibrium (3.2) is $\sum_{i=1}^{N} x_i^{\langle j\rangle}(\mathbf{p}) = 0$, for all $j = 1\ldots S$. Solving for $\mathbf{p}$, we arrive at the following proposition.

Proposition 3.6 (CARA $\Rightarrow$ LogOP) *In a market of disjoint securities, each agent having CARA c_i, the prices*

$$p^{\langle j\rangle} = \frac{\prod_{i=1}^{N}[\Pr_i(A_j)]^{\alpha_i}}{\prod_{i=1}^{N}[\Pr_i(A_0)]^{\alpha_i} + \prod_{i=1}^{N}[\Pr_i(A_1)]^{\alpha_i} + \cdots + \prod_{i=1}^{N}[\Pr_i(A_S)]^{\alpha_i}} \tag{3.5}$$

form a competitive equilibrium, where $\alpha_i = (1/c_i)/\sum_l(1/c_l)$ is the normalized risk tolerance of agent i.

Proof. These prices can be shown to satisfy (3.2) by substitution, and some lengthy algebra. $\square$

These prices are the same probabilities as derived by the LogOP (2.6), where the expert weights are normalized risk tolerances. The market model, therefore, provides one way

if $u(w)=w$ how do prices look like?

to ground a well-known centralized pooling mechanism in terms of individual behavior. Rubinstein (1974) derives a result analogous to Proposition 3.6 within a more general securities market framework. The major difference is that the normalization of prices falls out naturally in our derivation, yielding a legal probability distribution. Wilson (1968) also arrives at the weighted geometric average relationship when agents (with CARA) share the profits of a security that pays off contingent on a single, continuous random variable. Nau and McCardle (1991) derive a similar result for more than one continuous variable when all beliefs are multivariate normal, and quadratic options are available in addition to the standard securities.

Proposition 3.6 provides a decision-theoretic interpretation for the notoriously slippery concept of expert weights. In the usual interpretation, the exponents in (3.5) encode some sort of degree of expertise, confidence, or reliability, and are almost always chosen in an ad hoc manner (Benediktsson and Swain, 1992; French, 1985; Genest and Zidek, 1986; Ng and Abramson, 1994; Winkler, 1968). In arguing against the use of opinion pools, French (1985) presents as his first reason that

> ...they all introduce weights ...which are not operationally defined. How should they be chosen? To say that one expert is twice as good as another is a figure of speech, not an arithmetic statement. The problem is further complicated by the likelihood of correlation between the expert's [sic] opinions. Several pragmatic solutions have been proposed ..., but to my knowledge none avoid a certain arbitrariness.

In the market model, weights are directly related to a well-understood quantity, namely risk tollerance. Note also that the weights sum to unity, as is the standard convention.

It may seem that weights *should* relate to a notion of confidence, however loosely defined, rather than to risk tolerance. In Section 4.1, we propose an extension of the market model in which agents learn from prices, and show how a confidence-related interpretation arises naturally.

3.3.2 Generalized Logarithmic Utility

Next consider an assumption that all agents exhibit *generalized logarithmic utility* (GLU) for money: $u(\mu) = \ln(\mu + b)$. With the further constraint that $b > 0$, GLU is a standard *decreasingly* risk-averse form. Note that u is imaginary on the interval $(-\infty, -b)$. Since utility approaches $-\infty$ as $\mu \to -b$ from the right, and u is increasing, we can safely extrapolate that it is $-\infty$ for all $\mu \leq -b$. The upshot is that the agent would *never* put b or more dollars at risk, regardless of probability assessments. Note that any wealth accumulated in dollars is additive with b, so a natural interpretation of b is as the agent's initial wealth.

We start with the simplest case: one agent investing in one security. Utility (3.1) is

$$U(x) = -\Pr(A)\ln((1-p)x + b) - \Pr(\bar{A})\ln(-px + b).$$

The limit on maximum dollars at risk implies that $-b/(1-p) < x < b/p$, since outside this range $U(x) = -\infty$. These bounds can also be understood as follows: if an agent sells a quantity $-b/(1-p)$ of the security, it stands to lose \$$b$ in the worst case; similarly if it buys b/p units it also could lose \$$b$. Within this range, optimal demand is

$$x(p) = b\left(\frac{\Pr(A)}{p} - \frac{\Pr(\bar{A})}{1-p}\right). \tag{3.6}$$

For S securities, unique optimal demand is

$$x^{\langle j\rangle}(\mathbf{p}) = b\left(\frac{\Pr(A_j)}{p^{\langle j\rangle}} - \frac{\Pr(A_0)}{1 - p^{\langle 1\rangle} - p^{\langle 2\rangle} - \cdots - p^{\langle S\rangle}}\right) \tag{3.7}$$

and equilibrium prices can again be characterized in closed form.

Proposition 3.7 (GLU ⇒ LinOP) *In a market of disjoint securities, each agent having GLU with initial wealth b_i, the prices*

$$p^{\langle j\rangle} = \sum_{i=1}^{N} \beta_i \Pr_i(A_j) \tag{3.8}$$

NB

form a competitive equilibrium, where $\beta_i = b_i/\sum_l b_l$ is agent i's normalized initial wealth.

Proof. These prices satisfy (3.2) by substitution. □

This is exactly the LinOP formula (2.5), where the expert weights are again directly related to risk tolerance. This result can again be seen as a simplified derivation of that by Rubinstein (1974, 1975, 1976).

Thus both the linear and logarithmic opinion pools arise as special cases of our market mechanism; choosing still other utility forms seems a natural way to catalog a variety of aggregation functions.

3.3.3 Composite Agents

With disjoint securities, and under each of the utility assumptions (CARA, GLU) the market as a whole behaves as a single, *composite* agent with probabilities equal to prices. An outsider, offering securities to the group at various prices, cannot differentiate the aggregate demand response from that of an individual.

Proposition 3.8 (Composite CARA Agent) *Suppose that, in a market of disjoint securities, all agents exhibit CARA with risk aversion c_i, and effect equilibrium prices $\mathbf{q}$. Then their total demand for security $\langle A_j \rangle$ offered at price $p^{\langle j \rangle}$ is equal to that of a composite agent with beliefs $\mathrm{Pr}_0 = \mathbf{q}$ and CARA utility with risk tolerance $1/c_0 = \sum_{i=1}^{N} 1/c_i$.*

Proof. For each security, A_j, total demand is:

$$
\begin{aligned}
x_0^{\langle j \rangle} &= \sum_{i=1}^{N} \frac{1}{c_i} \ln \left(\frac{1 - p^{\langle 1 \rangle} - \cdots - p^{\langle S \rangle}}{p^{\langle j \rangle}} \cdot \frac{\mathrm{Pr}_i(A_j)}{\mathrm{Pr}_i(A_0)} \right) \\
&= \frac{1}{c_0} \ln \left(\frac{1 - p^{\langle 1 \rangle} - \cdots - p^{\langle S \rangle}}{p^{\langle j \rangle}} \right) + \frac{1}{c_0} \ln \left(\frac{\prod_{i=1}^{N} [\mathrm{Pr}_i(A_j)]^{c_0/c_i}}{\prod_{i=1}^{N} [\mathrm{Pr}_i(A_0)]^{c_0/c_i}} \right) \\
&= \frac{1}{c_0} \ln \left(\frac{1 - p^{\langle 1 \rangle} - \cdots - p^{\langle S \rangle}}{p^{\langle j \rangle}} \cdot \kappa \prod_{i=1}^{N} [\mathrm{Pr}_i(A_j)]^{\alpha_i} \cdot \frac{1}{\kappa \prod_{i=1}^{N} [\mathrm{Pr}_i(A_0)]^{\alpha_i}} \right) \\
& \quad \text{where } \kappa \equiv \frac{1}{\prod_{i=1}^{N} [\mathrm{Pr}_i(A_0)]^{\alpha_i} + \cdots + \prod_{i=1}^{N} [\mathrm{Pr}_i(A_S)]^{\alpha_i}} \\
&= \frac{1}{c_0} \ln \left(\frac{1 - p^{\langle 1 \rangle} - \cdots - p^{\langle S \rangle}}{p^{\langle j \rangle}} \cdot \frac{q^{(j)}}{1 - q^{(1)} - \cdots - q^{(S)}} \right) \text{(from (3.5))} \\
&= \frac{1}{c_0} \ln \left(\frac{1 - p^{\langle 1 \rangle} - \cdots - p^{\langle S \rangle}}{p^{\langle j \rangle}} \cdot \frac{\mathrm{Pr}_0(A_j)}{\mathrm{Pr}_0(A_0)} \right).
\end{aligned}
$$

□

Proposition 3.9 (Composite GLU Agent) *Suppose that, in a market of disjoint securities, all agents exhibit GLU with initial wealth b_i, and effect equilibrium prices $\mathbf{q}$. Then their total demand for security $\langle A_j \rangle$ offered at price $p^{\langle j \rangle}$ is:*

$$
x_0^{\langle j \rangle} = b_0 \left(\frac{\mathrm{Pr}_0(A_j)}{p^{\langle j \rangle}} - \frac{\mathrm{Pr}_0(A_0)}{1 - p^{\langle 1 \rangle} - p^{\langle 2 \rangle} - \cdots - p^{\langle S \rangle}} \right),
$$

or equal to that of a composite agent with beliefs $\mathrm{Pr}_0(A_j) = q^{(j)}$ and GLU with initial wealth $b_0 = \sum_{i=1}^{N} b_i$.

Proof. Analogous to the proof of Proposition 3.8. □

Furthermore, any subset of the agents can be aggregated together, interacting with the rest of the system as if it were an individual. Wilson (1968) first described conditions for a composite agent, which he called a *surrogate*, extended later by Rubinstein (1974) and covered in Huang and Litzenberger's textbook (1988). Notice that group rationality emerges from the individual decisions of its members, unlike the classic supra Bayesian, whose existence is imposed at the outset. Assuming positive c_i, the composite agent's risk aversion is strictly less than that of any individual; in other words, the group together is willing to take on more risk than any proper subset of its members. Raiffa (1968) elucidates the connection between risk sharing and group decision making, discussing the case of CARA in depth.

A mixed population, consisting of some agents with CARA and some with GLU, can be summarized as *two* composite agents, each representing its respective subpopulation. In this case, numeric solutions, or approximate formulas, for price are possible.

3.4 Arbitrary Securities

Deriving closed-form pricing formulae for arbitrary securities is difficult, even under homogeneous utility constraints. Although it is a straightforward task to write down each agent's expected utility for money, the number of terms is in general exponential in the number of events. Numeric maximization methods such as gradient descent or Newton's method will be well behaved, since utility for securities is concave, though will still suffer from exponential blowup in the size of the objective function.

Nevertheless a few things can be said for the general case. First, if $S = |\Omega| - 1$, and securities are linearly independent, then the market is *complete* and prices define a full probability distribution over Ω. Moreover, the distribution is irrespective of which $|\Omega| - 1$ securities are chosen (Mas-Colell et al., 1995; Varian, 1987). In particular, assuming either CARA or GLU, the consensus is the same as implied by (3.5) or (3.8), respectively, in a market with securities corresponding to all but one state.

Any prices that form incoherent probabilities (that allow arbitrage) cannot be part of an equilibrium. If an agent, with *any* utility function that is monotonically increasing for money, believes that some outcome is certain ($\Pr(A) \in \{0, 1\}$), or that one subset of events is logically equivalent to another, then it will effectively dictate that relationship in the consensus, by demanding infinite quantities if prices suggest otherwise.

Under the CARA restriction, I have also derived some results for two arbitrary events,

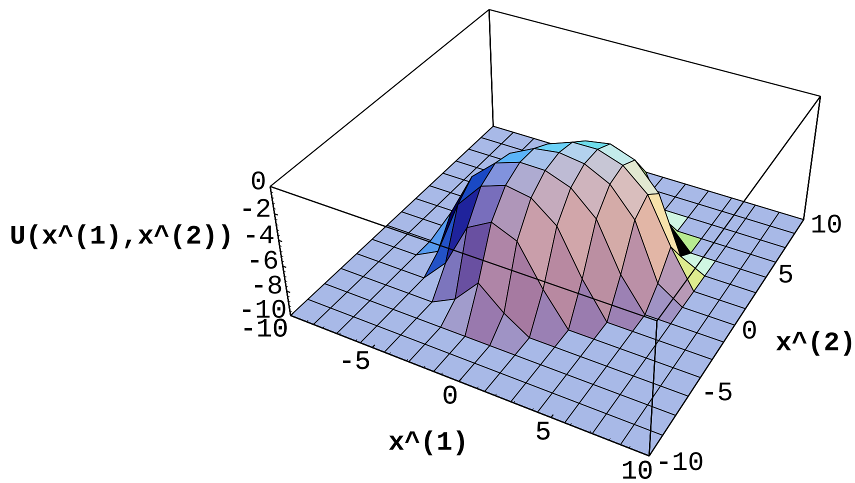


Figure 3.1: Utility (3.9) versus demands $x^{\langle 1\rangle}$ and $x^{\langle 2\rangle}$ for the parameters $\Pr(A_1) = \Pr(A_2) = 0.5$, $\Pr(A_1A_2) = 0.1$, $p^{\langle 1\rangle} = p^{\langle 2\rangle} = 0.5$, and $c = 1$. Utility is maximized at $\{x^{\langle 1\rangle}, x^{\langle 2\rangle}\} = \{0, 0\}$, as expected when price equals belief.

and for events with special independence structures.

3.4.1 Two Securities

We can solve for optimal demand of an agent with CARA, in a market of two arbitrary securities $\langle A_1\rangle$ and $\langle A_2\rangle$. Following the same progression above, the utility for the two securities is the expected utility of the corresponding lottery, which for CARA is given by

$$\begin{aligned} U(x^{\langle 1\rangle}, x^{\langle 2\rangle}) = E[u(L(x^{\langle 1\rangle}, x^{\langle 2\rangle}))] = \\ &- \Pr(A_1A_2)e^{-c\left[(1-p^{\langle 1\rangle})x^{\langle 1\rangle}+(1-p^{\langle 2\rangle})x^{\langle 2\rangle}\right]} \\ &- \Pr(A_1\bar{A}_2)e^{-c\left[(1-p^{\langle 1\rangle})x^{\langle 1\rangle}-p^{\langle 2\rangle}x^{\langle 2\rangle}\right]} \\ &- \Pr(\bar{A}_1A_2)e^{-c\left[-p^{\langle 1\rangle}x^{\langle 1\rangle}+(1-p^{\langle 2\rangle})x^{\langle 2\rangle}\right]} \\ &- \Pr(\bar{A}_1\bar{A}_2)e^{-c\left[-p^{\langle 1\rangle}x^{\langle 1\rangle}-p^{\langle 2\rangle}x^{\langle 2\rangle}\right]}. \end{aligned} \tag{3.9}$$

The decision variables in this optimization problem are coupled; optimal demand for A_1 may depend on the demand for A_2, and thus on the price of A_2 (and vice versa). Figure 3.1 graphs utility (3.9) as a function of $x^{\langle 1\rangle}$ and $x^{\langle 2\rangle}$ for a particular instantiation of beliefs, risk coefficient, and prices. By solving the first-order condition, we can find a closed form for optimal demand:

$$x^{\langle 1\rangle}(p^{\langle 1\rangle}, p^{\langle 2\rangle}) = \frac{1}{c}\ln\left(\frac{Q(1-p^{\langle 1\rangle}-p^{\langle 2\rangle}) + R(p^{\langle 2\rangle}-p^{\langle 1\rangle}) + \sqrt{Z}}{2\Pr(\bar{A}_1A_2)\Pr(\bar{A}_1\bar{A}_2)p^{\langle 1\rangle}}\right), \tag{3.10}$$

where

$$
\begin{aligned}
Z &\equiv Q^2(1-p^{\langle 1\rangle}-p^{\langle 2\rangle})^2+2QR[p^{\langle 1\rangle}(1-p^{\langle 1\rangle})+p^{\langle 2\rangle}(1-p^{\langle 2\rangle})]+R^2(p^{\langle 2\rangle}-p^{\langle 1\rangle})^2,\\
Q &\equiv \Pr(A_1\bar{A}_2)\Pr(\bar{A}_1A_2),\ \text{and}\\
R &\equiv \Pr(A_1A_2)\Pr(\bar{A}_1\bar{A}_2).
\end{aligned}
$$

Optimal demand for A_2 is analogous, with security subscripts switched. Note that, as expected, (3.10) coincides with (3.4) when A_1 and A_2 are disjoint, and $S=2$.

When the two events are dependent, demands for the two goods are correlated in the opposite direction of the dependence. That is, $x^{\langle 1\rangle}$ is increasing in $x^{\langle 2\rangle}$ when $\Pr(A_1|A_2) < \Pr(A_1|\bar{A}_2)$, and decreasing in $x^{\langle 2\rangle}$ when $\Pr(A_1|A_2) > \Pr(A_1|\bar{A}_2)$. This behavior arises because negatively correlated goods provide *insurance* for each other, whereas positively related events increase the exposed risk, and because both securities pay off in a common currency (dollars) for which the agent is risk averse.

For a moment, let us relax the assumption that no security is equivalent to a combination of other securities, in order to examine its implications. Consider two securities representing equivalent events, $A_1 = A_2$. If $p^{\langle 1\rangle} \neq p^{\langle 2\rangle}$, then the agent can achieve unbounded utility through arbitrage, buying infinite amounts of the cheaper security to sell in the higher priced market. If the prices coincide, then the agent demands the same total amount as if there were only one security, splitting its demand arbitrarily between the two available. Put another way, it still desires the same total amount of $A_1 = A_2$, but there also exists a risk-free portfolio (z of A_1 and $-z$ of A_2) for which the agent is indifferent between any z. Figures 3.2 and 3.3 graph expected utility (3.9) versus $x^{\langle 1\rangle}$ and $x^{\langle 2\rangle}$ for this situation. The former illustrates the case when $p^{\langle 1\rangle} = p^{\langle 2\rangle}$, the latter when $p^{\langle 1\rangle} > p^{\langle 2\rangle}$.

The same basic principles hold when $A_1 = \bar{A}_2$. If $p^{\langle 1\rangle} \neq 1 - p^{\langle 2\rangle}$, then an arbitrage opportunity exists. Otherwise, there exists a risk-free investment option (buy equal amounts of both), and optimal demand is not unique.

3.4.2 Independent Events

More can be said as well when agents have CARA, and certain independence conditions among events hold.

Proposition 3.10 (Separation of independent events) *If an agent with CARA holds that* $\Pr(A_1|A_2\cdots A_S) = \Pr(A_1)$*, then its optimization problem for* $x^{\langle 1\rangle}$ *is separable, and*

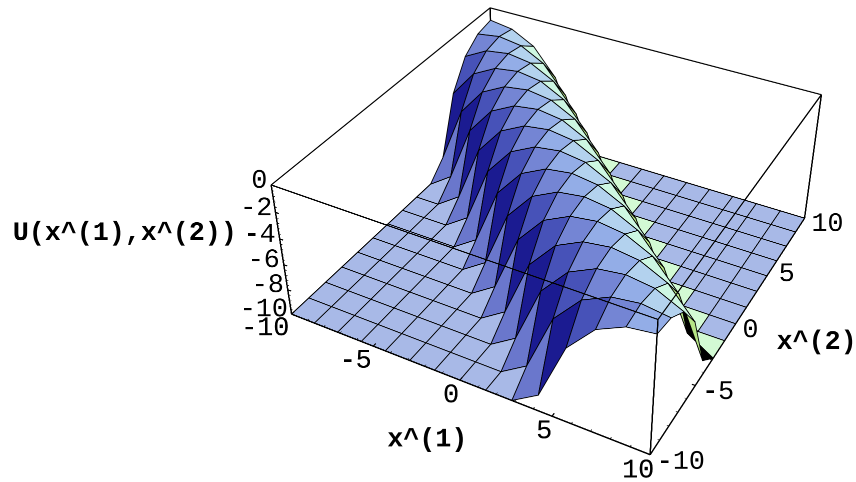


Figure 3.2: Utility (3.9) versus demands $x^{\langle 1\rangle}$ and $x^{\langle 2\rangle}$ for the parameters $\Pr(A_1) = \Pr(A_2) = \Pr(A_1A_2) = 0.5$ (thus implying that the events are equivalent), $p^{\langle 1\rangle} = p^{\langle 2\rangle} = 0.5$, and $c = 1$. Utility is a maximum along the line $x^{\langle 1\rangle} + x^{\langle 2\rangle} = 0$; the agent's total demand for $A_1 = A_2$ is zero, but it splits purchases arbitrarily between the equivalent markets.

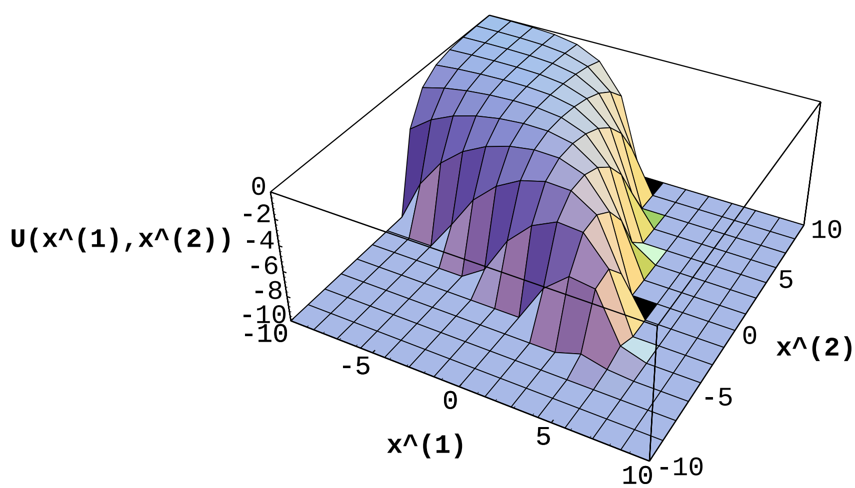


Figure 3.3: Utility (3.9) versus demands $x^{\langle 1\rangle}$ and $x^{\langle 2\rangle}$ for the parameters $\Pr(A_1) = \Pr(A_2) = \Pr(A_1A_2) = 0.5$ (thus implying that the events are equivalent), $p^{\langle 1\rangle} = 0.7$, $p^{\langle 2\rangle} = 0.3$, and $c = 1$. The prices are inconsistent with the logical relationship $A_1 = A_2$, and the agent maximizes utility by selling A_1 and buying A_2 in infinite quantities.

reduces to the single-security case.

Proof. Let $X = \{A_2, \ldots, A_S\}$ be the set of all events except A_1. Utility (3.1) is

$$\begin{aligned} -\sum_X &\left(\Pr(A_1)\Pr(X)e^{-c[(1-p^{\langle 1\rangle})x^{\langle 1\rangle}+\Upsilon^{\langle X\rangle}]} + \Pr(\bar{A}_1)\Pr(X)e^{-c[-p^{\langle 1\rangle}x^{\langle 1\rangle}+\Upsilon^{\langle X\rangle}]}\right) \\ = \quad & -\left(\Pr(A_1)e^{-c(1-p^{\langle 1\rangle})x^{\langle 1\rangle}} + \Pr(\bar{A}_1)e^{-c(-p^{\langle 1\rangle}x^{\langle 1\rangle})}\right)\sum_X \Pr(X)e^{-c\Upsilon^{\langle X\rangle}} \end{aligned}$$

where $\sum_X$ is a sum over all possible joint outcomes of events in X, and $\Upsilon^{\langle X\rangle}$ is the payoff in the selected outcome. The optimal choice for $x^{\langle 1\rangle}$ is independent of demand for other securities, and has the same solution as if $\langle A_1\rangle$ were offered in isolation. □

It follows that when *all* agents have CARA, and everyone agrees that A_1 is independent of all other events, the equilibrium price is given by (3.5) as if $\langle A_1\rangle$ were the only security present ($S = 1$). Moreover, because its price does not depend on the outcome of any other event, A_1 remains independent of all other events in the distribution implied by equilibrium prices. I derive more general independence preservation properties of the LogOP in Chapter 5 and investigate implications for securities markets in Chapter 7.

In general, including the case of GLU, independence of events does *not* lead to separable demand. Even if everyone agrees that an event is independent of all others, a dependence will generally be introduced among the equilibrium prices.

3.5 Exposing a Group

Under what conditions can the collective behavior of a group be differentiated from that of an individual? Can an outsider O, by querying the group's demand for securities at various prices, determine whether $N > 1$? The composite agent results in Section 3.3.3 identify conditions under which no distinction can be made within the context of a *fixed* market configuration. In this section, we discuss how O might deduce the makeup of the group by examining measurements collected under different configurations.

Assume that agents are myopic, optimizing (3.1) across *currently available* securities. The agents do not attempt to predict what O will offer next, but simply react according to whatever markets exist at the moment. O has the power to open and cancel markets at will; the agents are essentially memoryless, and blindly recompute their optimization problem in the new environment, with negligible animosity toward O's probing. Agents may have state-dependent utilities or prior stakes in events, though the latter change only via trades in available securities.

We might envision the group as subject to varying degrees of isolation from the outside world.

ISO Agents are isolated from outside markets, beyond those offered by O.

COMP Agents have outside access to a complete securities market.

NCOMP Agents have outside access to $1 \leq S < |\Omega| - 1$ securities.

Prices of outside securities, and agent beliefs, are assumed constant throughout the observation procedure.

Suppose that O first opens a market in a single security and observes the equilibrium price. O then withdraws this market and opens two more in disjoint securities that form a partition of the first. That is, O determines both of the following:

- the equilibrium price $p^{\langle A_1 \rangle}$ for a single security $\langle A_1 \rangle$
- the equilibrium prices $p^{\langle A_1 A_2 \rangle}$ and $p^{\langle A_1 \bar{A}_2 \rangle}$ for securities $\langle A_1 A_2 \rangle$ and $\langle A_1 \bar{A}_2 \rangle$

Under condition ISO, If $N = 1$ then, by de Finetti's argument,

$$p^{\langle A_1 \rangle} = p^{\langle A_1 A_2 \rangle} + p^{\langle A_1 \bar{A}_2 \rangle}. \tag{3.11}$$

Thus, if (3.11) does *not* hold, we can be certain that $N > 1$. Consensus beliefs that always adhere to (3.11) satisfy what is called the *marginalization property* (MP). Under ISO, in exceptional circumstances—for example, all agents have GLU and thus prices form a LinOP, or all beliefs are identical—Equation 3.11 is upheld by a group as well (Genest, 1984c). Generically, however, prices revealed by a group will *not* conform to (3.11). Under condition COMP, (3.11) will hold always, independent of N, since both the LHS and RHS must be consistent with the distribution over Ω fully specified by the prices in the complete market. Under NCOMP, (3.11) may *not* hold, again regardless of the cardinality of the group. Even if the "group" is a single agent, the prices it reveals will be affected by others via the outside markets.

Now imagine that O conducts, in turn, three experiments to elicit the following information:

- the equilibrium price $p^{\langle A_1 A_2 \rangle}$ for security $\langle A_1 A_2 \rangle$

- the equilibrium price $p^{\langle A_1|A_2\rangle}$ for *conditional* security[8] $\langle A_1|A_2\rangle$
- the equilibrium price $p^{\langle A_2\rangle}$ for security $\langle A_2\rangle$

Under condition ISO, an individual necessarily reveals

$$p^{\langle A_1A_2\rangle} = p^{\langle A_1|A_2\rangle}p^{\langle A_2\rangle}. \tag{3.12}$$

An opinion pool that always respects (3.12) satisfies what I call the *family aggregation* (FA) property, defined in Chapter 5. In that chapter, I show that *no* combination of beliefs can simultaneously satisfy FA, unanimity, and nondictatorship.[9] Unanimity (Proposition 3.5) and nondictatorship are implicit in the market design. It follows that, under ISO, even when the conditions for a standard composite agent are met (all have CARA or all have GLU), Equation 3.12 is generically true if and only if $N = 1$. Under COMP, (3.12) always holds, and under NCOMP it may be violated even by an individual.

3.6 Related Work

Eisenberg and Gale (1959) propose a betting-based aggregation mechanism built not upon the securities market framework, but rather on the common method for deriving odds in horse races. They consider a pari-mutuel scheme where agents place bets across a partition of events, yielding a consensus probability equal to the proportion of the total bet on each event. If event A obtains, agents share the total amount bet in proportion to their bets on A. Agents bet to maximize expected payoffs, where expectation is with respect to their probability distribution over the events, subject to a budget constraint limiting their total bets. Eisenberg and Gale show that this mechanism yields a unique set of equilibrium probabilities, and Norvig (1967) presents a dynamic process for reaching this equilibrium through iterated bids.

As the authors point out, however, this scheme can yield pathological (their word) results. For example, if there are two bettors with equal budgets, then whichever has more uniform probabilities will dictate the results. According to Genest and Zidek (1986), the

[8] A conditional security $\langle A_1|A_2\rangle$ pays off \$1 if A_1A_2 occurs and nothing if $\bar{A}_1A_2$ occurs. If $\bar{A}_2$ occurs, then the bet is called off, and any price paid for the security is returned.

[9] FA essentially requires *both* MP and external Bayesianity (EB) (Genest, 1984a) to hold. Genest (1984b), and indirectly Dalkey (1972, 1975), prove that MP, EB, unanimity, and nondictatorship are mutually incompatible.

pari-mutuel approach to belief aggregation "has never enjoyed much popularity" for this reason. The pathological behavior, we believe, can be attributed to the role of arbitrary budgets. In the approach developed here, we impose no budgets, but rather rely on risk aversion to limit bets to the finite range.

Plott et al. (1997) investigate whether pari-mutuel markets are able to aggregate information, as postulated by RE theory, in a laboratory setting. In one set of experiments, each subject was given inside knowledge that a subset of horses would definitely *not* win. Although all agents were uncertain as to the outcome, their *collective* information was enough to identify the winning horse with certainty. Information aggregation did occur, and RE-based predictions fit the data well. In a second set of experiments, subjects were given probabilistic evidence in the form of a sample of random draws from the true distribution. To free subjects from the burden of computing complex Bayesian updates—and to control for associated errors—correct posterior distributions were provided along with the evidence. Some weak aggregation did occur, but not to the extent predicted by the RE model; instead, a prior information model proved more accurate.

In earlier work, Plott and Sunder (1982, 1988) conducted laboratory experiments to test the reasonableness of the RE assumption in the context of a securities market. Agents were initially unaware which of three states would occur. A typical market consisted of one or three securities, three *types* of agents with different preferences, and four agents of each type. Within each of several rounds, one-period securities were traded. Securities paid off differently according to the realized state and an agent's type. In one study (1982), privileged insiders were given categorical knowledge of the underlying state. The RE model's predictions—that equilibrium prices and other economic variables converge as if everyone were aware of the true state—were significantly more accurate than those of other models, including the classical Walrasian or PI hypothesis that agents do not revise their beliefs based on prices. When insiders were given less than certain information, the results were not definitive. In a second study (1988), insiders were told only that one of the three states would *not* occur. The combined knowledge of all agents was sufficient to logically infer the true state, though no single insider could directly do so. It was found that, in a complete market of three securities, the RE predictions were again the most accurate. In a single security market with agents of different preference types, the RE equilibrium was *not* realized. On the other hand, even in this last condition, Forsythe and Lundholm (1990), with a similar experimental design, found that RE *was* verified, as long as subjects were experienced and knew the possible types of agents.

Beyond the controlled setting of the laboratory, researchers have investigated the aggregative properties of public markets. Hanson (1995, 1999) proposes an *Idea Futures* market, where participants trade in securities that pay off contingent on future developments in science, technology, or other arenas of public interest. He argues that the reward structure of such a market encourages honest revelation of opinions among scientists, and provides more accurate probability assessments for use by funding agencies, public policy leaders, the media, and other interested parties. The concept is operational as a game (i.e., played with fake money), called the "Foresight Exchange", on the world-wide web at `http://www.ideosphere.com/`. Hanson also describes a scenario employing a securities market mechanism for coordinating computational agents (Hanson, 1991).

In a similar vein, though run with real money, the Iowa Electronic Market (IEM) (`http://www.biz.uiowa.edu/iem/`) supports trading in securities tied to the outcome of political and financial events. Their 1988 market, open only to University of Iowa students and employees, offered securities that paid off proportionally to the percentage of votes received by various candidates in that year's US Presidential election. The final prices matched Bush's final percent margin of victory more closely than any of the six major polls (Forsythe et al., 1992). Since opening to the public, subsequent US Presidential election markets have attracted wide participation and following.

West (1984) presents one axiomatic justification for the (unnormalized) LogOP, based on the betting behavior of a group. He defines S_i as the value for which agent i is indifferent between D_i units of security $\langle A \rangle$, and S_i dollars for sure. From this and u_i, one can derive $\Pr_i(A)$. Consider an assumption that the group as a whole will be indifferent between the allocation of $D = \langle D_1, \ldots, D_N \rangle$ units of the security, and the certain payoff $S = \langle S_1, \ldots, S_N \rangle$. From the premise that this is true for any D (along with some other conditions), West proves that the group's $\Pr_0(A)$ must obey an *unnormalized* geometric mean or unnormalized LogOP. In contrast to the normalized version (2.6), this pooling function does not in general yield a probability (e.g., if any agents disagree, $\Pr(A) + \Pr(\bar{A}) < 1$). West goes on to show that the degree by which the aggregate diverges from a probability can be used as a measure of disagreement among the agents.[10] However, we would make the case that the group should not be indifferent between S dollars and security distribution D, unless the individual beliefs happen to be the same. In fact, the

[10] One of the purposes of West's investigation was to demonstrate the pitfalls of blindly aggregating probabilities according to simple formulas (Mike West, personal communication). We share this motivation, and agree with the idea of deriving aggregate measures from behavioral postulates.

group should prefer the latter. The reason is that if the group has the security distribution, then there will be trading opportunities that can make everyone better off. We suspect that this preference can account for the normalization factor in the pooling procedure, a consequence of the market model presented here.

Varian (1985, 1989) analyzes the effects of divergence of opinion among agents on asset valuation and trading volume in a securities market, under various preference models. He shows that increased dispersion of beliefs is usually associated with higher volumes of trade. As long as agents' risk does not decrease too rapidly, increased dispersion also coincides with lower prices for securities.

Chapter 4

Agent Learning; Market Adaptation

This chapter extends the model of Chapter 3 to include agents that learn and multiperiod markets. Section 4.1 examines the case when agents update their beliefs based on market signals, by regarding prices as if they were the beliefs of another individual. Under particular assumptions, pricing equations retain the form of standard opinion pools, though the *confidence* that agents express in their own beliefs relative to market prices factor into the expert weights. Section 4.2 describes analysis and simulation of the dynamics of a securities market over time. In some circumstances, the redistribution of wealth among agents leads to price dynamics that can be rationalized as the Bayesian updates of a composite agent.

4.1 Learning from Prices: A Supra Bayesian Model of Rational Expectations

Chapter 3 assumes a classical *prior information* (PI) equilibrium in which agents' beliefs are fixed. Yet my advocacy of a market-based pooling device presupposes that exogenous prices do reveal meaningful aggregate information. If the agents themselves agree, then we would expect them to update their beliefs based on the observed prices. Anecdotally, an agent whose probability is poorly aligned with the going price may begin to suspect that everyone else knows something that it does not, and adjust its belief accordingly. This is a motivating premise of the theory of *rational expectations* (RE) (Grossman, 1981; Lucas, 1972). The equilibrium conditions remain that all agents are optimizing and that total demand is zero. Implicit in the former requirement however is the dependence of each agent's *posterior* probability on the equilibrium prices. In terms of the model outlined in Section 3.1, we replace each fixed prior probability $\Pr(\cdot)$ in all equations with the conditional probability $\Pr(\cdot|\mathbf{p})$. This substitution in general transforms closed-form solutions for price into implicit equations. In some cases, explored below, we are able to resolve $\mathbf{p}$.

What is a reasonable model for $\Pr(\cdot|\mathbf{p})$? Recall that $\mathbf{p}$ has all the properties of a probability distribution and that, at least under certain homogeneity conditions, a formal equivalency between the market and a composite individual can be established. In a strong sense, the agent faces the same problem as one decision maker updating beliefs based on the probability assessments of one expert, where the "expert" in this case is the market itself. This is precisely the scenario where the supra Bayesian method (2.4) is both justified and well-developed. The philosophical concerns reviewed in Section 2.3.1 surrounding the identification of the supra Bayesian are not relevant here; moreover, some of the computational burden is reduced, since the agent updates based only on one composite distribution rather than all $N-1$ distributions. Note that this interpretation makes sense when the agent does not consider the effect of its own decisions on prices.

Depending on how the situation is modeled, several supra Bayesian formulations are plausible (Berger, 1985; Clemen and Winkler, 1993; Lindley, 1988; Morris, 1977; West and Crosse, 1992). We adopt one similar to that proposed by Morris (1983), and later extended by Rosenblueth and Ordaz (1992), as it ultimately yields closed-form results for equilibrium prices.

4.1.1 Logarithmic Utility and Beta Updates

Consider first a market of agents with generalized logarithmic utility for money (GLU); Section 4.1.2 addresses the parallel situation of agents with constant absolute risk aversion (CARA).

We restrict attention for now to a single-security market. The agent envisions the corresponding event A as one trial in a series of identical, independent Bernoulli trials with probability of success Π. The parameter Π is itself uncertain; assume that, after witnessing s positive outcomes or *successes* among n repetitions of the event, the agent's probability is $\Pr(\Pi = \pi) \propto \pi^{s-1}(1-\pi)^{n-s-1}$, that is, characterized by a Beta distribution with parameters s and $n-s$. The agent is myopic, acting as if the market is a one-shot game, so uses $\Pr(A) = E[\Pi] = s/n$ to guide its decisions. The independent Bernoulli trials interpretation need not be taken literally; one may prefer to view s/n and n as subjective measures of probability and precision, respectively. In particular, the parameters need not be constrained to the integers.

The agent models the "market" (i.e., everyone else) as if it has collectively observed s' successful outcomes of the event A in a total of n' trials. The likelihood function $\Pr(p|\Pi = \pi) \propto \pi^{s'}(1-\pi)^{n'-s'}$ is the corresponding Binomial distribution. For consistency, it must be that $p = s'/n'$. The desired posterior probability

$$\Pr(\Pi = \pi|p) \propto \Pr(p|\Pi = \pi) \Pr(\Pi = \pi)$$

is straightforward to compute, as the Beta and Binomial distributions form a natural conjugate class (Raiffa and Schlaifer, 1961). The result is another Beta distribution with parameters $s + s'$ and $n + n' - (s + s')$.

Reformulating in terms of A, we find that

$$\begin{aligned}
\Pr(A|p) &= E[\Pi|p] \\
&= \frac{s+s'}{n+n'} = \frac{n}{n+n'} \cdot \frac{s}{n} + \frac{n'}{n+n'} \cdot \frac{s'}{n'} \\
&= w \Pr(A) + (1-w)p,
\end{aligned}$$

where $w \equiv n/(n+n')$. The agent's posterior probability given the price is just a weighted average of its prior and the price. The weighting term captures the agent's perception of its own confidence, expressed in terms of number of observations, as compared to the market's.

The argument generalizes easily to a market of S disjoint securities, where the update rule becomes

$$\Pr(A_j|p^{\langle j\rangle}) = w \Pr(A_j) + (1-w)p^{\langle j\rangle}. \tag{4.1}$$

Weights could conceivably vary across events as well (intuitively, the expert may desire to express greater confidence in certain assessments), though the above justification does not support this extension, and (4.1) would require normalization.

To determine the demand of an agent with GLU, we substitute the posterior $\Pr(A_j|p^{\langle j\rangle})$ for the prior $\Pr(A_j)$ in (3.7). After some algebraic manipulation, we find that

$$x^{\langle j\rangle}(\mathbf{p}) = wb\left(\frac{\Pr(A_j)}{p^{\langle j\rangle}} - \frac{\Pr(A_0)}{1 - p^{\langle 1\rangle} - \cdots - p^{\langle S\rangle}}\right).$$

In words, the relative stock that an agent places on its own prior probability versus public prices serves to attenuate demand by a multiplicative factor.

In a market of likewise agents, the equilibrium prices still obey (3.8), though these are now implicit equations:

$$
\begin{aligned}
p^{\langle j\rangle} &= \sum_{i=1}^{N} \beta_i \mathrm{Pr}_i(A_j|p^{\langle j\rangle}) \\
&= \sum_{i=1}^{N} \frac{w_i b_i}{\sum_l b_l} \mathrm{Pr}_i(A_j) + \sum_{i=1}^{N} \frac{(1-w_i)b_i}{\sum_l b_l} p^{\langle j\rangle}.
\end{aligned}
$$

Solving for $p^{\langle j\rangle}$, we find that

$$
p^{\langle j\rangle} = \sum_{i=1}^{N} \frac{w_i b_i}{\sum_l w_l b_l} \mathrm{Pr}_i(A_j). \tag{4.2}
$$

Prices retain the form of the LinOP, but with weights proportional to the product of risk tolerance and self-assessed confidence. The introduction of a weighting factor that depends on the precision of probability assessments rather than solely on parameters of utility may be more palatable to advocates of the standard opinion pools. Why should consensus probabilities depend on utilities at all? Hylland and Zeckhauser's impossibility theorem (1979), outlined in Section 2.1.6, demonstrates that confounding belief and utility aggregation is simply *unavoidable* if the resulting composite agent is to (1) uphold unanimous decisions, (2) maintain unanimous agreement on probabilities, and (3) allow Pr_0 to depend on more than one agent. Moreover, standard opinion pools are themselves not exempt from the influence of utilities; to the contrary, typical elicitation procedures (de Finetti, 1974; Winkler, 1968) reveal the experts' *risk-neutral* probabilities (2.3) rather than their true probabilities (Kadane and Winkler, 1988), as noted in Section 2.1.5.

Under some conditions we can show that prices do encode the "correct" aggregate probabilities. That is, the market reaches a RE equilibrium where prices are exactly as if all agents had witnessed all observations for all events. Each agent has seen $s_i^{\langle j\rangle}$ positive occurrences of A_j out of n_i total observations. Suppose that agents' observations are independent, that each agent knows its relative number of observations (i.e., $w_i = n_i/\sum_h n_h$), and that all agents have the same risk tolerance. Then (4.2) becomes

$$
p^{\langle j\rangle} = \frac{\sum_{i=1}^{N} s_i^{\langle j\rangle}}{\sum_{i=1}^{N} n_i}.
$$

Even though the agents' posterior probabilities are *not* in complete agreement, this is precisely the same equilibrium that would have arisen had all agents (or a composite agent) seen all observations.

4.1.2 Exponential Utility and Lognormal Updates

When agents have CARA and the posterior update rule is a weighted *geometric* average of prior and price,

$$\Pr(A_j|p^{\langle j\rangle}) \propto [\Pr(A_j)]^w [p^{\langle j\rangle}]^{(1-w)}, \tag{4.3}$$

then each agent's demand function has the form (3.4), though attenuated by a factor of w, and equilibrium prices remain a weighted geometric average (3.5) with expert weights proportional to the product of risk tolerance and the weighting term w.

Morris (1983) mentions as an aside that (4.3) can be justified in a manner analogous to that in Section 4.1.1, if the *log-odds* of Π, and the log-odds of price, are *normally* distributed. In attempting to formalize his statement, I found a derivation that *almost* goes through, save for one step. Of course, my lack of success does not falsify Morris's claim. I present the near-justification here anyway, to inform any future attempt at resolving the discrepancy.

The agent again imagines the event A to be one trial in a series of Bernoulli trials with probability of success Π. We formulate probability $\Pr(\Pi = \pi)$ indirectly in terms of log-odds $\Pr\left[\ln\left(\frac{\Pi}{1-\Pi}\right) = \ln\left(\frac{\pi}{1-\pi}\right)\right]$. The agent's prior log-odds is normally distributed with mean q_1 and precision ρ_1, where precision is the reciprocal of variance. That is,

$$\Pr\left[\ln\left(\frac{\Pi}{1-\Pi}\right) = \chi\right] \propto e^{-\frac{1}{2}\rho_1(\chi-q_1)^2},$$

where $\chi \equiv \ln\left(\frac{\pi}{1-\pi}\right)$. The precision ρ_1 is interpreted as a subjective measure of the agent's confidence in its assessment, analogous to the number of observed samples in Section 4.1.1. A second precision parameter ρ_2 captures the agent's appraisal of the *market's* confidence. Denote P as the random variable for price, and reserve p for its value. Given that the true log-odds $\ln\left(\frac{\Pi}{1-\Pi}\right)$ equals χ, the log-odds of price P is normally distributed with mean χ and precision ρ_2:

$$\Pr\left[\ln\left(\frac{P}{1-P}\right) = q_2 \middle| \ln\left(\frac{\Pi}{1-\Pi}\right) = \chi\right] \propto e^{-\frac{1}{2}\rho_2(q_2-\chi)^2}.$$

where $q_2 \equiv \ln\left(\frac{p}{1-p}\right)$. The agent's posterior log-odds of Π given the price is a multiplication of the two normal distributions. The result is another normal distribution with mean

$$\frac{\rho_1}{\rho_1+\rho_2}\cdot q_1 + \frac{\rho_2}{\rho_1+\rho_2}\cdot q_2.$$

As before, the agent reasons only one period ahead, so presumably calculates the probability of A as the expected value of Π. And, for consistency, price is the expected value of P. Imagine for the moment, however, that the agent instead calculates the *log-odds* of A as the expected *log-odds* of Π:

$$\ln\left(\frac{\Pr(A)}{1-\Pr(A)}\right) = E\left[\ln\left(\frac{\Pi}{1-\Pi}\right)\right], \tag{4.4}$$

and the log-odds of p as the expected log-odds of P. Under these translations, we find that (4.3) is in fact justified, where the weighting term, $w = \rho_1/(\rho_1+\rho_2)$, is naturally construed as a measure of relative confidence. Nevertheless, I cannot conceive of a rationalization whereby the agent should translate from Π to A using (4.4) rather than using $\Pr(A) = E[\Pi]$. Generically, the two procedures will *not* return the same result.

4.2 Wealth Dynamics and Market Adaptation

We have seen that markets can mimic both the LinOP and LogOP formulas, and that simple learning from prices can lead to confidence-based weights. These results highlight the market's flexibility, showing that popular methods for deriving consensus are in fact special cases of the market mechanism. More heterogeneous mixtures of agents entail still other pooling functions, most of which are unlikely to admit closed-form expression. Yet this begs the question of which aggregation rule is the most appropriate and/or accurate. In contrived scenarios, one or the other is clearly warranted. For example, if the agents' beliefs are distributed such that their mean is the "true" probability, then the LinOP with equal weights will recover it.[1] Similarly, if the agents' log-odds are distributed such that their mean is the true log-odds, then the LogOP is justified. Of course, it is unreasonable to assume knowledge of the distribution of agent beliefs in relation to the true probability. In this section, I argue that, over time, the market may implicitly *learn* this distribution. Markets are not only flexible but can be *adaptive*, rewarding accuracy and punishing inaccuracy. For example, in an economy of decreasingly risk averse agents, those with better probabilistic models will gain greater influence on prices over time, as their relative wealth (and, consequently, their risk tolerance) increases.

I report simulation results in support of this claim. The market consists of a single

[1]This is precisely the construction of the comparative simulation experiments of Ng and Abramson (1992); in this light, their finding that the LinOP is best is perhaps not surprising.

security traded repeatedly over the course of T *periods*, or rounds. With only one security per period, we can let the subscript on events A_t and the superscripts on price $p^{\langle t\rangle}$ and demand $x^{\langle t\rangle}$ index time. Unbeknownst to the agents, each event unfolds as an independent Bernoulli trial with probability of success π. At the beginning of time period t, the realization of A_t is unknown and agents trade in $\langle A_t\rangle$ until equilibrium. Then the outcome—either A_t or $\bar{A}_t$—is revealed, and the agents' holdings pay off accordingly. As time period $t+1$ begins, the outcome of A_{t+1} is uncertain. Agents trade in the $t+1$ period security until equilibrium, the event outcome is revealed, payoffs are collected, and the process repeats.

Agents are *myopic*, or make decisions as if the current time period were the final one, maximizing expected utility (3.1) for each period in isolation. They do not update their beliefs based on observations. Discussion is again organized according to agent type.

4.2.1 Logarithmic Utility and the Bayesian Composite Agent

Recall the generalized logarithmic utility (GLU) form: $u(\mu) = \ln(\mu + b)$. As discussed in Section 3.3.2, the parameter b can be construed as the agent's initial wealth in dollars. Any accumulated gain or loss is naturally reincorporated as an addition or subtraction, respectively, to this wealth parameter. Rewrite initial wealth as $b^{\langle 0\rangle}$, and let $b^{\langle t\rangle}$ indicate the agent's wealth after period t ends. Negative wealth entails infinite disutility, so, at any given time t, the agent will never demand a quantity such that more than $b^{\langle t\rangle}$ of its dollars are at risk. All else being equal, the wealthier of two agents demands a proportionately larger quantity of the security, as dictated by (3.6). In an economy composed solely of agents with GLU, the equilibrium price is a wealth-weighted average (3.8). Thus, as an agent accrues relatively more earnings than the others, its influence on price increases. In the next two subsections, we examine how this adaptive process unfolds; first, when agents' beliefs are fixed and, second, when agents learn from price. In the former case, prices actually react as if the market were a single agent updating according to Bayes' rule.

Fixed beliefs

Figure 4.1 plots the price of one security over fifty time periods, in a market composed of one hundred agents, each with GLU, initial wealth $b_i^{\langle 0\rangle} = 1$, and $\Pr_i(A)$ generated randomly and uniformly on $(0, 1)$. In this simulation $\pi = 0.5$. For comparison, the figure

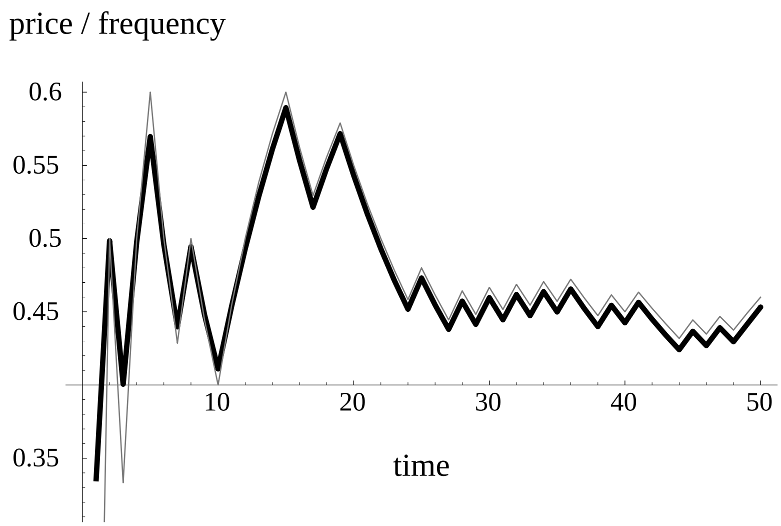


Figure 4.1: Price (black line) versus the observed frequency (gray line) of the event over fifty time periods. The market consists of one hundred fixed-belief agents with GLU and $b_i^{\langle 0 \rangle} = 1$.

also shows the *observed frequency*, or the number of times that A has occurred divided by the number of periods. The market price tracks the observed frequency extremely closely. Note that price changes are due entirely to a transfer of wealth from inaccurate agents to accurate agents, who then wield more power in the market; individual beliefs remain fixed.

Figure 4.2 illustrates the nature this transfer of wealth. The graph provides a snapshot of agents' wealth $b_i^{\langle 15 \rangle}$ versus their belief $\Pr_i(A)$ after period 15. In this run, A has occurred in three out of the fifteen trials. The maximum in wealth obtains at a belief near 3/15. The solid line in the figure is a Beta distribution with parameters $3+1$ and $12+1$, scaled only by a constant. This distribution is precisely the posterior probability of success that results from the observation of 3 successes out of 15 independent Bernoulli trials, when the prior probability of success is uniform on (0,1). The fit is essentially perfect, and can be proved in the limit.

Proposition 4.1 (Market Beta) *Suppose that agents with GLU, identical initial wealth $b_i^{\langle 0 \rangle} = b$, and $\Pr_i(A)$ generated randomly and uniformly on $(0,1)$, trade in a single-security market over n time periods. For notational brevity, let $\pi_i \equiv \Pr_i(A)$. Then for a sufficiently large number of agents N, the ratio of any two agents' wealth is*

$$\frac{b_h^{\langle n \rangle}}{b_i^{\langle n \rangle}} \approx \frac{(\pi_h)^s (1-\pi_h)^{n-s}}{(\pi_i)^s (1-\pi_i)^{n-s}}, \tag{4.5}$$

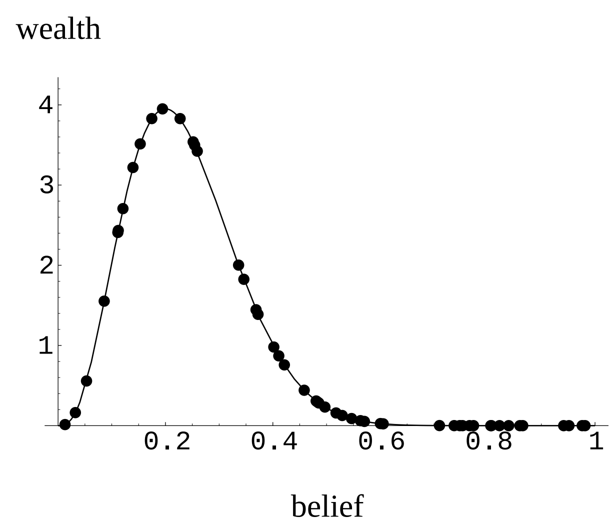


Figure 4.2: Wealth $b_i^{\langle 15 \rangle}$ after fifteen time periods versus belief $\Pr_i(A)$ for fifty agents with GLU and $b_i^{\langle 0 \rangle} = 1$. The event has occurred in three of the fifteen trials. The solid line is the posterior Beta distribution consistent with the observation of three successes in fifteen independent Bernoulli trials.

where s is the number of positive occurrences of A observed in the n periods.

Proof. We proceed by induction. The base case $n = 0$ holds by construction. Assume that (4.5) holds for $n = t$. Therefore, agent h's wealth after period t is $b_h^{\langle t \rangle} = \kappa(\pi_h)^s(1-\pi_h)^{t-s}$, where κ is independent of h. The price $p^{\langle t+1 \rangle}$ is a weighted sum of beliefs $(\sum_{i=1}^N \pi_i b_i^{\langle t \rangle})/\sum_{l=1}^N b_l^{\langle t \rangle}$ as prescribed by (3.8). Because beliefs π_i are generated uniformly on (0,1), as $N \to \infty$, this summation approaches

$$p^{\langle t+1 \rangle} \to \int_0^1 \pi \cdot \frac{\kappa(\pi)^s(1-\pi)^{t-s}}{\int_0^1 \kappa(u)^s(1-u)^{t-s}du} d\pi = E[\mathtt{Beta}(s+1, t-s+1)] = \frac{s+1}{t+2}.$$

The agent's demand for $\langle A_{t+1} \rangle$ is according to (3.4) with $S = 1$. We substitute the $t+1$ price into this demand equation and compute the agent's revised wealth $b_h^{\langle t+1 \rangle}$ after appropriate payoffs under both possible outcomes:

Case A_{t+1}**.** After a positive outcome, the agent's new wealth is $b_h^{\langle t+1 \rangle} = \kappa(\pi_h)^{s+1}(1 - \pi_h)^{t-s}(1/p^{\langle t+1 \rangle})$, which is proportional to a Beta distribution reflecting one additional successful trial than at time t. The rightmost factor depends only on price, and will be identical for all agents.

Case $\bar{A}_{t+1}$**.** Revised wealth is $b_h^{\langle t+1 \rangle} = \kappa(\pi_h)^s(1-\pi_h)^{t-s+1}(1/(1-p^{\langle t+1 \rangle}))$, or proportional to a Beta distribution reflecting an additional trial without success.

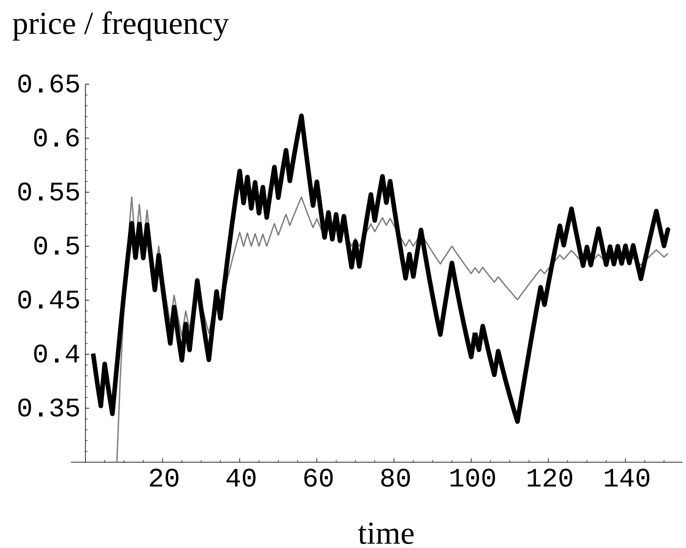


Figure 4.3: Price (black line) versus observed frequency (gray line) over one hundred fifty time periods for one hundred agents with GLU and learning parameter $w = 0.2$. As the frequency converges to $\pi = 0.5$, the price remains volatile.

The period $t+1$ wealth of agent i is analogous, and so (4.5) holds for $n = t+1$. □

Although individual agents are not adaptive, the market's composite agent appears to compute a proper Bayesian update. Specifically, wealth is reallocated proportionally to a Beta distribution corresponding to the observed number of successes and trials, and price is approximately the expected value of this Beta distribution.[2] Moreover, this correspondence holds regardless of the number of successes or failures, or the temporal order of their occurrence. A kind of collective Bayesianity *emerges* from the interactions of a group of myopic, fixed-belief agents subject to the reward structure of a securities market.

Learning from Price

In this section, we consider agents with GLU that *learn* from the price according to the weighted average update rule (4.1) derived in Section 4.1.1. Figure 4.3 graphs the dynamics of price in an economy of of such agents, along with the observed frequency. Over time, the price remains significantly more volatile than the frequency, which converges toward $\pi = 0.5$. Below, we characterize the transfer of wealth that precipitates this added volatility; for now concentrate on the price signal itself. Inspecting Figure 4.3,

[2]As t grows, this expected value rapidly approaches the observed frequency plotted in Figure 4.1.

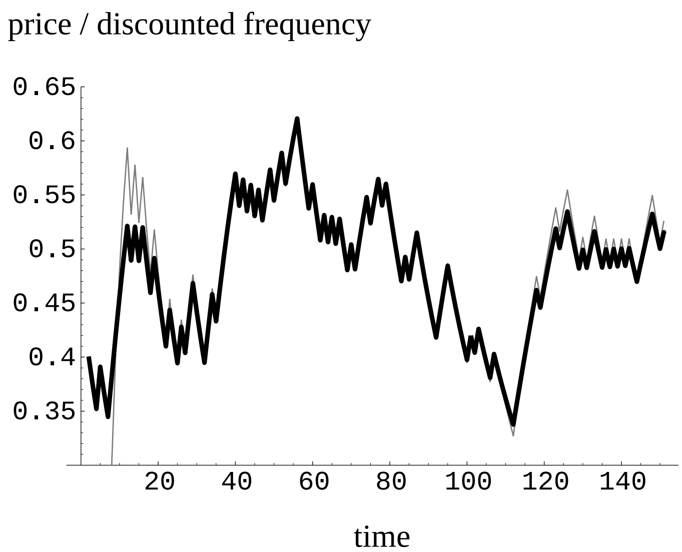


Figure 4.4: Price (black line) versus discounted frequency (gray line), with discount factor $\gamma = 0.96$, for the same experiment as Figure 4.3.

price changes still exhibit a marked dependence on event outcomes, though at any given period the effect of recent history appears magnified, and the past discounted, as compared with the observed frequency. Working from this intuition, I attempt to fit the data to an appropriately modified measure of frequency. Define the *discounted frequency* at period n as

$$d_n = \frac{\sum_{t=1}^{n} \gamma^{n-t}(1_{A_t})}{\sum_{t=1}^{n} \gamma^{n-t}(1_{A_t}) + \sum_{t=1}^{n} \gamma^{n-t}(1_{\bar{A}_t})}, \tag{4.6}$$

where 1_{A_t} is the indicator function for the event at period t, and γ is the *discount factor*. Note that $\gamma = 1$ recovers the standard observed frequency.

Figure 4.4 illustrates a very close correlation between discounted frequency, with $\gamma = 0.96$ (hand tuned), and the same price curve of Figure 4.3. While standard frequency provides a provably good model of price dynamics in an economy of fixed-belief agents with GLU (Proposition 4.1), discounted frequency (4.6) appears an excellent model for GLU agents that learn from price.

Prices are dependent on the apportionment of wealth among agents and also the agents' *posterior* beliefs. Figures 4.5, 4.6, and 4.7 display the dependence of wealth on posterior belief $\Pr_i(A|p^{\langle t \rangle})$ at periods $t = 9$, $t = 70$, and $t = 112$, respectively, from the same experiment as in Figure 4.3. For comparison, each figure also shows the distribution $\texttt{Beta}(s+1, t-s+1)$ appropriate for the number of successes s observed so far in t trials. While prior beliefs are distributed uniformly on $(0, 1)$, the update rule (4.1) serves to compress posterior beliefs into a smaller range centered around the price. Notice that

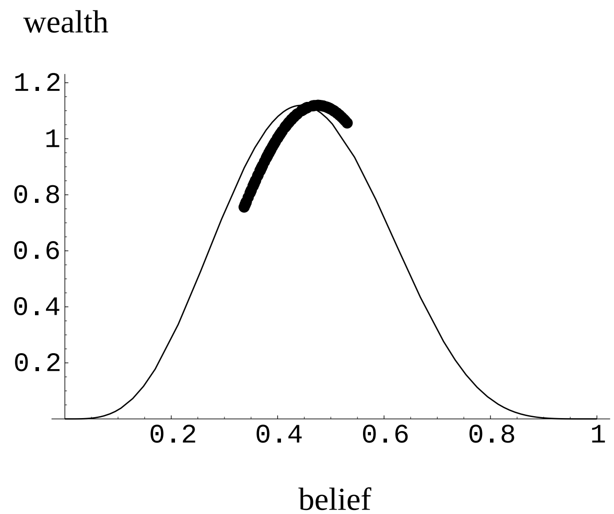


Figure 4.5: Wealth $b_i^{\langle 9 \rangle}$ versus belief $\Pr_i(A|p^{\langle 9 \rangle})$ at period nine of the same experiment as Figure 4.3. The observed frequency is $4/9$ and the solid line is $\texttt{Beta}(4+1, 5+1)$. The shapes of the two distributions match, though wealth is concentrated at the center and shifted right.

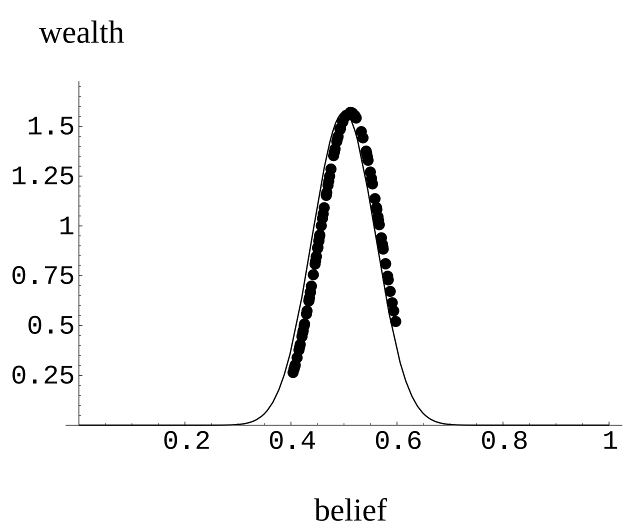


Figure 4.6: Wealth $b_i^{\langle 70 \rangle}$ versus belief $\Pr_i(A|p^{\langle 70 \rangle})$. The observed frequency is $35/70$ and the solid line is $\texttt{Beta}(35+1, 35+1)$. Within the window of posterior beliefs, the two distributions are well aligned.

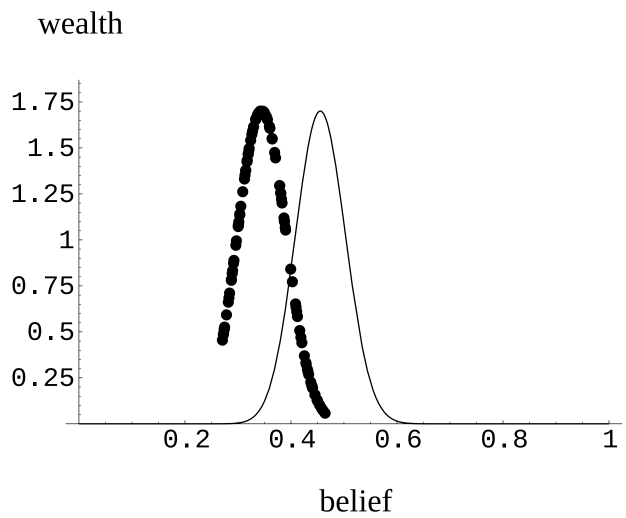


Figure 4.7: Wealth $b_i^{\langle 112 \rangle}$ versus belief $\Pr_i(A|p^{\langle 112 \rangle})$. The observed frequency is 51/112 (0.455) and the solid line is $\texttt{Beta}(51+1, 61+1)$. The shapes match, but wealth is shifted significantly left, explaining the disparity between frequency and price (0.359) in Figure 4.3.

learning from prices is of most benefit to the highly inaccurate agents (in this case, near $\Pr_i(A) = \{0, 1\}$), as might be expected. Even the worst agents fare reasonably well as compared to the fixed-belief case (Figure 4.2), where their fortunes rapidly converge to zero. Within the reduced range of posterior beliefs, the *shape* of the distribution of wealth mirrors the shape of the Beta distribution, although the position of the wealth curve may appear shifted to the left or right. The magnitude and direction of this shift correlates with features in Figure 4.3. For example, when the wealth distribution is shifted significantly left, as in Figure 4.7 (period 112), the price is considerably less than the observed frequency. In Figure 4.6 (period 70), the two distributions are well aligned, and price and frequency are similar.

The close fit of price to a discounted frequency can probably *not* be modeled as if the market's composite agent "forgets" long past observations, since the shape of the wealth distribution does seem to reflect all previous outcomes. Increased volatility in price arises instead from the horizontal translations evident in Figures 4.5, 4.6, and 4.7.

4.2.2 Exponential Utility

The demand function for an agent with CARA (3.4) does not depend on its wealth. In an economy of such agents with fixed beliefs, price remains constant from period to period. Nevertheless, it is instructive to examine how wealth is redistributed among the

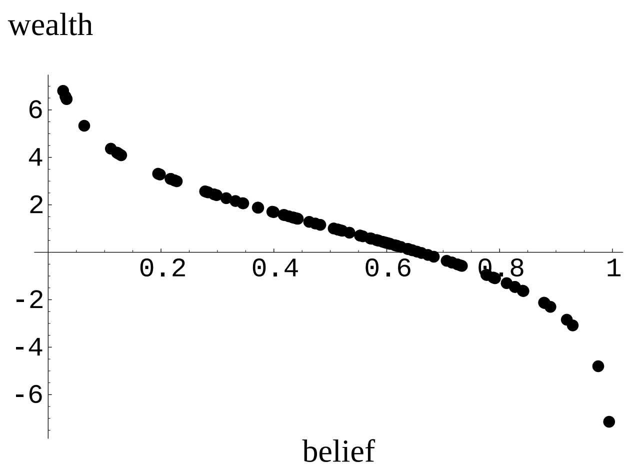


Figure 4.8: Wealth $b_i^{\langle 15 \rangle}$ after fifteen time periods versus belief $\Pr_i(A)$ for one hundred agents with CARA and $b_i^{\langle 0 \rangle} = 1$. At period 15, the price (0.507) is greater than the observed frequency (6/15), so the agents that sold most fared best.

agents over time. One might expect that wealth would again accumulate among the most accurate agents. Figures 4.8 and 4.9 suggest otherwise. In the former, the agents with the smallest probability assessments fared best; in the latter, the agents with the largest assessments gained the most. It turns out that winners at any given period are determined simply based on whether the price is greater than or less than the frequency of successes up to that time. If the price is greater, then big sellers win; conversely, if the price is smaller, big buyers win. Agents with beliefs near 0 tend to sell while agents with beliefs near 1 tend to buy. The difference in the reward structure for markets of agents with CARA as compared to markets of agents with GLU likely relates to the difference between improper and proper scoring rules for eliciting probabilities (Winkler, 1968).

4.2.3 Mixed Populations

Figure 4.10 graphs wealth versus belief for two subpopulations within a mixture of fifty agents of each type. It appears that the relationships seen in homogeneous populations in Sections 4.2.1 and 4.2.2 carry over to the mixed case. The relative wealth among the subgroup of GLU agents fits the appropriate Beta distribution, even though the *absolute* fortune of the agents with GLU depends on those with CARA, and even though the price at each period no longer matches the expectation of that Beta distribution. The wealth of the subgroup of agents with CARA is skewed toward the buyers, since the price in the displayed period is less than the observed frequency.

Figure 4.11 displays the price and observed frequency over time in a market of fifty of

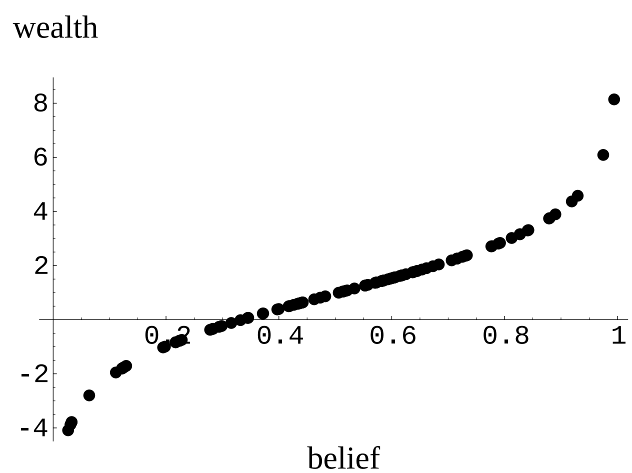


Figure 4.9: Setup identical to Figure 4.8, with different realizations of outcomes. At period 15, the price (0.507) is less than the observed frequency (9/15), so the agents that bought most fared best.

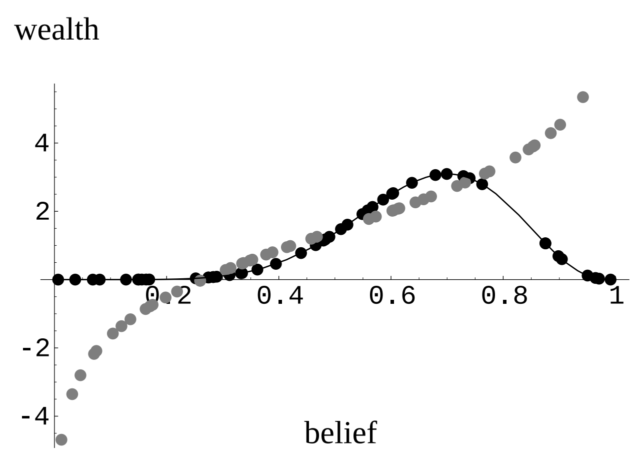


Figure 4.10: Wealth $b_i^{\langle 10 \rangle}$ after ten time periods versus belief $\Pr_i(A)$ for fifty agents with CARA (gray points) and fifty with GLU (black points). The period 10 price (0.528) is less than the observed frequency (7/10). The solid line is $\texttt{Beta}(7+1, 3+1)$.

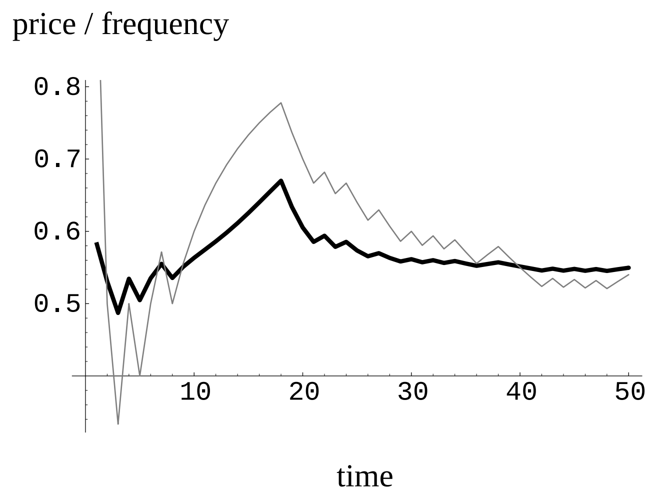


Figure 4.11: A single security's price (black line) and the observed frequency (gray line) of the event over fifty time periods. The market consists of fifty fixed-belief agents with GLU and fifty with CARA.

each type of agent. We see that the price does follow the frequency, but not as closely as in Figure 4.1. Close inspection reveals that the *relative* proportion of changes are similar for both price and observed frequency. All of the fluctuations in price are due to wealth redistribution among the agents with GLU. The presence of agents with CARA serves to shift the absolute price and dampen the magnitude of changes.

4.2.4 Implementation Issues

Simulation results in Sections 4.2.1, 4.2.2, and 4.2.3 were generated in *Mathematica*. Equilibrium prices are computed directly using the closed-form solutions (3.5) and (3.8) for experiments involving homogeneous populations of CARA and GLU agents, respectively. For the mixed population experiments, agents conceptually submit their closed-form demand functions, (3.3) or (3.6), and price is computed numerically as the unique zero of $\sum_i x_i(p)$. The implementation conforms almost exactly to the analytic model, including arbitrarily divisible securities (e.g., an agent might purchase 3.1415926 units of the security).

I have also implemented the model via a C++ interface to the *Michigan Internet AuctionBot* (`http://auction.eecs.umich.edu/`), which provides fairly general facilities to set up and run distributed computational markets. The AuctionBot supports a web interface for human users and a TCP interface and associated API for computer agents (Wurman et al., 1998b). My code consists of three C++ classes: `agent`, `market`, and

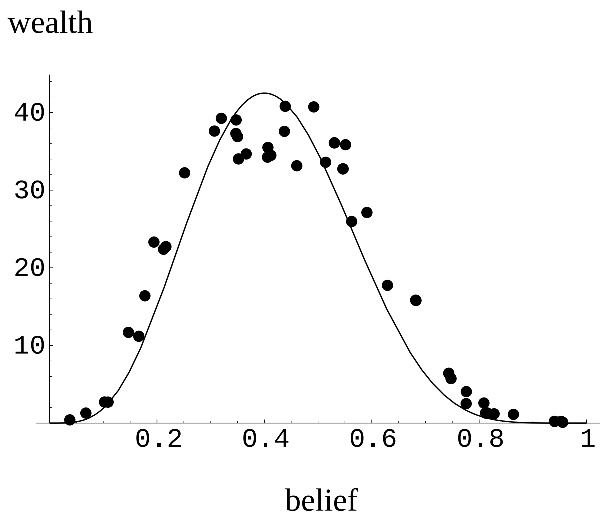


Figure 4.12: Wealth $b_i^{\langle 10 \rangle}$ versus belief $\Pr_i(A)$ for fifty agents with GLU and $b_i^{\langle 0 \rangle} = 20$. The fit to a Beta distribution is imperfect due to discretization inherent in the AuctionBot implementation.

`experiment`, and some utilities to make the API calls. I have attempted to design the program in a modular fashion, with an eye toward flexibility and future extensibility. Currently, experiments are orchestrated from a single process on a single machine, although the same basic code could support physically separated and/or proprietary agents designed by separate owners.

Few real markets accept closed-form demand functions as bids, and most restrict trades to discrete quantities. In the AuctionBot implementation, agents submit approximately twenty price–quantity bid *points* of the form $(p, x(p))$, rather than a full curve. Securities can be bought and sold only in integer quantities. The price of each security is determined according to the *(M+1)st price auction* (Wurman et al., 1998a). This auction protocol specifies the price as the $(M+1)$st highest bid among all agents (both buyers and sellers), where M is the number of units of the security offered for sale. The AuctionBot currently does not support larger numbers of bid points or continuous-good auctions, though plans to do so are underway.

All of the features of wealth and price dynamics reported in Sections 4.2.1, 4.2.2, and 4.2.3, are evident in AuctionBot experiments, as exemplified in Figures 4.12, 4.13, and 4.14. However, due to the discretization of bids and quantities, the AuctionBot results are quite a bit noisier.

The AuctionBot offers a truly distributed platform for empirical studies. Communication between agents and the auction must conform to the API protocol and are conducted

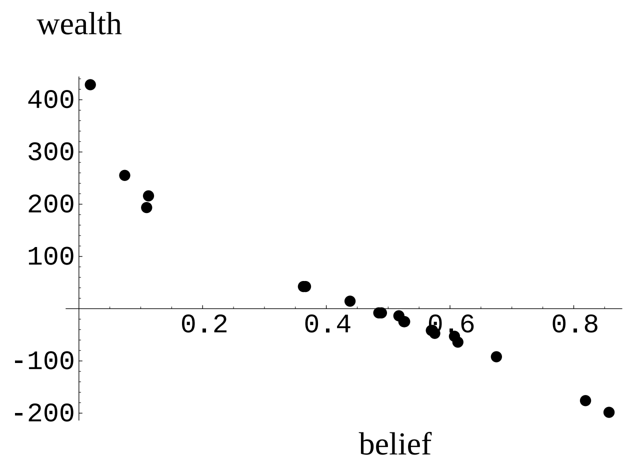


Figure 4.13: Results from AuctionBot. Wealth $b_i^{\langle 30 \rangle}$ versus belief $\Pr_i(A)$ for twenty agents with CARA and $b_i^{\langle 0 \rangle} = 20$.

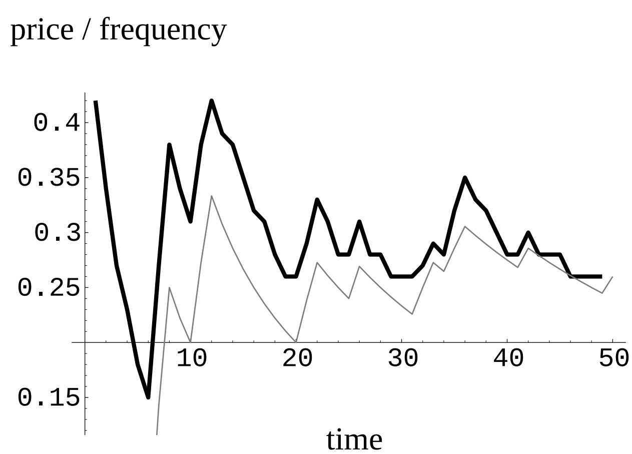


Figure 4.14: [Price (black line) versus observed frequency (gray line) over fifty time periods for fifty agents with GLU. The correspondence is weaker than that displayed in Figure 4.1, due to discretization effects.

over TCP connections, subject to real network delays. In such an environment, even detecting equilibrium becomes nontrivial (Wellman and Walsh, 1999). I employed a simple constant waiting time—reasonable for the single-security case. Waiting times dominated all computational considerations. For example, a one hundred period experiment, with a ninety second wait per period, requires over two and a half hours to complete. For comparison, the same simulation executes within several seconds in *Mathematica*. The *Mathematica* implementation proved largely sufficient for exploring a direct multiperiod extension to the model of Chapter 3. On the other hand, the AuctionBot implementation is a much more realistic foundation on which to build a *real* system for aggregating the beliefs of self-interested and proprietary agents.

Chapter 5

Graphical Representations of Aggregate Belief

In this chapter, we put aside the market framework for the moment, in order to address the problem of representing aggregate beliefs concisely. The implications of results in this chapter for securities markets will be examined in Chapter 7.

We presume that each agents' beliefs are given as a graphical model, and that the combined beliefs are to be represented as well with a graphical model. Two intuitively reasonable assumptions in this context, made *a priori* by other authors, are (1) if all agents agree on a single topology, then that structure should be maintained, and (2) probability aggregation can be isolated within each conditional probability table (CPT). Section 5.2 demonstrates that each of these properties leads to an impossibility theorem when combined with other reasonable, oft-invoked assumptions. On a more positive note, Section 5.3 shows that the logarithmic opinion pool (LogOP) maintains all agreed-upon Markov independencies, and describes procedures for constructing consensus Markov networks (MNs) and consensus Bayesian networks (BNs) that are consistent with the LogOP. Section 5.4 presents an algorithm that can, in some cases, compute the LogOP exponentially faster than the brute force approach. That section also characterizes the computation complexity of the linear opinion pool (LinOP). Section 5.1 defines the properties necessary to state the various impossibility and possibility results appearing later in the chapter.

5.1 Property Definitions

Recall the independence preservation property (IPP), defined in Section 2.3.1. For an aggregation function f to satisfy IPP, any independencies that are agreed-upon by all agents must be maintained within the consensus distribution. Genest and Wagner (1987) prove that *no* aggregation function can simultaneously satisfy IPP, proportional dependence on states (PDS) (Property 2.5), and nondictatorship (ND) (Property 2.2). But one might argue that IPP is overly strong. It requires preservation of, for example, a unanimous independence between the events $E = A_3\bar{A}_7$ and $F = \bar{A}_2A_4 \vee A_7$. This kind of independence seems of little descriptive value to a modeler, and indeed cannot be represented with a BN. One may be willing to forgo preserving *all* independencies, being content to preserve independencies among the *primary* events, $A_1, A_2, \ldots, A_M$. With this in mind, I define a weaker independence property.

Property 5.1 (Event independence preservation property (EIPP)) *If* $\Pr_i(A_j|A_k) = \Pr_i(A_j)$ *for all agents i, then* $\Pr_0(A_j|A_k) = \Pr_0(A_j)$*.*

Note that, when $|\Omega| = 4$, the conditions IPP and EIPP are essentially equivalent. In this situation, the only way for two events to be independent is if each consists of exactly two atomic states, and if they overlap at exactly one state (Genest and Wagner, 1987).

In Section 5.2, we see that substituting EIPP for IPP does admit a possibility that is consistent with both PDS and ND, though not a very satisfactory one. In search of a nontrivial possibility, I define two even weaker independence conditions.

Property 5.2 (Markov event independence preservation property (MEIPP)) *If* $\Pr_i(A_j|WA_k) = \Pr_i(A_j|W)$ *for all agents i and for all $W \subseteq Z$ (including $W = \emptyset$), then* $\Pr_0(A_j|A_k) = \Pr_0(A_j)$*.*

Property 5.3 (Non-Markov event independence preservation property (NMEIPP)) *If* $\Pr_i(A_j|A_k) = \Pr_i(A_j)$ *for all agents i, and* $\Pr_h(A_j|WA_k) \neq \Pr_h(A_j|W)$*, for some agent h and some $W \subseteq Z$, then* $\Pr_0(A_j|A_k) = \Pr_0(A_j)$*.*

These two properties are purposely constructed so that EIPP $\Leftrightarrow$ (MEIPP $\wedge$ NMEIPP). We see in Section 5.2 that the source of the impossibility lies entirely within the latter. Finally, I define a stronger version of the MEIPP.

Property 5.4 (Markov independence preservation property (MIPP)) *Let $W, X \subseteq Z - A_j$ be disjoint sets of events such that $A_j \cup W \cup X = Z$. If* $\Pr_i(A_j|WX) = \Pr_i(A_j|W)$ *for all agents i, then* $\Pr_0(A_j|WX) = \Pr_0(A_j|W)$*.*

The relative strengths of these various independence conditions can be summarized as follows:

$$\text{IPP} \Rightarrow \text{EIPP} \Leftrightarrow (\text{MEIPP} \wedge \text{NMEIPP})$$

$$\text{MIPP} \Rightarrow \text{MEIPP}$$

Finally, I define a property that captures what seems to be a natural assumption within the context of graphical models, advocated independently by other authors (Matzkevich and Abramson, 1992). We say that an aggregator satisfies the *family aggregation* (FA) property if it operates locally, within each conditional probability table (CPT) of the consensus structure.

Property 5.5 (Family aggregation (FA))

$$\begin{aligned}&\Pr_0(A_j|\mathbf{pa}(A_j)) = \\ &\qquad f(\Pr_1(A_j|\mathbf{pa}(A_j)), \ldots, \Pr_N(A_j|\mathbf{pa}(A_j))).\end{aligned}$$

In Sections 5.2.1 and 5.2.2, we consider the implications of the properties EIPP and FA, respectively.

5.2 Combining Bayesian Networks: Examples and Impossibility

5.2.1 Event Independence Preservation

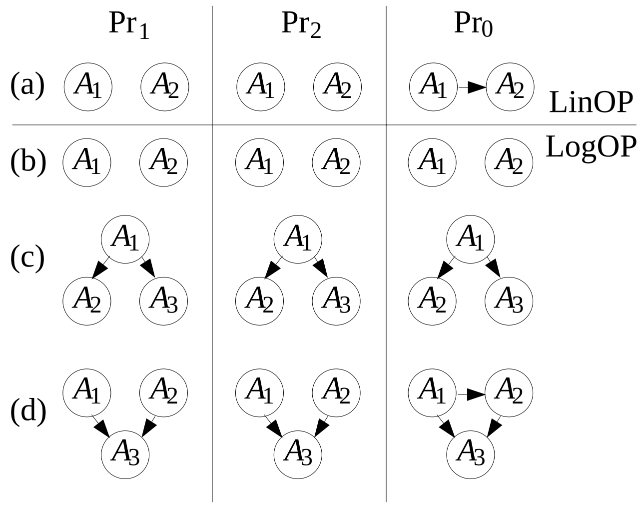
We begin with an example the build the underlying intuition.

Example 5.1 (EIPP and the LinOP)

Suppose that two agents agree that two primary events, A_1 and A_2, are independent, as pictured in Figure 5.1(a), but disagree on the associated marginal probabilities.

For concreteness, let the first agent hold beliefs $\Pr_1(A_1) = \Pr_1(A_2) = 0.5$, and the second $\Pr_2(A_1) = 0.8$ and $\Pr_2(A_2) = 0.6$. Thus,

$$\begin{array}{ll}\Pr_1(A_1A_2) = 0.25 & \Pr_2(A_1A_2) = 0.48 \\ \Pr_1(A_1\bar{A}_2) = 0.25 & \Pr_2(A_1\bar{A}_2) = 0.32 \\ \Pr_1(\bar{A}_1A_2) = 0.25 & \Pr_2(\bar{A}_1A_2) = 0.12 \\ \Pr_1(\bar{A}_1\bar{A}_2) = 0.25 & \Pr_2(\bar{A}_1\bar{A}_2) = 0.08.\end{array}$$



1

Figure 5.1: Independence preservation behavior of (a) LinOP and (b)–(d) LogOP. If two agents' beliefs $\Pr_1$ and $\Pr_2$ have the dependency structures shown, then the consensus $\Pr_0$ will in general have the dependency structure depicted in column three.

Now if we apply the LinOP (2.5) with, say, equal weights of $w_1 = w_2 = 0.5$, we get:

$$\Pr_0(A_1A_2) = 0.365$$
$$\Pr_0(A_1\bar{A}_2) = 0.41$$
$$\Pr_0(\bar{A}_1A_2) = 0.185$$
$$\Pr_0(\bar{A}_1\bar{A}_2) = 0.165.$$

In particular, $\Pr_0(A_1)\Pr_0(A_2) \neq \Pr_0(A_1A_2)$, and so the two events are *not* independent in the consensus.[1] Even though the precondition of the EIPP is met, the postcondition is not: a BN representation of the derived consensus would have to include an edge between A_1 and A_2. $\square$

Example 5.2 (EIPP and the LogOP)

Suppose that two agents' beliefs over two primary events are as described in Example 5.1. If we apply the LogOP with equal weights, we get:

$$\Pr_0(A_1A_2) = 0.367007$$
$$\Pr_0(A_1\bar{A}_2) = 0.29966$$

[1]As early as Yule (1903) it was recognized that averaging two distributions may mask a commonly held independence.

$$\mathrm{Pr}_0(\bar{A}_1 A_2) = 0.183503$$
$$\mathrm{Pr}_0(\bar{A}_1 \bar{A}_2) = 0.14983.$$

In this case, $\mathrm{Pr}_0(A_1)\,\mathrm{Pr}_0(A_2) = \mathrm{Pr}_0(A_1 A_2)$, and the two events remain independent, as shown in Figure 5.1(b). This is not a numerical coincidence; in fact, independence between only two events is always maintained by the LogOP (Genest and Wagner, 1987). Now suppose that among three primary events, both agents agree that A_3 is independent of A_2 given A_1. That is, both agents agree that dependencies conform to a tree structure, with A_1 the parent of both A_2 and A_3, as depicted in Figure 5.1(c). Then once again, the LogOP will maintain this structure. One might conjecture that the LogOP maintains all BN structures, but this is not the case. For example, suppose that, among three primary events, the two agents agree that A_1 and A_2 are mutually independent, and that A_3 depends on both A_1 and A_2. That is, both agents agree on the polytree structure in Figure 5.1(d). In this case, when we compute the consensus with the LogOP, A_1 and A_2 will in general become mutually dependent, the EIPP is not satisfied, and a consensus BN will require an arc between the two nodes. □

Having seen that both the LinOP and the LogOP violate the EIPP, we seek a more general characterization of the class of functions that do obey it. We begin by showing that Lemma 3.2 in (Genest and Wagner, 1987), originally proved with respect to the IPP, is also applicable under the weaker EIPP.

Lemma 5.1 (Adapted from (Genest and Wagner, 1987)) *If f obeys EIPP and PDS, then there exist constants $\alpha_1, \alpha_2, \ldots, \alpha_N$, and c such that*

$$\mathrm{Pr}_0(\omega_j) = \sum_{i=1}^{N} \alpha_i \mathrm{Pr}_i(\omega_j) + c. \tag{5.1}$$

Proof. Consider three events A_1, A_2, and A_3, with agents' beliefs described as follows:

$$\begin{aligned}
\mathrm{Pr}_i(A_1 A_2 A_3) = \mathrm{Pr}_i(A_1 A_2 \bar{A}_3) &= \frac{(1-z_i)^2}{4(1+z_i)} \\
\mathrm{Pr}_i(A_1 \bar{A}_2 A_3) = \mathrm{Pr}_i(A_1 \bar{A}_2 \bar{A}_3) &= \frac{1-z_i}{4} \\
\mathrm{Pr}_i(\bar{A}_1 \bar{A}_2 A_3) &= x_i \\
\mathrm{Pr}_i(\bar{A}_1 \bar{A}_2 \bar{A}_3) &= y_i,
\end{aligned} \tag{5.2}$$

where $z_i = x_i + y_i$ for all i. In this case, all agents agree that A_1 and A_2 are independent and, as long as $z_i < 1$, these equations describe a legal probability distribution. Since f obeys PDS, there must be some function g such that,

$$\Pr_0(\bar{A}_1\bar{A}_2A_3) = \frac{g(x_1, x_2, \ldots, x_N)}{\sum_{k=1}^{8} g(\Pr_1(\omega_k), \ldots, \Pr_N(\omega_k))}$$

and similarly for $\Pr_0(\bar{A}_1\bar{A}_2\bar{A}_3)$. Now imagine a second situation exactly as in (5.2), except with $\Pr_i(\bar{A}_1\bar{A}_2A_3) = x'_i$ and $\Pr_i(\bar{A}_1\bar{A}_2\bar{A}_3) = y'_i$. Genest and Wagner show that, as long as $x_i + y_i = x'_i + y'_i < 1$, then

$$\begin{aligned} & g(x_1, x_2, \ldots, x_N) + g(y_1, y_2, \ldots, y_N) \\ = \quad & g(x'_1, x'_2, \ldots, x'_N) + g(y'_1, y'_2, \ldots, y'_N). \end{aligned} \tag{5.3}$$

From here, they show that *since x_i and y_i can be chosen arbitrarily* (as long as their sum is less than one), then f must have the form specified. □

Genest and Wagner go on to show, without further assumption, that f must be a dictatorship. However, that proof does *not* carry through under the weaker condition EIPP. This can be seen via a simple counterexample. Let f always ignore the agents' opinions, and simply assign a uniform distribution over all $\omega \in \Omega$. In this case, the consensus distribution holds that *all* primary events A_j are independent, and thus any agreed upon independencies are trivially maintained. One might wonder whether EIPP admits any other, more appealing, aggregation functions. The following proposition essentially establishes that it does not.

Proposition 5.2 *No aggregation function f can simultaneously satisfy EIPP, PDS, UNAM, and ND.*

Proof. With the addition of UNAM, it is clear that c must be zero in (5.1), and thus f must have the form of a standard LinOP (2.5). From Example 5.1, we know that the LinOP does not maintain independence even between just two events. The fact that the LinOP cannot satisfy both IPP and ND is proved formally by several authors (Genest, 1984c; Lehrer and Wagner, 1983; Wagner, 1984). Their proofs are applicable to EIPP as well, since they hold even when $|\Omega| = 4$, in which case EIPP and IPP coincide. □

A careful examination of the proof of Lemma 5.1 also suggests one more possibility when the full generality of IPP is relaxed. Suppose that all agents agree that *all* *three* events, A_1, A_2, and A_3, are completely independent. Then it can be shown that $\Pr_i(\bar{A}_1 A_2 A_3) = z_i/(1+z_i) + y_i$ and, furthermore, that $x_i = y_i$ for all i. In this case, (5.3) holds only vacuously, since $x_i' = x_i$ and $y_i' = y_i$. Moreover, since x_i and y_i are no longer arbitrary, the proof does not go through. Thus, under this fully independent condition, the conclusion of Lemma 5.1 is no longer valid.

This insight leads us to characterize the inherent impossibility more sharply, by dividing EIPP into two, weaker conditions, NMEIPP and MEIPP, and showing that the former retains the impossibility while the latter does not.

Corollary 5.3 *No aggregation function f can simultaneously satisfy NMEIPP, PDS, UNAM, and ND.*

Proof. The proof of Lemma 5.1 still follows under NMEIPP, and thus so does the proof of Proposition 5.2. □

Section 5.3 demonstrates that in fact, MEIPP is perfectly consistent with PDS, UNAM, and ND in a nontrivial way. Indeed, the stronger MIPP is consistent as well.

5.2.2 Family Aggregation

Example 5.3 (Family aggregation)

Consider two agents, each with a BN consisting of two primary events, with A_1 the parent of A_2 and with beliefs as follows:

$$\begin{aligned} \Pr_1(A_1) &= 0.2 & \Pr_2(A_1) &= 0.8 \\ \Pr_1(A_2|A_1) &= 0.4 & \Pr_2(A_2|A_1) &= 0.8 \\ \Pr_1(A_2|\bar{A}_1) &= 0.6 & \Pr_2(A_2|\bar{A}_1) &= 0.3 \end{aligned}$$

We compute each consensus CPT as an average of the corresponding individual CPTs. That is, $\Pr_0(A_1) = (.2+.8)/2 = .5$, $\Pr_0(A_2|A_1) = (.4+.8)/2 = .6$, etc. This results in the following consensus joint distribution:

$$\begin{aligned} \Pr_0(A_1 A_2) &= 0.3 \\ \Pr_0(A_1 \bar{A}_2) &= 0.2 \\ \Pr_0(\bar{A}_1 A_2) &= 0.225 \\ \Pr_0(\bar{A}_1 \bar{A}_2) &= 0.275. \end{aligned}$$

Next suppose that both agents reverse their edge between the two events, such that A_2 is the parent of A_1, but that their joint distributions remain unchanged. Now the agents' CPTs are:

$$\begin{aligned} \Pr_1(A_2) &= 0.56 & \Pr_2(A_2) &= 0.7 \\ \Pr_1(A_1|A_2) &= 0.142857 & \Pr_2(A_1|A_2) &= 0.914286 \\ \Pr_1(A_1|\bar{A}_2) &= 0.272727 & \Pr_2(A_1|\bar{A}_2) &= 0.533333 \end{aligned}$$

and if we average locally within each CPT, we get a different consensus distribution:

$$\begin{aligned} \Pr_0(A_1 A_2) &= 0.333 \\ \Pr_0(A_1 \bar{A}_2) &= 0.149121 \\ \Pr_0(\bar{A}_1 A_2) &= 0.297 \\ \Pr_0(\bar{A}_1 \bar{A}_2) &= 0.220878. \end{aligned}$$

Thus averaging only within each family of the BN violates the form of the opinion pool itself (2.1), which insists that the consensus joint distribution depend only on the underlying joint distributions of the agents involved. □

We now show that this inconsistency is not confined solely to the averaging aggregator.

Proposition 5.4 *No aggregation function f can simultaneously satisfy FA, UNAM, and ND.*

Proof. Let the first event in the consensus BN be A_{j_1} the second A_{j_2}, ..., and the last A_{j_M}. The FA property requires both of the following:

$$\begin{aligned} &\Pr_0(A_{j_1}) \\ &\quad = \quad f(\Pr_1(A_{j_1}), \Pr_2(A_{j_1}), \ldots, \Pr_N(A_{j_1})) \qquad (5.4) \\ &\Pr_0(A_{j_M}|Z - A_{j_M}) \\ &\quad = \quad f(\Pr_1(A_{j_M}|Z - A_{j_M}), \ldots, \Pr_N(A_{j_M}|Z - A_{j_M})). \qquad (5.5) \end{aligned}$$

By the definition of an opinion pool (2.1), the consensus belief depends only on the agents' underlying joint distributions, and not on the particular ordering of events in each BN. Thus, we must arrive at the same consensus distribution as long as $\{j_1, j_2, \ldots, j_M\}$ is some permutation of $\{1, 2, \ldots, M\}$. Consider two permutations, one where $j_1 = 1$ and one where $j_M = 1$. Then (5.4) and (5.5) become:

$$\begin{aligned} &\Pr_0(A_1) \\ &\quad = \quad f(\Pr_1(A_1), \Pr_2(A_1), \ldots, \Pr_N(A_1)) \qquad (5.6) \\ &\Pr_0(A_1|Z - A_1) \\ &\quad = \quad f(\Pr_1(A_1|Z - A_1), \ldots, \Pr_N(A_1|Z - A_1)). \qquad (5.7) \end{aligned}$$

Dalkey (1975) proves that no function can simultaneously satisfy (5.6), (5.7), UNAM, and ND. Alternatively, the two equations essentially require that f satisfy both MP and EB, defined in Section 2.3.1, which Genest (1984b) shows are incompatible with UNAM and ND. □

5.3 The LogOP and Consensus Markov Networks

The results in Section 5.2 suggest that insisting upon general event independence preservation has rather severe consequences. In this section, we see that preserving *Markov* independencies is in fact compatible with PDS, UNAM, and ND. Let A_j be a primary event, and $W \subseteq Z - A_j$ and $X = Z - W - A_j$ be sets of events. Then A_j is Markov independent of X given W if $\Pr(A_j|WX) = \Pr(A_j|W)$.

Proposition 5.5 *The LogOP satisfies MIPP.*

Proof. Since the LogOP is defined in terms of atomic states ω, we make use of the following two identities:

$$\Pr_0(A|WX) \equiv \frac{\Pr_0(AWX)}{\Pr_0(AWX)+\Pr_0(\bar{A}WX)}$$
$$\Pr_0(A|W) \equiv \frac{\sum_X \Pr_0(AWX)}{\sum_X \Pr_0(AWX)+\sum_X \Pr_0(\bar{A}WX)}$$

where $\sum_X$ represents a sum over all possible combinations of outcomes of events in the set X. Then we have that,

$$
\begin{aligned}
\Pr_0(A|WX) &= \frac{\prod_{i=1}^N [\Pr_i(AWX)]^{\alpha_i}}{\prod_{i=1}^N [\Pr_i(AWX)]^{\alpha_i}+\prod_{i=1}^N [\Pr_i(\bar{A}WX)]^{\alpha_i}} \\
&= \frac{\prod\left[\frac{\Pr_i(AW)\Pr_i(WX)}{\Pr_i(W)}\right]^{\alpha_i}}{\prod\left[\frac{\Pr_i(AW)\Pr_i(WX)}{\Pr_i(W)}\right]^{\alpha_i}+\prod\left[\frac{\Pr_i(\bar{A}W)\Pr_i(WX)}{\Pr_i(W)}\right]^{\alpha_i}} \\
&= \frac{\prod[\Pr_i(AW)]^{\alpha_i}}{\prod[\Pr_i(AW)]^{\alpha_i}+\prod[\Pr_i(\bar{A}W)]^{\alpha_i}} \\
&= \frac{\prod[\Pr_i(AW)]^{\alpha_i}}{\prod[\Pr_i(AW)]^{\alpha_i}+\prod[\Pr_i(\bar{A}W)]^{\alpha_i}} \cdot \frac{\sum_X \prod[\Pr_i(WX)]^{\alpha_i}}{\sum_X \prod[\Pr_i(WX)]^{\alpha_i}} \\
&= \frac{\sum_X \prod[\Pr_i(AW)\Pr_i(WX)]^{\alpha_i}}{\sum_X \prod[\Pr_i(AW)\Pr_i(WX)]^{\alpha_i}+\sum_X \prod[\Pr_i(\bar{A}W)\Pr_i(WX)]^{\alpha_i}} \\
&= \frac{\sum_X \prod\left[\frac{\Pr_i(AW)\Pr_i(WX)}{\Pr_i(W)}\right]^{\alpha_i}}{\sum_X \prod\left[\frac{\Pr_i(AW)\Pr_i(WX)}{\Pr_i(W)}\right]^{\alpha_i}+\sum_X \prod\left[\frac{\Pr_i(\bar{A}W)\Pr_i(WX)}{\Pr_i(W)}\right]^{\alpha_i}} \\
&= \frac{\sum_X \prod[\Pr_i(AWX)]^{\alpha_i}}{\sum_X \prod[\Pr_i(AWX)]^{\alpha_i}+\sum_X \prod[\Pr_i(\bar{A}WX)]^{\alpha_i}} \\
&= \frac{\sum_X \Pr_0(AWX)}{\sum_X \Pr_0(AWX)+\sum_X \Pr_0(\bar{A}WX)} \\
&= \Pr_0(A|W)
\end{aligned}
$$

$\square$

Suppose that each agent's belief is given as a MN, and we wish to generate a consensus MN structure that can encode the results of the LogOP. As discussed in Section 2.3.4, graph connectivity in a MN represents probabilistic dependence, and the neighborhood relation represents direct influence. For each node A_j, the set of its neighbors plays the role of W in Proposition 5.5, and all other nodes constitute the set X. The proposition ensures that, if all agents agree on a common MN structure, then the consensus distribution derived by the LogOP will respect the same structure. When agents are not in complete agreement on the structure, then the consensus can be represented as a MN defined by the union of all the individual MNs. In other words, there is an edge between A_j and A_k in the consensus MN if and only if there is an edge between those two nodes in at least one of the agents' MNs.

Pearl (1988) gives axiomatic descriptions of both MNs and BNs. Only the former includes an axiom called *strong union*, which states that if $\Pr(A_j|A_k) = \Pr(A_j)$, then $\Pr(A_j|WA_k) = \Pr(A_j|W)$ for all $W \subseteq Z$. Notice that, if the precondition of the EIPP is met, and strong union holds for all agents, then the precondition of the MEIPP must also hold. This axiom is the key distinction that allows common MN structures to be maintained in the LogOP consensus, whereas common BN structures in general are not.

Given a collection of BNs, generating a consensus BN structure that is consistent with the LogOP is also relatively straightforward. We first convert each BN into a MN by moralizing the graphs, or fully connecting each node's parents and dropping edge directionality (Lauritzen and Spiegelhalter, 1988; Neapolitan, 1990). Next, we compute the union of the individual MNs, and finally we convert the resulting consensus MN back into a BN by filling in or triangulating the network, reintroducing directionality according to the fill-in order[2] (Jensen, 1996; Lauritzen and Spiegelhalter, 1988; Neapolitan, 1990; Pearl, 1988).

I have outlined how to derive consensus MN or BN *structures*; what of computing the associated probabilities? In Section 5.4, I give an algorithm for computing the probabilities in a consensus BN that is polynomial in the size of its CPTs. Note that, even when all agents agree on a BN structure, the size of the final representation may grow exponentially during fill-in, and computing the union of the intermediate MNs when agents disagree will only exacerbate this problem. Nevertheless, even a decomposable representation can be exponentially smaller than the full joint distribution, and the most popular algorithms for exact Bayesian inference do operate on decomposable models in practice.

5.4 Computing LogOP and LinOP

Since the LinOP (2.5) and LogOP (2.6) are defined over atomic states, computing, for example, the consensus marginal probability of a single event involves in the worst case a summation over 2^{M-1} terms. Moreover, even computing the LogOP consensus for a single state requires a normalization factor that is itself a sum over all 2^M states. In this section, we see that if each agent's belief is represented as a BN, the LinOP and LogOP consensus for any probabilistic query can be computed more efficiently. In particular, for the LogOP, we can compute the CPTs of a consensus BN with time complexity $O\left(NM^2 \cdot 2^{\max\{q(j)\}}\right)$, where $q(j)$ is the number of parents of A_j in the consensus structure.

We focus first on the task of generating a LogOP-consistent consensus BN. We compute its structure as described in Section 5.3. Consider computing the CPT at A_j, that is, $\Pr_0(A_j|\mathbf{pa}(A_j))$ for all combinations of outcomes of events in $\mathbf{pa}(A_j)$. From Proposition 5.4, we know that simply combining each agent's assessment of this condi-

[2] I do not claim that these consensus structures are minimal, or even that LogOP is the preferred aggregation method. My goal is more to guide a modeler's decision process by delineating what representations are consistent under what circumstances.

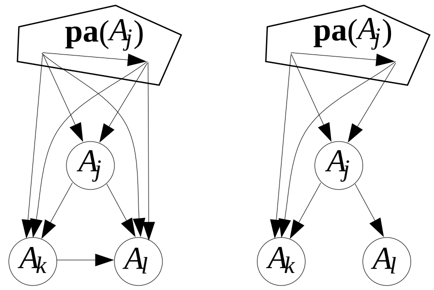


Figure 5.2: Two potential sections of a decomposable BN. A_j's children can be either in the same clique or in separate cliques.

tional probability will not succeed in general. However, we *can* compute the *last* CPT, $\Pr_0(A_M|\mathbf{pa}(A_M))$, in terms of only the $\Pr_i(A_M|\mathbf{pa}(A_M))$, by computing the LogOP over the single event A_M:

$\Pr_0(A_M|\mathbf{pa}(A_M)) =$

$$\frac{\prod_{i=1}^{N}\left[\Pr_i(A_M|\mathbf{pa}(A_M))\right]^{\alpha_i}}{\prod\left[\Pr_i(A_M|\mathbf{pa}(A_M))\right]^{\alpha_i}+\prod\left[\Pr_i(\bar{A_M}|\mathbf{pa}(A_M))\right]^{\alpha_i}}. \tag{5.8}$$

Because the LogOP satisfies EB, if we condition on *all* other events $Z - A_M$ in the network, then the LogOP over just A_M will return the same result as if we had computed the LogOP over all events, and then conditioned on $Z - A_M$. Equation 5.8 also reflects the fact that $\Pr_0(A_M|\mathbf{pa}(A_M)) = \Pr_0(A_M|Z - A_M)$ and $\Pr_i(A_M|\mathbf{pa}(A_M)) = \Pr_i(A_M|Z - A_M)$, by the semantics of the BNs.

We can compute the remainder of the CPTs in reverse index order. Assume that the CPTs $\Pr_0(A_k|\mathbf{pa}(A_k))$ have been calculated for all $k > j$, and that next we need to calculate $\Pr_0(A_j|\mathbf{pa}(A_j))$. To simplify the discussion, let A_j have exactly two children, A_k and A_l, with $j < k < l$; the analysis generalizes easily to more children (or one child). Since the BN is decomposable, its topology is a tree of cliques (Chyu, 1991b; Pearl, 1988; Shachter et al., 1991), and A_k and A_l can either be in the same clique or in separate cliques, as depicted in Figure 5.2. Note that decomposability also ensures that A_j's neighbors, $A_l \cup A_k \cup \mathbf{pa}(A_j)$, constitute its Markov blanket. We can query each of the agent's BNs for the probabilities $\Pr_i(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))$ using a standard BN inference algorithm. From these, we can compute the corresponding consensus probability as a LogOP only over A_j, as before:

$$\Pr_0(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))$$

$$\propto \prod_{i=1}^{N} [\mathrm{Pr}_i(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))]^{\alpha_i}. \tag{5.9}$$

We now need only eliminate the conditioning on A_l and A_k. By Bayes's rule, we have that

$$\begin{aligned}
&\frac{\mathrm{Pr}_0(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))}{\mathrm{Pr}_0(\bar{A}_j|A_l \cup A_k \cup \mathbf{pa}(A_j))} \\
&\quad = \frac{\mathrm{Pr}_0(A_l \cup A_k|A_j \cup \mathbf{pa}(A_j))}{\mathrm{Pr}_0(A_l \cup A_k|\bar{A}_j \cup \mathbf{pa}(A_j))} \cdot \frac{\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j))}{\mathrm{Pr}_0(\bar{A}_j|\mathbf{pa}(A_j))} \\
&\quad = \frac{\mathrm{Pr}_0(A_l|A_k \cup A_j \cup \mathbf{pa}(A_j))}{\mathrm{Pr}_0(A_l|A_k \cup \bar{A}_j \cup \mathbf{pa}(A_j))} \cdot \frac{\mathrm{Pr}_0(A_k|A_j \cup \mathbf{pa}(A_j))}{\mathrm{Pr}_0(A_k|\bar{A}_j \cup \mathbf{pa}(A_j))} \\
&\qquad \cdot \frac{\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j))}{\mathrm{Pr}_0(\bar{A}_j|\mathbf{pa}(A_j))}.
\end{aligned}$$

Because the BN is decomposable, and regardless of whether A_k and A_l are in the same or different cliques, $\mathrm{Pr}_0(A_l|A_k \cup A_j \cup \mathbf{pa}(A_j)) = \mathrm{Pr}_0(A_l|\mathbf{pa}(A_l))$ and $\mathrm{Pr}_0(A_k|A_j \cup \mathbf{pa}(A_j)) = \mathrm{Pr}_0(A_k|\mathbf{pa}(A_k))$, both of which have already been computed. Therefore we can calculate the CPT at A_j as follows:

$$\begin{aligned}
&\frac{\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j))}{\mathrm{Pr}_0(\bar{A}_j|\mathbf{pa}(A_j))} \\
&\quad = \frac{\mathrm{Pr}_0(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))}{\mathrm{Pr}_0(\bar{A}_j|A_l \cup A_k \cup \mathbf{pa}(A_j))} \cdot \frac{\mathrm{Pr}_0(A_l|\tilde{\mathbf{pa}}(A_l))}{\mathrm{Pr}_0(A_l|\mathbf{pa}(A_l))} \\
&\qquad \cdot \frac{\mathrm{Pr}_0(A_k|\tilde{\mathbf{pa}}(A_k))}{\mathrm{Pr}_0(A_k|\mathbf{pa}(A_k))},
\end{aligned} \tag{5.10}$$

where $\tilde{\mathbf{pa}}(A_k)$ and $\tilde{\mathbf{pa}}(A_l)$ contain $\bar{A}_j$, and $\mathbf{pa}(A_k)$ and $\mathbf{pa}(A_l)$ contain A_j. Once we compute the likelihood ratio on the LHS of (5.10), the desired probabilities are uniquely determined, since $\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j)) + \mathrm{Pr}_0(\bar{A}_j|\mathbf{pa}(A_j)) = 1$. The psuedocode for the full algorithm is given in Figure 5.3.

A consensus BN consistent with the LinOP would in general be fully connected, and thus not an object of particular value. However, if all agents' beliefs are given as BNs, we can retain their separation and still compute LinOP queries more efficiently. We exploit the fact that the LinOP obeys MP, and thus that the LinOP of any compound, marginal event can be computed as a LinOP over only that event. For example,

$$\mathrm{Pr}_0(A_2\bar{A}_5A_9) = \sum_{i=1}^{N} \alpha_i \mathrm{Pr}_i(A_2\bar{A}_5A_9),$$

LOGOP-CONSENSUS-BN$(\mathrm{Pr}_1, \mathrm{Pr}_2, \ldots, \mathrm{Pr}_N)$

INPUT: N Bayesian networks: $\mathrm{Pr}_1, \mathrm{Pr}_2, \ldots, \mathrm{Pr}_N$

OUTPUT: LogOP-consistent consensus BN: Pr_0

1. Structure of $\mathrm{Pr}_0 =$ TRIANGULATE $[\cup_{i=1}^{N}$ MORALIZE $[\mathrm{Pr}_i]]$
2. $\mathrm{Pr}_0(A_M|\mathbf{pa}(A_M)) \propto \prod_{i=1}^{N} [\mathrm{Pr}_i(A_M|\mathbf{pa}(A_M))]^{\alpha_i}$
3. **for** $j = M-1$ **downto** 1
4. $\quad \mathrm{Pr}_0(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j)) \propto \prod_{i=1}^{N} [\mathrm{Pr}_i(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))]^{\alpha_i}$
5. $\quad \frac{\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j))}{\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j))} = \frac{\mathrm{Pr}_0(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))}{\mathrm{Pr}_0(A_j|A_l \cup A_k \cup \mathbf{pa}(A_j))} \cdot \frac{\mathrm{Pr}_0(A_l|\tilde{\mathbf{pa}}(A_l))}{\mathrm{Pr}_0(A_l|\mathbf{pa}(A_l))} \cdot \frac{\mathrm{Pr}_0(A_k|\tilde{\mathbf{pa}}(A_k))}{\mathrm{Pr}_0(A_k|\mathbf{pa}(A_k))}$

Figure 5.3: Algorithm for computing the CPTs of a LogOP-consistent consensus BN.

where the terms on the RHS are calculated using a standard algorithm for Bayesian inference. Any conditional probability can be computed as the division of two compound, marginal probabilities.

Finally, I characterize the computational complexity of LinOP when all input models are BNs. Clearly, computing an arbitrary query $\mathrm{Pr}_0(E|F)$ is NP-hard. Proposition 5.6 establishes that, even when all topologies agree, and even when only computing the LinOP of a CPT entry, the problem remains intractable.

Proposition 5.6 *Let all input BNs have identical topologies. Then computing* $\mathrm{Pr}_0(A_j|\mathbf{pa}(A_j))$ *consistent with LinOP is NP-hard.*

Proof. Suppose that $N = 2$. Let Pr_1 be an arbitrary BN and let Pr_2 have an identical topology, but encode a uniform distribution—that is, $\mathrm{Pr}_2(\omega) = 1/2^M$. I have shown that, if $\mathrm{Pr}_0(A_M|\mathbf{pa}(A_M))$ were computable in polynomial time, then $\mathrm{Pr}_1(A_M)$ could be inferred in polynomial time. Computing the later query is NP-hard (Cooper, 1990), and so the former must be as well. □

5.5 Related Work

Faria and Smith (1996) examine a group decision making situation where agents agree on a common decomposable BN structure and have identical preferences. They define a weaker form of EB, called *conditional external Bayesianity* (CEB), which requires EB to hold only for CPT entries, and only when evidence updates are based on *cutting* likelihood functions—those which can be factored according to the model structure. They show that a generalized LogOP, called a conditional LogOP, is the only pooling function that satisfies both CEB and UNAM. The conditional modified LogOP preserves the agreed-upon structure and allows expert weights to vary across families in the structure. The authors also present an associated procedure for iteratively revising weights that reflects the relative alignment of the experts' predictions with actual observed outcomes.

Ng and Abramson (1994) describe an architecture called the *probabilistic multi-knowledge-base system*, which consists of a collection of BNs, each encoding the knowledge of a single expert. The BNs are kept separate and probabilities are combined *at run time* with a variable-weight variant of the LinOP. The authors address a variety of engineering issues, including the elicitation and propagation of expert confidence information, and build a working prototype to diagnose pathologies of the lymph system. Xiang (1996) describes conditions under which *multiply sectioned Bayesian networks*, originally developed for single agent reasoning, can represent the combined beliefs of multiple agents. The main assumption is that, whenever two agents' BNs contain some of the same events, they must agree on the joint distribution over these common events. Bonduelle (1987) prescribes both normative and behavioral techniques for a decision maker (DM) to identify and reconcile differences of opinion among experts. When those opinions are expressed as graphical models, he suggests that the DM first choose a consensus topology, and then calculate aggregate probabilities. Jacobs (1995) compares the LinOP and supra Bayesian approaches as methods for combining the multiple feature analyzers found in real and artificial neural systems.

Matzkevich and Abramson (1992) give an algorithm for explicitly combining two BN DAGs into a single DAG, or *fusing* the two topologies. The algorithm transfers one arc at a time from the second DAG to the first, possibly reversing the arc in order to remain consistent with the current partial ordering. Reversing arcs may add new arcs to the second DAG (Shachter, 1988), which would in turn need to be transferred. In a second paper, the same authors show (1993) that the task of minimizing the number of arcs in their combined DAG is NP-hard, as are several other related tasks. They argue that, intuitively, the consensus model should capture independencies agreed upon by at least $c \leq n$ of the agents; in particular, when $c = n$ and the orderings are mutually consistent, the consensus DAG should be a union of the individual DAGs. In both of these papers, and in Bonduelle's work, it is essentially assumed that the EIPP, or a stronger version thereof, should hold.

Though Matzkevich and Abramson make no commitment on how to combine probabilities, they do give an example (1992) where the LinOP is applied *locally*, or separately within each CPT, thus satisfying the FA property. Although such a constraint on aggregation may seem natural, we saw in Section 5.2 that it actually has very severe implications.

Chapter 6

Compact Market Representations of Probability: Top-Down Approach

We have seen that the equilibrium prices of securities are interpretable as probabilities. But, as described in Section 2.3.3, in order to uniquely determine a complete probability distribution over Ω (likelihoods for all states $\omega \in \Omega$) we generally require a complete market of $|\Omega| - 1 = 2^M - 1$ securities, where M is the number of primary events.

Bayesian networks (BNs) can implicitly define a complete probability distribution over Ω, often with far fewer than $2^M - 1$ parameters, by exploiting conditional independencies among the primary events. A natural question is whether securities markets can benefit from a similar reduction in size. Is it possible for a market of fewer than $2^M - 1$ securities to implicitly (and uniquely) describe a full joint distribution?

This chapter answers the question in the affirmative, by outlining a procedure for constructing such markets. The construction is *top-down*, meaning that it specifies a collection of agents and securities such that the implied price equilibrium conforms to any given probability distribution. This chapter is essentially an exercise in existence. This methodology diverges from the bottom-up philosophy embraced in the rest of the dissertation.

Also in contrast to other chapters, the market framework adopted here is that of general equilibrium economics under *certainty* (Section 2.3.2), rather than under uncertainty. In fact, the flow of all other chapters is not affected by skipping the current chapter. Agents do not have explicit beliefs or risk profiles, and utility for securities is not defined as expected utility for money. Rather, agents have utility directly for securities, as if securities were standard consumable commodities. Henceforth, I will refer to securities instead as *goods*. Although the good $\langle A \rangle$ still pays off \$1 contingent on the occurrence of A, agent preferences are defined over the goods themselves, rather than dollars.

The compact market construction is presented in terms of a mapping from any BN to what I call a *MarketBayes* economy: a configuration of goods, consumers, and producers such that equilibrium prices in the market equal the probabilities in the BN. I also refer to the mapping itself as the "MarketBayes mapping". Section 6.1 outlines the basic economic constructs that form the building blocks of a MarketBayes economy. Section 6.2 describes the complete mapping, and characterizes the equivalence between the resulting market price system and the given probability distribution. Events in the BN are encoded as goods, conditional probabilities as consumers, and the laws of probability as arbitrage producers.[1] The mapping introduces one consumer per conditional probability, although multiple disagreeing consumers are possible, in which case equilibrium prices would constitute an aggregate probability distribution. Section 6.3 walks through the construction of an example MarketBayes economy.

6.1 MarketBayes Building Blocks

Sections 6.1.1 and 6.1.2 define the MarketBayes consumers and producers, respectively.

6.1.1 The Consumers

In a BN, the basic unit of information is a conditional probability. For example, an event A_2 with sole predecessor A_1 is accompanied by the information $\Pr(A_2|A_1) = \pi$, where π is some probability. Using the definition of conditional probability, the same equation can be rewritten as $\Pr(A_1A_2) = \pi \Pr(A_1)$. In the MarketBayes economy, we wish to enforce the same ratio between the prices of the goods:

$$p^{\langle A_1A_2\rangle} = \pi p^{\langle A_1\rangle}. \tag{6.1}$$

Our approach to maintaining this relation is to introduce a consumer that considers the relative value of the good $\langle A_1A_2\rangle$ to be π times the value of the good $\langle A_1\rangle$. If the ratio of the prices diverges from π, the consumer will buy or sell accordingly, tending to drive the ratio toward π.

[1] A MarketBayes economy can be seen as a vehicle for distributed Bayesian inference. However, employing it as such is unlikely to yield any speedup, even with separate processors for every agent and good. The purpose of the MarketBayes construction is rather to generate compact market structures and facilitate belief aggregation. For multiprocessor speedups, look elsewhere (Delcher et al., 1996; Pennock, 1998).

The MarketBayes economy employs CES (constant elasticity of substitution) consumers for this purpose. The CES utility function for two goods takes the form[2]

$$u(x^{\langle 1\rangle}, x^{\langle 2\rangle}) = \left(\alpha^{\langle 1\rangle}\left[x^{\langle 1\rangle}\right]^{\frac{\sigma-1}{\sigma}} + \alpha^{\langle 2\rangle}\left[x^{\langle 2\rangle}\right]^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}, \tag{6.2}$$

where the $\alpha^{\langle j\rangle}$ are coefficients dictating the relative values of the two goods, and σ is a global substitution parameter dictating the degree to which consumption in one good (at proportions dictated by the $\alpha^{\langle j\rangle}$) can substitute for the other.

Let $p^{\langle 1\rangle}$ and $p^{\langle 2\rangle}$ be the prices of the two goods. The consumer's optimization problem (2.7) is to maximize its utility function (6.2) subject to its budget constraint, as described in Section 2.3.2. For CES consumers this problem has a closed-form solution:

$$x^{\langle 1\rangle}(p^{\langle 1\rangle}, p^{\langle 2\rangle}) = \frac{\left[\alpha^{\langle 1\rangle}\right]^{\sigma}\left(p^{\langle 1\rangle}e^{\langle 1\rangle} + p^{\langle 2\rangle}e^{\langle 2\rangle}\right)}{\left[p^{\langle 1\rangle}\right]^{\sigma}\left(\left[\alpha^{\langle 1\rangle}\right]^{\sigma}\left[p^{\langle 1\rangle}\right]^{1-\sigma} + \left[\alpha^{\langle 2\rangle}\right]^{\sigma}\left[p^{\langle 2\rangle}\right]^{1-\sigma}\right)} \tag{6.3}$$

$$x^{\langle 2\rangle}(p^{\langle 1\rangle}, p^{\langle 2\rangle}) = \frac{\left[\alpha^{\langle 2\rangle}\right]^{\sigma}\left(p^{\langle 1\rangle}e^{\langle 1\rangle} + p^{\langle 2\rangle}e^{\langle 2\rangle}\right)}{\left[p^{\langle 2\rangle}\right]^{\sigma}\left(\left[\alpha^{\langle 1\rangle}\right]^{\sigma}\left[p^{\langle 1\rangle}\right]^{1-\sigma} + \left[\alpha^{\langle 2\rangle}\right]^{\sigma}\left[p^{\langle 2\rangle}\right]^{1-\sigma}\right)} \tag{6.4}$$

To implement an equation of the form (6.1), we introduce a CES consumer interested in the two goods $\langle A_1A_2\rangle$ and $\langle A_1\rangle$. The consumer is endowed with an equal amount $e^{\langle A_1A_2\rangle} = e^{\langle A_1\rangle} = e$ of each good (the exact value does not matter for the current purpose, as long as $e > 0$). By setting $\alpha^{\langle A_1A_2\rangle} = \pi$ and $\alpha^{\langle A_1\rangle} = 1$, we encode the desired conditional probability. Although the relation is strictly enforced (according to Proposition 6.1 below) only as $\sigma \to \infty$, we have found in practice that convergence to the correct price ratio typically obtains for values of $\sigma > 4$ or so.

Proposition 6.1 *Let $p^{\langle A_1A_2\rangle^*}$ and $p^{\langle A_1\rangle^*}$ be equilibrium prices for the two goods in an economy containing the CES consumer defined above, with $\sigma \to \infty$. If $\lim_{\sigma\to\infty} \frac{p^{\langle A_1A_2\rangle^*}}{p^{\langle A_1\rangle^*}}$ is finite and bounded away from zero, then it is π.*

Proof. In competitive equilibrium, by definition, the consumer is solving its optimization problem. The first-order conditions for that optimization problem dictate that the marginal utility per unit price be constant across goods. In particular,

$$\frac{\frac{\partial u}{\partial x^{\langle A_1A_2\rangle}}}{p^{\langle A_1A_2\rangle^*}} = \frac{\frac{\partial u}{\partial x^{\langle A_1\rangle}}}{p^{\langle A_1\rangle^*}}.$$

[2]CES forms are commonly employed in general equilibrium modeling (Shoven and Whalley, 1992), due to their flexibility and convenient analytical properties.

Substituting the marginal utilities and rearranging,

$$\begin{aligned}\frac{p^{\langle A_1A_2\rangle *}}{p^{\langle A_1\rangle *}} &= \frac{\left(\pi\left[x^{\langle A_1A_2\rangle}\right]^{\frac{\sigma-1}{\sigma}}+\left[x^{\langle A_1\rangle}\right]^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{\sigma-1}}\pi\left[x^{\langle A_1A_2\rangle}\right]^{-\frac{1}{\sigma}}}{\left(\pi\left[x^{\langle A_1A_2\rangle}\right]^{\frac{\sigma-1}{\sigma}}+\left[x^{\langle A_1\rangle}\right]^{\frac{\sigma-1}{\sigma}}\right)^{\frac{1}{\sigma-1}}\left[x^{\langle A_1\rangle}\right]^{-\frac{1}{\sigma}}}\\ &= \frac{\pi\left[x^{\langle A_1A_2\rangle}\right]^{-\frac{1}{\sigma}}}{\left[x^{\langle A_1\rangle}\right]^{-\frac{1}{\sigma}}},\end{aligned}$$

which approaches π as $\sigma \to \infty$, as long as the $x^{\langle j\rangle}$ are bounded away from zero. For CES consumers, this will be true as long as the price ratio is finite and bounded away from zero. Thus, $p^{\langle A_1A_2\rangle *} = \pi p^{\langle A_1\rangle *}$, which is Equation (6.1) exactly. □

The alert reader may observe that the same result could have been obtained more directly using the linear utility function, $u(x^{\langle A_1A_2\rangle}, x^{\langle A_1\rangle}) = \pi x^{\langle A_1A_2\rangle} + x^{\langle A_1\rangle}$. Indeed, the CES utility function approaches linearity in the limit. However, with linear utility the equilibrium would be more fragile—two consumers with different π would prevent existence. Moreover, for linear utility functions the optimal demand function is discontinuous in prices, and reaching equilibrium through a distributed, incremental bidding process becomes more difficult.

Note that Proposition 6.1 requires only that there exist one such CES consumer. It is true vacuously when there is no equilibrium. For situations with more than one CES consumer connecting the same pair of goods with differing π values, the result still holds because we take the σ for only one of them to infinity.

6.1.2 The producers

Whereas the role of consumers in a MarketBayes economy is to encode conditional probabilities, we employ producers to implement identities of probability theory. Producers act as *arbitrageurs*, converting between logically equivalent bundles of goods.[3] For example, the identity

$$A_1 \equiv A_1A_2 \vee A_1\bar{A}_2$$

can be expressed in a producer with the technology to convert a unit of $\langle A_1\rangle$ into one unit each of $\langle A_1A_2\rangle$ and $\langle A_1\bar{A}_2\rangle$, or *vice versa*. This producer's technology exhibits *constant*

[3] Our expression of the laws of probability in arbitrageurs can be viewed as a direct embodiment of Nau and McCardle's argument (1991) that all rationality or coherence principles ultimately reduce to "no-arbitrage" postulates.

returns to scale, that is, it can perform this transform at any volume level.

The corresponding probabilistic identity, $\Pr(A_1) = \Pr(A_1A_2) + \Pr(A_1\bar{A}_2)$, can likewise be expressed in terms of a constraint on prices of goods:

$$p^{\langle A_1 \rangle} = p^{\langle A_1 A_2 \rangle} + p^{\langle A_1 \bar{A}_2 \rangle}. \tag{6.5}$$

The arbitrageur effectively enforces this equation by its bidding policy. If the price $p^{\langle A_1 \rangle}$ diverges from the sum $p^{\langle A_1 A_2 \rangle} + p^{\langle A_1 \bar{A}_2 \rangle}$, the producer can make profits by transforming one side to the other. Its resulting demand behavior will tend to drive the respective input and output prices towards equality.

The producer's goal is to maximize profits (2.8). For the producer associated with the identity above, the profits when transforming y units of $\langle A_1 \rangle$ into y each of $\langle A_1 A_2 \rangle$ and $\langle A_1 \bar{A}_2 \rangle$ (note that if y is negative, the transformation goes the other way) are simply

$$y\left(p^{\langle A_1 A_2 \rangle} + p^{\langle A_1 \bar{A}_2 \rangle} - p^{\langle A_1 \rangle}\right). \tag{6.6}$$

Proposition 6.2 *Let $p^{\langle A_1 \rangle^*}$, $p^{\langle A_1 A_2 \rangle^*}$, and $p^{\langle A_1 \bar{A}_2 \rangle^*}$ be the equilibrium prices for the three goods in an economy containing the arbitrage producer defined above. Then $p^{\langle A_1 \rangle^*} = p^{\langle A_1 A_2 \rangle^*} + p^{\langle A_1 \bar{A}_2 \rangle^*}$.*

Proof. Competitive producers must be maximizing profits in equilibrium. But the profit function (6.6) has a bounded maximizer y (finite production) only if Equation 6.5 holds. $\square$

Note that the producer always makes zero profit in equilibrium (this is true in general for competitive, constant-returns producers). Results analogous to Proposition 6.2 can be derived for arbitrageurs representing identities of the form (6.5) but with arbitrary numbers of event conjunctions on the right-hand side.

6.2 Mapping from Probabilities to Markets

In this section we piece together the individual components described above into an interconnected market price system. We describe a general mapping from any BN to a MarketBayes configuration of goods, consumers, and producers. We show how the economy effectively represents the same information as the BN.

We are interested in three general properties that such a mapping may possess.

Property 6.1 (Existence) *There exists a competitive equilibrium in the MarketBayes economy such that the price of each good equals the probability of the corresponding event in the BN.*

Property 6.2 (Uniqueness) *There is a* unique *competitive equilibrium in the MarketBayes economy, satisfying the conditions of Property 6.1.*

Property 6.3 (Convergence) *The unique competitive equilibrium of the MarketBayes economy can be derived via an iterative, distributed bidding process, such as tatonnement or variants (Mas-Colell et al., 1995; Cheng and Wellman, 1998).*

Each property subsumes the previous and, in general, the properties are successively harder to verify. In the following sections, we construct the MarketBayes economy incrementally in two stages. After stage one, we prove that Property 6.1 holds; after stage two we prove that Property 6.2 holds. The goods in the economy—specified in Section 6.2.1—remain constant across both stages. In Section 6.2.2, we define stage one of the mapping, a set of consumers sufficient to establish the Existence Property for arbitrary BNs. In Section 6.2.3 we define stage two of the mapping, adding producers to the economy to ensure the Uniqueness Property for a non-restrictive class of BNs. We do not yet have a general proof of the Convergence Property (except for complete graphs, not presented), but we conjecture that the WALRAS price adjustment algorithm (Cheng and Wellman, 1998) does converge for a broad class of MarketBayes economies. Our computational experience supports this conjecture, and in Section 6.3 we present a concrete example that does indeed converge when implemented as a computational economy.

6.2.1 The Goods

The goods in a MarketBayes economy are securities contingent on conjunctions of events. Rather than include all such conjunctions explicitly as goods, we attempt to exploit the independencies present in a (possibly sparse) BN graph.

Let $A_1, \ldots, A_M$ be events in a BN. Denote the parents of node A_j by $A_{j_1}, \ldots, A_{j_q}$. For each node A_j in the network we add to the economy goods for all 2^{q+1} possible conjunctions of A_j and its parents:[4]

$$\langle A_j A_{j_1} \cdots A_{j_q} \rangle, \langle A_j A_{j_1} \cdots \bar{A}_{j_q} \rangle, \ldots, \langle \bar{A}_j \bar{A}_{j_1} \cdots \bar{A}_{j_q} \rangle.$$

[4]Though each node may have a different number of parents, we conserve subscripts and use q for the number of parents of the *current* node.

We also add to the economy all 2^q possible conjunctions of the parents of A_j alone, if these goods are not already included in those defined previously.

$$\langle A_{j_1} \cdots A_{j_q} \rangle, \langle A_{j_1} \cdots \bar{A}_{j_q} \rangle, \ldots, \langle \bar{A}_{j_1} \cdots \bar{A}_{j_q} \rangle$$

Finally we add a single good $\langle T \rangle$ which corresponds to the proposition **true**. This security always pays off \$1, regardless of the outcomes of any events. $\langle T \rangle$ naturally plays the role of the numeraire of the economy and its price is kept at \$1 by definition.

The economy then consists of $O(M \cdot 2^q)$ goods, where q is the maximum number of parents of any node, and M is the number of nodes.

6.2.2 Stage one: The Consumers

For each node A_j, the given BN provides us with 2^q conditional probabilities, of the form $\Pr(A_j | A_{j_1} A_{j_2} \cdots A_{j_q}) = \pi$. These conditional probabilities dictate that the following 2^q ratios between prices of goods must hold:

$$\begin{array}{rcl} p^{\langle A_j A_{j_1} A_{j_2} \cdots A_{j_q} \rangle} & = & \pi_1 p^{\langle A_{j_1} A_{j_2} \cdots A_{j_q} \rangle} \\ p^{\langle A_j A_{j_1} A_{j_2} \cdots \bar{A}_{j_q} \rangle} & = & \pi_2 p^{\langle A_{j_1} A_{j_2} \cdots \bar{A}_{j_q} \rangle} \\ \vdots & & \vdots \\ p^{\langle A_j \bar{A}_{j_1} \bar{A}_{j_2} \cdots \bar{A}_{j_q} \rangle} & = & \pi_{2^q} p^{\langle \bar{A}_{j_1} \bar{A}_{j_2} \cdots \bar{A}_{j_q} \rangle} \end{array} \tag{6.7}$$

From the above equations, it is trivial to derive the complementary equations that contain $\bar{A}_j$:

$$\begin{array}{rcl} p^{\langle \bar{A}_j A_{j_1} A_{j_2} \cdots A_{j_q} \rangle} & = & (1 - \pi_1) p^{\langle A_{j_1} A_{j_2} \cdots A_{j_q} \rangle} \\ p^{\langle \bar{A}_j A_{j_1} A_{j_2} \cdots \bar{A}_{j_q} \rangle} & = & (1 - \pi_2) p^{\langle A_{j_1} A_{j_2} \cdots \bar{A}_{j_q} \rangle} \\ \vdots & & \vdots \\ p^{\langle \bar{A}_j \bar{A}_{j_1} \bar{A}_{j_2} \cdots \bar{A}_{j_q} \rangle} & = & (1 - \pi_{2^q}) p^{\langle \bar{A}_{j_1} \bar{A}_{j_2} \cdots \bar{A}_{j_q} \rangle} \end{array} \tag{6.8}$$

Root nodes in the BN can be handled in the same way by considering them to be children of the proposition **true**, with "conditional" probability $\Pr(A_j | \mathbf{true}) = \pi$.

In the MarketBayes economy, CES consumers effectively implement equations of the form (6.7) and (6.8). Specifically, we add $2 \cdot 2^q$ consumers to the economy for each node A_j—one for each of the equations in (6.7) and (6.8). The consumers are instantiated as described in Section 6.1.1.

Proposition 6.3 *Property 6.1 (Existence) holds for the mapping from any BN to the set* *of goods and consumers defined above.*

Proof. We need to show that the probabilities of events in the BN form a (possibly nonunique) price equilibrium in the MarketBayes economy. A set of prices constitutes an equilibrium if the resulting demand is in material balance, that is, there is no *excess demand* in the economy at those prices. Probabilities represented by the BN obey the ratios described in Equations (6.7) and (6.8), by construction. Then it suffices to show that any set of prices that satisfies these equations implies zero excess demand in the economy. Consider the CES consumer associated with the first conditional probability in (6.7). Let $p^{\langle 1\rangle} = p^{\langle A_j A_{j_1} A_{j_2} \cdots A_{j_q}\rangle}$ and $p^{\langle 2\rangle} = p^{\langle A_{j_1} A_{j_2} \cdots A_{j_q}\rangle}$ be the prices of the two relevant goods. From (6.7) we have that $p^{\langle 1\rangle} = \pi_1 p^{\langle 2\rangle}$. The demands for each good (from (6.3) and (6.4)) are then:

$$
\begin{aligned}
x^{\langle 1\rangle}(\pi_1 p^{\langle 2\rangle}, p^{\langle 2\rangle}) &= \frac{{\pi_1}^{\sigma}(\pi_1 p^{\langle 2\rangle} e + p^{\langle 2\rangle} e)}{\left[\pi_1 p^{\langle 2\rangle}\right]^{\sigma}\left({\pi_1}^{\sigma}\left[\pi_1 p^{\langle 2\rangle}\right]^{1-\sigma} + \left[p^{\langle 2\rangle}\right]^{1-\sigma}\right)} \\
&= \frac{{\pi_1}^{\sigma} p^{\langle 2\rangle} e(\pi_1 + 1)}{{\pi_1}^{\sigma}\pi_1 p^{\langle 2\rangle} + {\pi_1}^{\sigma} p^{\langle 2\rangle}} = \frac{e(\pi_1+1)}{\pi_1+1} = e
\end{aligned}
$$

$$
\begin{aligned}
x^{\langle 2\rangle}(\pi_1 p^{\langle 2\rangle}, p^{\langle 2\rangle}) &= \frac{\pi_1 p^{\langle 2\rangle} e + p^{\langle 2\rangle} e}{\left[p^{\langle 2\rangle}\right]^{\sigma}\left({\pi_1}^{\sigma}\left[\pi_1 p^{\langle 2\rangle}\right]^{1-\sigma} + \left[p^{\langle 2\rangle}\right]^{1-\sigma}\right)} \\
&= \frac{p^{\langle 2\rangle} e(\pi_1 + 1)}{p^{\langle 2\rangle}\pi_1 + p^{\langle 2\rangle}} = \frac{p^{\langle 2\rangle} e(\pi_1 + 1)}{p^{\langle 2\rangle}(\pi_1 + 1)} = e
\end{aligned}
$$

In other words, at the specified price ratio, the consumer demands exactly the amounts it is endowed with. Since the same argument applies to every consumer in the economy, the total excess demand must be zero. Thus, any set of prices that satisfies (6.7) and (6.8) constitutes a competitive equilibrium for the economy, and the Existence Property is established. □

Note that this result does not require $\sigma \to \infty$, as did Proposition 6.1. If consumers other than the ones specified in our construction are present, then the network probabilities may not constitute a price equilibrium.

6.2.3 Stage Two: The producers

Although the existence of an equilibrium corresponding to the probability distribution is somewhat encouraging, the MarketBayes construction will be of limited use if other,

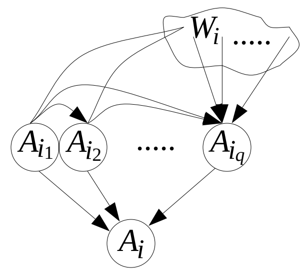


Figure 6.1: A section of a decomposable BN. Undirected links could be oriented either way.

incorrect, probabilities can also represent an equilibrium. In this section we extend the economy to ensure the Uniqueness Property. To do so, we must in general add some consumers and producers to enforce the remaining constraints. These additions correspond to the arcs necessary to render the network *decomposable*, as defined in Section 2.3.4. Any graph can be made decomposable so, in principle, the MarketBayes mapping applies to arbitrary networks. Note that the transformation can be performed with respect to any ordering θ of the events. Finding the *optimal* ordering is equivalent to finding a minimum triangulation, and is itself NP-hard, though reasonable heuristics are available. Moreover, in some cases, even the optimal ordering necessitates an exponential blowup in the size of the representation (Kloks, 1994).

As above, let $A_1, \ldots, A_M$ be the events of the BN. Without loss of generality, we can assume the index labels on the nodes are consistent with the partial order represented by the directed graph: if A_j is an ancestor of A_k then $j < k$. Note that A_1 is always a root node (no parents) and A_M is always a leaf node (no children). Since the network is decomposable, the parents of A_j—$A_{j_1}, \ldots, A_{j_q}$—form a complete subset of the graph. Furthermore, (taking, again without loss of generality, $j_1 < \cdots < j_q$), the nodes $A_{j_1}, A_{j_2}, \ldots, A_{j_{q-1}}$ must all be parents of the node A_{j_q}. Let W_j be the set of parents of the node A_{j_q} that are *not* also parents of A_j. This structure is depicted in Figure 6.1.

The set W_j represents the additional nodes required to specify a joint distribution over the parents of A_j. For each node A_j, we generate arbitrageurs (according to the scheme of Section 6.1.2) representing the following probabilistic identities:

$$\begin{array}{rcl} p^{\langle A_{j_1} A_{j_2} \cdots A_{j_q} \rangle} & = & \sum_{W_j} p^{\langle W_j A_{j_1} A_{j_2} \cdots A_{j_q} \rangle} \\ p^{\langle A_{j_1} A_{j_2} \cdots \bar{A}_{j_q} \rangle} & = & \sum_{W_j} p^{\langle W_j A_{j_1} A_{j_2} \cdots \bar{A}_{j_q} \rangle} \\ \vdots & & \vdots \\ p^{\langle \bar{A}_{j_1} \bar{A}_{j_2} \cdots \bar{A}_{j_q} \rangle} & = & \sum_{W_j} p^{\langle W_j \bar{A}_{j_1} \bar{A}_{j_2} \cdots \bar{A}_{j_q} \rangle} \end{array} \tag{6.9}$$

The summations are over all possible combinations of outcomes of events in the set W_j. Note that if the set W_j is empty for some j, then we need not add any producers to the economy for node A_j. If the BN is a complete graph, then the set W_j will be empty for all j; in this case producers are simply not necessary.

Proposition 6.4 *Property 6.2 (Uniqueness) holds for the mapping from any decomposable BN to the set of goods, consumers, and producers defined above.*

Proof. From Propositions 6.1 and 6.2 it is clear that the consumer equations in (6.7) and (6.8), along with the producer equations in (6.9), must hold simultaneously in equilibrium. The probabilities of the events in the BN must satisfy these equations since they correspond directly to the given conditional probabilities plus identities in probability theory. Therefore, we need only show that there is a *unique* set of prices that satisfies this set of equations instantiated for the configuration of this economy. The proof is by induction. Let G_j be the set of goods added to the economy for event A_j, namely, all possible conjunctions of A_j with its parents plus all possible conjunctions of its parents alone. Mathematically,

$$\begin{aligned} G_j &= \{\langle A_j A_{j_1} \cdots A_{j_q}\rangle, \ldots, \langle \bar{A}_j \bar{A}_{j_1} \cdots \bar{A}_{j_q}\rangle\} \cup \\ &\quad \{\langle A_{j_1} \cdots A_{j_q}\rangle, \ldots, \langle \bar{A}_{j_1} \cdots \bar{A}_{j_q}\rangle\}. \end{aligned}$$

- **Base Case.** Define the set G_0 to be $\{\langle T\rangle\}$. The good $\langle T\rangle$ is the numeraire and its price is maintained at unity by definition. Thus the price of the good in the set G_0 is uniquely determined.

- **Induction.** Assume that all of the prices of goods in the sets $G_0, G_1, \ldots, G_{j-1}$ are uniquely determined. We want to prove that all of the prices of goods in the set G_j are uniquely determined. Consider the general situation as depicted in Figure 6.1. The prices of the goods in the set G_{j_q} are uniquely determined since $j_q < j$. Among the goods in the set G_{j_q} are the goods in the sets

$$\bigcup_{W_j} \langle W_j A_{j_1} A_{j_2} \cdots A_{j_q}\rangle, \ldots, \bigcup_{W_j} \langle W_j \bar{A}_{j_1} \bar{A}_{j_2} \cdots \bar{A}_{j_q}\rangle,$$

  where $\cup_{W_j}$ is the union over all possible outcomes of the events in the set W_j. These are exactly the goods on the right hand side of (6.9). Thus the prices of the goods on the left hand side of (6.9) must be uniquely determined. These goods are in turn

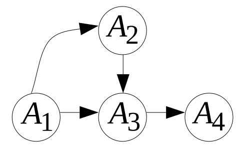


Figure 6.2: An example BN, to be transformed into a MarketBayes economy.

exactly those on the right hand sides of (6.7) and (6.8). Thus the prices of the goods on the left hand sides of (6.7) and (6.8) must be uniquely determined. The goods on the left hand sides of(6.7), (6.8), and (6.9) are exactly those goods in the set G_j. □

Note that the price system in an equilibrium MarketBayes economy is sufficient to recover the complete joint distribution, as it specifies all conditional probabilities in the original BN.

6.3 An Example Economy

In this section we construct a concrete MarketBayes economy using the technique described in the previous section. We provide empirical verification of the system by reporting results of running the example in an actual computational economy. For this example, the prices of goods converge correctly to the probabilities of the corresponding events in the BN.

The example BN is pictured in Figure 6.2. It is already decomposable, so we need not add any additional links. Let the conditional probabilities associated with the example network be as follows:

$$\begin{array}{ll}
\Pr(A_1) = 0.4 & \\
\Pr(A_2|A_1) = 0.2 & \Pr(A_2|\bar{A}_1) = 0.3 \\
\Pr(A_3|A_1A_2) = 0.11 & \Pr(A_3|\bar{A}_1A_2) = 0.22 \\
\Pr(A_3|A_1\bar{A}_2) = 0.33 & \Pr(A_3|\bar{A}_1\bar{A}_2) = 0.44 \\
\Pr(A_4|A_3) = 0.25 & \Pr(A_4|\bar{A}_3) = 0.85
\end{array}$$

The goods in the economy consist of all combinations of each node with its parents,

$$
\begin{array}{llll}
\langle A_1\rangle, & \langle \bar{A}_1\rangle, & & \\
\langle A_1A_2\rangle, & \langle A_1\bar{A}_2\rangle, & \langle \bar{A}_1A_2\rangle, & \langle \bar{A}_1\bar{A}_2\rangle, \\
\langle A_1A_2A_3\rangle, & \langle A_1A_2\bar{A}_3\rangle, & \langle A_1\bar{A}_2A_3\rangle, & \langle A_1\bar{A}_2\bar{A}_3\rangle, \\
\langle \bar{A}_1A_2A_3\rangle, & \langle \bar{A}_1A_2\bar{A}_3\rangle, & \langle \bar{A}_1\bar{A}_2A_3\rangle, & \langle \bar{A}_1\bar{A}_2\bar{A}_3\rangle, \\
\langle A_3A_4\rangle, & \langle A_3\bar{A}_4\rangle, & \langle \bar{A}_3A_4\rangle, & \langle \bar{A}_3\bar{A}_4\rangle
\end{array}
$$

and all combinations of each node's parents alone, if not already included in the group above. In this example, the combinations of the parent of node A_4 still need to be added.

$$\langle A_3\rangle, \langle \bar{A}_3\rangle$$

Finally we add the numeraire good $\langle T\rangle$ to the economy.

For each conditional probability in the BN we add a consumer. For this example we have consumers enforcing the following relationships:

$$
\begin{array}{ll}
p^{\langle A_1\rangle} = 0.4p^{\langle T\rangle} & \\
p^{\langle A_1A_2\rangle} = 0.2p^{\langle A_1\rangle} & p^{\langle \bar{A}_1A_2\rangle} = 0.3p^{\langle \bar{A}_1\rangle} \\
p^{\langle A_1A_2A_3\rangle} = 0.11p^{\langle A_1A_2\rangle} & p^{\langle \bar{A}_1A_2A_3\rangle} = 0.22p^{\langle \bar{A}_1A_2\rangle} \\
p^{\langle A_1\bar{A}_2A_3\rangle} = 0.33p^{\langle A_1\bar{A}_2\rangle} & p^{\langle \bar{A}_1\bar{A}_2A_3\rangle} = 0.44p^{\langle \bar{A}_1\bar{A}_2\rangle} \\
p^{\langle A_3A_4\rangle} = 0.25p^{\langle A_3\rangle} & p^{\langle \bar{A}_3A_4\rangle} = 0.85p^{\langle \bar{A}_3\rangle}
\end{array}
$$

We also add the complementary consumers defined by Equation (6.8).

$$
\begin{array}{ll}
p^{\langle \bar{A}_1\rangle} = 0.6p^{\langle T\rangle} & \\
p^{\langle A_1\bar{A}_2\rangle} = 0.8p^{\langle A_1\rangle} & p^{\langle \bar{A}_1\bar{A}_2\rangle} = 0.7p^{\langle \bar{A}_1\rangle} \\
p^{\langle A_1A_2\bar{A}_3\rangle} = 0.89p^{\langle A_1A_2\rangle} & p^{\langle \bar{A}_1A_2\bar{A}_3\rangle} = 0.78p^{\langle \bar{A}_1A_2\rangle} \\
p^{\langle A_1\bar{A}_2\bar{A}_3\rangle} = 0.67p^{\langle A_1\bar{A}_2\rangle} & p^{\langle \bar{A}_1\bar{A}_2\bar{A}_3\rangle} = 0.56p^{\langle \bar{A}_1\bar{A}_2\rangle} \\
p^{\langle A_3\bar{A}_4\rangle} = 0.75p^{\langle A_3\rangle} & p^{\langle \bar{A}_3\bar{A}_4\rangle} = 0.15p^{\langle \bar{A}_3\rangle}
\end{array}
$$

Consider the first equation above, $p^{\langle A_1\rangle} = 0.4p^{\langle T\rangle}$. The CES consumer representing this relationship has an interest in the two goods $\langle A_1\rangle$ and $\langle T\rangle$, with CES α coefficients of 0.4 and 1, respectively. In our computational market we endow the consumer with an amount $e = 10$ of each good, and set the global substitution parameter σ to 50. The remaining consumers are instantiated in the same way.

For each node in the BN we add the arbitrage producers defined in Equation 6.9. In this example, we need to add producers only for A_4, since the set W_j is empty for $j \neq 4$.

$$p^{\langle A_3 \rangle} = p^{\langle A_1 A_2 A_3 \rangle} + p^{\langle \bar{A}_1 A_2 A_3 \rangle} + p^{\langle A_1 \bar{A}_2 A_3 \rangle} + p^{\langle \bar{A}_1 \bar{A}_2 A_3 \rangle}$$

$$p^{\langle \bar{A}_3 \rangle} = p^{\langle A_1 A_2 \bar{A}_3 \rangle} + p^{\langle \bar{A}_1 A_2 \bar{A}_3 \rangle} + p^{\langle A_1 \bar{A}_2 \bar{A}_3 \rangle} + p^{\langle \bar{A}_1 \bar{A}_2 \bar{A}_3 \rangle}$$

This collection of goods, consumers, and producers forms a complete MarketBayes economy. We have implemented this example in our market-oriented programming environment, WALRAS, which provides some general facilities for specifying computational markets (Wellman, 1993). Given the specification of agents and goods, the WALRAS distributed bidding protocol attempts to find a competitive equilibrium via an asynchronous, iterative, price-adjustment process (Cheng and Wellman, 1998). In the WALRAS bidding protocol, each agent submits demand functions for each of the goods they are interested in to the auctions for the respective goods. The auction then sets the price so as to clear its market. When the prices change, the agents may submit new bids, and the process iterates. For this example economy, the MarketBayes prices indeed converge correctly to within 0.001 of the correct probabilities.

We can also use the market results to recover probabilities that are not explicitly represented as goods in the system. Since the conditional probabilities in a BN capture the complete joint distribution, the probability of any propositional expression can be recovered through additions and multiplications of MarketBayes prices. For example, $p^{\langle \bar{A}_1 A_3 \rangle}$ can be computed by summing $p^{\langle \bar{A}_1 A_2 A_3 \rangle} + p^{\langle \bar{A}_1 \bar{A}_2 A_3 \rangle}$. For simple summations like this, we can generate an explicit arbitrageur to produce the desired good. However, determining the probabilities of general expressions and conditionals may require new types of agents, or even off-line calculations using the prices of existing goods.

Chapter 7

Compact Securities Markets

As already discussed, complete securities markets define complete probability distributions over Ω. First, this chapter addresses the same question as Chapter 6: can a compact securities market, structured in the manner of a Bayesian network (BN), define a complete probability distribution over Ω? In addition, complete securities markets support Pareto optimal allocations of risk. Then this chapter attacks the second natural question: can compact markets also support all desirable exchanges of risk?[1] If a compact market both defines a complete probability distribution and supports Pareto optimal exchange, we will say that the market is *operationally complete*.

In this chapter, we return to the market framework of Chapter 3; in particular, we return to our standard view of agents as expected utility maximizers, with beliefs over Ω and utility for money. In contrast to Chapter 6, we do not explicitly enforce independencies in the market; rather we characterize what market structures are operationally complete given the agents' beliefs. Results in this chapter draw on those in Chapter 5. The impossibility and possibility of combining BNs have direct bearing on whether and when compact markets are operationally complete.

Section 7.1 reexamines the concepts of equilibrium and completeness in a securities market, stating them in terms of the agents' risk-neutral probabilities. Section 7.2 describes how securities markets can be structured exactly as BNs. Section 7.3 establishes that structured markets are operationally complete if, in equilibrium, all agents' risk-neutral independencies agree with those encoded in the market. Section 7.4 examines when agreement on true independencies is sufficient for operational completeness.

[1]This second question is not relevant in the context of Chapter 6, where agents do not explicitly have beliefs or risk.

7.1 Equilibrium in a Securities Market—Revisited

7.1.1 Equilibrium as Consensus

In Chapter 3, we invoked the standard formulation of competitive equilibrium (3.2) as a fixed point where each agent's demand is optimal at current prices, and each security's price balances aggregate demand. Here we examine an alternative characterization of equilibrium, recognized first by Drèze (1987). Agent i's first-order condition for $x_i^{\langle j\rangle}$ is:

$$\frac{\partial U_i(\mathbf{x})}{\partial x_i^{\langle j\rangle}} = \sum_{\omega\in\Omega} \Pr_i(\omega) \frac{\partial u_i\left(\Upsilon_i^{\langle\omega\rangle}\right)}{\partial x_i^{\langle j\rangle}} = 0,$$

where $\Upsilon_i^{\langle\omega\rangle} = \sum_k \left(1_{\omega\in A_k} - p^{\langle k\rangle}\right) x_i^{\langle k\rangle}$ is its payoff in state ω. Applying the chain rule

$$\begin{aligned}\sum_{\omega\in\Omega} \Pr_i(\omega)\left(1_{\omega\in A_j} - p^{\langle j\rangle}\right) u_i'\left(\Upsilon_i^{\langle\omega\rangle}\right) &= 0\\ \sum_{\omega\in A_j} \Pr_i(\omega) u_i'\left(\Upsilon_i^{\langle\omega\rangle}\right) - p^{\langle j\rangle} \sum_{\omega\in\Omega} \Pr_i(\omega) u_i'\left(\Upsilon_i^{\langle\omega\rangle}\right) &= 0,\end{aligned}$$

and solving for $p^{\langle j\rangle}$, we find that:

$$p^{\langle j\rangle} = \frac{\sum_{\omega\in A_j} \Pr_i(\omega) u_i'\left(\Upsilon_i^{\langle\omega\rangle}\right)}{\sum_{\omega\in\Omega} \Pr_i(\omega) u_i'\left(\Upsilon_i^{\langle\omega\rangle}\right)} = \Pr_i^{\mathrm{RN}}(A_j). \tag{7.1}$$

In words, equilibrium can also be considered a fixed point where exchanges among agents induce a *consensus* on risk-neutral probabilities across available securities, and where the security prices themselves match these agreed-upon values.

7.1.2 Complete Markets, Complete Consensus, and Pareto Optimality

As described in Section 2.3.3, a securities market is *complete* when $S = |\Omega| - 1$ and all securities are *linearly independent*. In such a market, equilibrium allocations of risk are Pareto optimal: any gamble, contingent on *any* event $E \subseteq \Omega$, that is an acceptable purchase for one agent is *not* an acceptable sale for any other (Arrow, 1964).

A probability distribution over Ω has dimensionality $|\Omega| - 1$ (normalized likelihoods for the $|\Omega|$ states). Prices of securities in a complete market constitute $|\Omega| - 1$ linearly independent equations for these $|\Omega| - 1$ unknowns, and thus define unique probabilities for all states $\omega \in \Omega$, also called the *state prices* (Huang and Litzenberger, 1988; Varian, 1987). Denote these probabilities as $\Pr_0(\omega)$, and let $\Pr_0(E) = \sum_{\omega \in E} \Pr_0(\omega)$ be the price-probability of any event E, perhaps not directly corresponding to an available security.

The agents' risk-neutral distributions also have dimensionality $|\Omega| - 1$, subject to the S constraints defined by (7.1). If the market is complete, it follows that $\Pr_i^{\mathrm{RN}}$ is uniquely determined, and equals $\Pr_0$ for all i. That is, a complete market induces a compete consensus on risk-neutral probabilities. This suggests an intuitive explanation of why equilibrium allocations are Pareto optimal. All agents behave *as if* they are risk-neutral (payoff-maximizing) with identical beliefs. In such a situation, there are simply no differences of risk-preference or opinion on which to trade.

If $S < |\Omega| - 1$, then the consensus on risk-neutral probabilities is generally incomplete. Whenever $\Pr_h^{\mathrm{RN}}(\omega) \neq \Pr_i^{\mathrm{RN}}(\omega)$ for any ω, there exists an acceptable exchange between agents h and i, though perhaps not supported by the S available securities. An equilibrium allocation in an incomplete market is not necessarily Pareto optimal.[2] But it *can* be, depending on the particular belief structures of the agents. Call a market *operationally complete* if its competitive equilibrium $(\mathbf{x}, \mathbf{p})$ is Pareto optimal (with respect to the agents involved), even if the market contains less than $|\Omega| - 1$ securities. As a degenerate example, an empty market is operationally complete for an economy of completely identical agents. Although such a market does not support all *conceivable* trades, it does support all *acceptable* trades among the given agents.

7.2 Structured Markets: An Analogy to Graphical Models

Achieving completeness is, practically speaking, all but impossible; the required number of securities—exponential in the number of primary events—is simply too huge.[3]

In attempting to represent probability distributions over Ω, researchers in uncertain reasoning are faced with an analogous combinatorial explosion. The typical solution is to work with the factored event space, rather than the state space, and to exploit any

[2] Allocations are always efficient with respect to *available* securities, but not necessarily with respect to all states.

[3] Actually, the requirement of "only" $|\Omega| - 1$ securities is itself a significant reduction. In an economy with a set G of standard goods, the most straightforward complete market contains $|G| \cdot |\Omega|$ securities, each paying off in one good under one state realization. Arrow (1964) showed that a market where securities and goods are essentially separated, with $|\Omega|$ securities paying off in a single numeraire good plus $|G|$ *spot* markets in the standard goods, is also complete. For the purposes of belief aggregation, we need consider only the securities market. Moreover, since arbitrage arguments ensure that prices are normalized, $|\Omega| - 1$ securities are sufficient.

independencies among events using graphical models (Section 2.3.4).

Continuing the analogy, securities markets can be structured according to the directed acyclic graph D of any BN. Simply introduce one conditional security $\langle A_j|\mathbf{pa}(A_j)\rangle$ for every conditional probability $\Pr(A_j|\mathbf{pa}(A_j))$ in the network. For each event A_j with $q(j) = |\mathbf{pa}(A_j)|$ parents, this adds $2^{q(j)}$ securities, one for each possible combination of outcomes of events in $\mathbf{pa}(A_j)$. Call such a market *D-structured.* Imagine for the moment that D is fully connected (that is, no independencies are represented). Then a D-structured market contains $\sum_{j=1}^{M} 2^{j-1} = 2^M - 1 = |\Omega| - 1$ linearly independent securities, and is thus complete.

The benefit of a BN representation, and likewise a structured market, obtains when D is less than fully connected. What can be said in this case? Certainly, depending on the beliefs and utilities of the agents, inefficient allocations are possible. Nonetheless, under circumstances explored below, the smaller market may suffice for operational completeness.[4]

7.3 Compact Markets I: Consensus on Risk-Neutral Independencies

7.3.1 Risk-Neutral Independency Markets

Call a D-structured market a *risk-neutral independency market*, or an *RNI-market*, if, in equilibrium, D is an I-map of $\Pr_i^{\mathrm{RN}}$ for all agents i. That is, all agents' risk-neutral distributions agree with the independencies encoded in the market's structure. Paralleling our notation for true conditional independence, let $\mathrm{CI}^{\mathrm{RN}}[A_j, W, X]$ denote the risk-neutral conditional independence $\Pr^{\mathrm{RN}}(A_j|WX) = \Pr^{\mathrm{RN}}(A_j|W)$.

Proposition 7.1 *At equilibrium in an RNI-market,* $\Pr_h^{\mathrm{RN}}(\omega) = \Pr_i^{\mathrm{RN}}(\omega)$ *for all agents h, i and all states $\omega \in \Omega$.*

Proof. The market contains $\sum_{j=1}^{M} 2^{q(j)}$ securities, imposing an equal number of constraints on every agent's risk-neutral distribution via (7.1). For each event, I-mapness

[4] Sandholm and Lesser (1996) take a different approach to reducing the number of financial instruments required, while still allowing agents to handle unforeseen future outcomes in certain negotiation situations. Agents purchase *leveled commitment contracts* which allow them to pay a penalty fee to decommit from their obligations after the fact. In this setup, many fewer than $|\Omega| - 1$ contracts are needed, though Pareto optimal reallocation of risk may not be assured, and aggregate probabilistic information is not available. Andersson and Sandholm (1998) empirically verify that leveled commitment contracts support more (economically) efficient outcomes than full commitment and commitment-free protocols.

further imposes $2^{q(j)}(2^{j-1-q(j)}-1)$ conditional independence constraints of the form $\text{CI}_i^{\text{RN}}[A_j, \mathbf{pa}(A_j), \mathbf{pred}(A_j) - \mathbf{pa}(A_j)]$, for all combinations of outcomes of events in $\mathbf{pa}(A_j)$ and all but one combination of outcomes of events in $\mathbf{pred}(A_j) - \mathbf{pa}(A_j)$ (the remaining one is implied by the others). Then every agent's risk-neutral distribution is subject to

$$\sum_{j=1}^{M} 2^{q(j)} + 2^{q(j)}(2^{j-1-q(j)}-1) = \sum_{j=1}^{M} 2^{j-1} = 2^M - 1 = |\Omega| - 1$$

identical, linearly independent constraints. Therefore $\Pr_h^{\text{RN}} = \Pr_i^{\text{RN}}$ for all h, i. $\square$

In an RNI-market, define the *state prices* $\Pr_0(\omega) = \Pr_i^{\text{RN}}(\omega)$ as the unique probabilities over Ω that are consistent with the prices of available securities and the independencies of D. The following corollary establishes that equilibrium prices for any of the $|\Omega| - 1 - S$ "missing" securities are also derivable from $\Pr_0$.

Corollary 7.2 *Let $\langle p^{\langle 1 \rangle}, \ldots, p^{\langle S \rangle} \rangle$ be the equilibrium prices in an RNI-market. Introduce a new security $\langle E \rangle$. Then $\langle p^{\langle 1 \rangle}, \ldots, p^{\langle S \rangle}, \Pr_0(E) \rangle$ are equilibrium prices in the expanded market.*

Proof. Before the extra security is introduced, all agents' risk-neutral probabilities $\Pr_i^{\text{RN}}(E)$ already equal $\Pr_0(E)$, without buying or selling any quantity of the security. It follows that, with the additional security, the equilibrium condition (7.1) is satisfied with $x_i^{\langle E \rangle} = 0$ for all i, $p^{\langle E \rangle} = \Pr_0(E)$, and all other prices unchanged. $\square$

The number of securities in an RNI-market, $O\left(M \cdot 2^{\max\{q(j)\}}\right)$, can be exponentially smaller than the $2^M - 1$ required for traditional completeness. The following corollary shows that the more compact market supports allocations that are equally efficient.

Corollary 7.3 *Every RNI-market is operationally complete. That is, the equilibrium allocations $\mathbf{x}$ and state prices $\Pr_0$ in an RNI-market constitute an equilibrium in a (truly) complete market composed of the same agents.*

Proof. By repeated application of Corollary 7.2, we can add the $|\Omega| - 1 - S$ securities necessary to complete the market.[5] For each new security, a price consistent with $\Pr_0$, coupled with zero demand from all agents, satisfies (7.1). All complete markets, regardless of structure, support the same equilibrium allocations and state prices (Huang and Litzenberger, 1988; Mas-Colell et al., 1995; Varian, 1987). $\square$

[5]A natural set to add are the $\sum_{j=1}^{M} 2^{q(j)}(2^{j-1-q(j)}-1)$ securities of the form $\langle A_j | \mathbf{pred}(A_j) \rangle$, for all events A_j, all combinations of outcomes of $\mathbf{pa}(A_j)$, and all but one combination of outcomes of $\mathbf{pred}(A_j) - \mathbf{pa}(A_j)$.

Proposition 7.1 and its corollaries are equilibrium results only. I sketch here one possible procedure for *reaching* agreement on the market structure.[6] Begin with securities in only the M events: $\langle A_1 \rangle, \ldots, \langle A_M \rangle$. If any agent's demand for $\langle A_k | A_j \rangle$ (for any $j < k$) at price $p^{\langle A_k \rangle}$ is nonzero, then it creates a new market in $\langle A_k | A_j \rangle$. If, at some future time, the agent has zero demand for its new security, then it may retract the security. An additional condition for equilibrium is that no agent desires to create or withdraw any markets. Then, in equilibrium, it should be the case that all agents' risk-neutral independencies agree with the market structure, and that the market is operationally complete. We might want to add a transaction cost for opening new markets, so that equilibrium only ensures that risks are hedged up to a threshold cost. In some sense, this may be how real financial markets operate. New derivative assets, and customized exotic options, are created to meet the demands of investors and institutions to hedge particular risks. The absence of a market is evidence that any advantage of a more precise hedge is outweighed by the cost of opening the new market. As Varian (1987) reasons, "if people really care about achieving a certain distribution of wealth across states of nature, doesn't it seem likely that the market will offer an asset that will achieve such a pattern?"

7.3.2 Computational Complexity of Arbitrage

In this section, I show that identifying arbitrage opportunities in an RNI-market is #P-hard.[7] The analysis is from the perspective of an outsider observer O, searching for any source of risk-free profit within the market.

Imagine that, after equilibrium is reached in an RNI-market, a redundant security is introduced, say $\langle A_M \rangle$. The equilibrium price of $\langle A_M \rangle$ is already determined (Corollary 7.2): it must equal $\Pr_0(A_M) = \Pr_i^{\mathrm{RN}}(A_M)$; furthermore, if the current price does *not* equal $\Pr_0(A_M)$, then the market is not in equilibrium, and arbitrage is possible. For example, if $p^{\langle A_M \rangle} < \Pr_0(A_M)$, then O could purchase it at the going price and sell it to any of the agents at price p^* such that $p^{\langle A_M \rangle} < p^* < \Pr_i^{\mathrm{RN}}(A_M) = \Pr_0(A_M)$. Although O does not have direct access to $\Pr_0(A_M)$, it is uniquely computable given the other prices and the independence structure of D.

If O can find an arbitrage opportunity by correctly pricing the redundant security,

[6] This procedure is similar to Geiger's (1990) protocol for eliciting independence structures from experts.

[7] A problem is #P-hard if there is a polynomial-time reduction to it from the problem of counting the number of solutions to a satisfiability problem. Determining whether a solution to a satisfiability problem exists is NP-complete, so any problem that is #P-hard is also NP-hard.

then O can perform Bayesian inference, which is #P-complete (Cooper, 1990; Neapolitan, 1990).

7.4 Compact Markets II: Consensus on True Independencies

Equilibrium agreement on risk-neutral independencies may seem a somewhat awkward assumption, especially considering that the $\Pr_i^{\mathrm{RN}}$ are changing as transactions occur. Some authors argue that, since agents appear to act according to $\Pr_i^{\mathrm{RN}}$ and standard elicitation techniques reveal $\Pr_i^{\mathrm{RN}}$ (Section 2.1.5), risk-neutral probabilities are in fact no less "real" than true probabilities (Kadane and Winkler, 1988; Nau and McCardle, 1990; Nau and McCardle, 1991; Nau, 1995). However, while it seems reasonable that agents would have true independencies in common (Smith, 1990; Pearl, 1993), it is harder to justify why their risk-neutral independencies would coincide. This section develops a theory of compact markets based on consensus on true independencies. If, despite any quantitative differences between Pr and $\Pr^{\mathrm{RN}}$, an agent's true independencies were always manifest as risk-neutral independencies, then results concerning RNI-markets would carry over unchanged. Section 7.4.1 demonstrates that this is indeed the case for a subclass of agents and a subset of independencies. Section 7.4.2 discusses how the limitations on combining graphical models, developed in Chapter 5, restrict the possibility of obtaining compact markets under more general circumstances.

7.4.1 Possibility Revisited: Consensus on Markov Independencies

Section 3.3.1 established an equivalence between equilibrium prices in a securities market and the probabilities derived by the logarithmic opinion pool (LogOP), when agents have constant absolute risk aversion (CARA). Section 5.3 demonstrated that the LogOP preserves agreed-upon Markov independencies, and described how the LogOP consensus can be represented concisely with a decomposable BN. In this section, I show that a parallel result applies to structured securities markets: in economies composed of agents with CARA, markets structured according to agreed upon (true) *Markov* independencies are operationally complete. Define an *independency market*, or an *I-market*, as a D-structured market such that D is an I-map of $\Pr_i$ for all agents i (i.e., all agents' true distributions agree with the independencies in D). An I-market is *decomposable* if D is decomposable—every node's parents are fully connected (Section 2.3.4.

Consider a single agent. Let $Z = \{A_1, \ldots, A_M\}$ be the set of all events, $A_j \in Z$ a particular event, and $W \subseteq Z - A_j$ and $X = Z - W - A_j$ subsets of events. We are interested in whether the agent's Markov independencies $\mathrm{CI}[A_j, W, X]$ are reflected as a risk-neutral independencies $\mathrm{CI}^{\mathrm{RN}}[A_j, W, X]$, and are thus observable.

Proposition 7.4

$$\mathrm{CI}[A_j, W, X] \& \left(\frac{u'\left(\Upsilon^{\langle \bar{A}_j W X \rangle}\right)}{u'\left(\Upsilon^{\langle A_j W X \rangle}\right)} = \frac{u'\left(\Upsilon^{\langle \bar{A}_j W \tilde{X} \rangle}\right)}{u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right)} \right) \Rightarrow \mathrm{CI}^{\mathrm{RN}}[A_j, W, X], \tag{7.2}$$

where the second precondition must hold for all possible joint outcomes of the events in W, and all pairs $(X, \tilde{X})$ of different joint outcomes of events in X.

Proof.

$$\frac{u'\left(\Upsilon^{\langle \bar{A}_j W X \rangle}\right)}{u'\left(\Upsilon^{\langle A_j W X \rangle}\right)} = \frac{u'\left(\Upsilon^{\langle \bar{A}_j W \tilde{X} \rangle}\right)}{u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right)}$$

$$\Pr(A_j W) + \Pr(\bar{A}_j W) \frac{u'\left(\Upsilon^{\langle \bar{A}_j W X \rangle}\right)}{u'\left(\Upsilon^{\langle A_j W X \rangle}\right)} = \Pr(A_j W) + \Pr(\bar{A}_j W) \frac{u'\left(\Upsilon^{\langle \bar{A}_j W \tilde{X} \rangle}\right)}{u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right)}$$

$$\frac{\frac{\Pr(A_j W)\Pr(WX)}{\Pr(W)} u'\left(\Upsilon^{\langle A_j W X \rangle}\right)}{\frac{\Pr(A_j W)\Pr(WX)}{\Pr(W)} u'\left(\Upsilon^{\langle A_j W X \rangle}\right) + \frac{\Pr(\bar{A}_j W)\Pr(WX)}{\Pr(W)} u'\left(\Upsilon^{\langle \bar{A}_j W X \rangle}\right)} =$$

$$\frac{\frac{\Pr(A_j W)\Pr(W\tilde{X})}{\Pr(W)} u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right)}{\frac{\Pr(A_j W)\Pr(W\tilde{X})}{\Pr(W)} u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right) + \frac{\Pr(\bar{A}_j W)\Pr(W\tilde{X})}{\Pr(W)} u'\left(\Upsilon^{\langle \bar{A}_j W \tilde{X} \rangle}\right)}$$

$$\frac{\Pr(A_j W X) u'\left(\Upsilon^{\langle A_j W X \rangle}\right)}{\Pr(A_j W X) u'\left(\Upsilon^{\langle A_j W X \rangle}\right) + \Pr(\bar{A}_j W X) u'\left(\Upsilon^{\langle \bar{A}_j W X \rangle}\right)} =$$

$$\frac{\Pr(A_j W \tilde{X}) u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right)}{\Pr(A_j W \tilde{X}) u'\left(\Upsilon^{\langle A_j W \tilde{X} \rangle}\right) + \Pr(\bar{A}_j W \tilde{X}) u'\left(\Upsilon^{\langle \bar{A}_j W \tilde{X} \rangle}\right)}$$

$$\frac{\Pr^{\mathrm{RN}}(A_j W X)}{\Pr^{\mathrm{RN}}(A_j W X) + \Pr^{\mathrm{RN}}(\bar{A}_j W X)} = \frac{\Pr^{\mathrm{RN}}(A_j W \tilde{X})}{\Pr^{\mathrm{RN}}(A_j W \tilde{X}) + \Pr^{\mathrm{RN}}(\bar{A}_j W \tilde{X})}$$

$$\Pr^{\mathrm{RN}}(A_j | W X) = \Pr^{\mathrm{RN}}(A_j | W \tilde{X})$$

□

The second precondition in (7.2) says that the ratio of marginal utility in states where A_j does not occur to marginal utility in states where A_j does occur cannot depend of the outcomes of events in X. This is true (and indeed $\Pr^{\mathrm{RN}} = \Pr$) if the agent's marginal utility u' is constant across states. This holds if the agent is risk neutral, and holds approximately if utility is state-independent and $\Upsilon^{\langle\omega_j\rangle} \approx \Upsilon^{\langle\omega_k\rangle}$. But this approximation is not realistic for an agent engaged in trading securities, since a central role of the market is precisely to enable the transfer of wealth across states.

Let $\Upsilon^{\langle A_j W\rangle}$ be the agent's payoff from all securities that depend only the outcomes of events in $A_j \cup W$. Examples are $\langle A_j\rangle$, $\langle A_j W\rangle$, and $\langle A_j|W\rangle$, which return the same dollar amount regardless of the realizations of events in $X = Z - W - A_j$. Similarly, let $\Upsilon^{\langle WX\rangle}$ be the payoff from securities that do not depend on A_j.

Suppose that the agent exhibits CARA, and that its payoffs are separable according to $\Upsilon^{\langle A_j WX\rangle} = \Upsilon^{\langle A_j W\rangle} + \Upsilon^{\langle WX\rangle} - \Upsilon^{\langle W\rangle}$. Separability essentially means that any of the agent's securities (or prior stakes) whose payoff depends on A_j cannot also depend on events in X. In this case,

$$\frac{u'\left(\Upsilon^{\langle \bar{A}_j WX\rangle}\right)}{u'\left(\Upsilon^{\langle A_j WX\rangle}\right)} = \frac{u'\left(\Upsilon^{\langle \bar{A}_j W\rangle} + \Upsilon^{\langle WX\rangle} - \Upsilon^{\langle W\rangle}\right)}{u'\left(\Upsilon^{\langle A_j W\rangle} + \Upsilon^{\langle WX\rangle} - \Upsilon^{\langle W\rangle}\right)} = \frac{ce^{-c\Upsilon^{\langle \bar{A}_j W\rangle}} e^{-c\Upsilon^{\langle WX\rangle}} e^{c\Upsilon^{\langle W\rangle}}}{ce^{-c\Upsilon^{\langle A_j W\rangle}} e^{-c\Upsilon^{\langle WX\rangle}} e^{c\Upsilon^{\langle W\rangle}}}$$

$$= \frac{ce^{-c\Upsilon^{\langle \bar{A}_j W\rangle}} e^{-c\Upsilon^{\langle W\tilde{X}\rangle}} e^{c\Upsilon^{\langle W\rangle}}}{ce^{-c\Upsilon^{\langle A_j W\rangle}} e^{-c\Upsilon^{\langle W\tilde{X}\rangle}} e^{c\Upsilon^{\langle W\rangle}}} = \frac{u'\left(\Upsilon^{\langle \bar{A}_j W\rangle} + \Upsilon^{\langle W\tilde{X}\rangle} - \Upsilon^{\langle W\rangle}\right)}{u'\left(\Upsilon^{\langle A_j W\rangle} + \Upsilon^{\langle W\tilde{X}\rangle} - \Upsilon^{\langle W\rangle}\right)} = \frac{u'\left(\Upsilon^{\langle \bar{A}_j W\tilde{X}\rangle}\right)}{u'\left(\Upsilon^{\langle A_j W\tilde{X}\rangle}\right)}$$

Thus the constraint on utility in (7.2) is satisfied, and any Markov independencies are observable.

We are now in a position to derive the main result of this section.

Proposition 7.5 *When all agents have CARA, every decomposable I-market is an RNI-market.*

Proof. Let W_j be the set of direct parents and direct children of event A_j, and X_j all other events. From decomposability and I-mapness, we can infer that

1. $\mathrm{CI}_i[A_j, W_j, X_j]$ for all agents i and events j,
2. none of the securities $\langle A_j|\mathbf{pa}(A_j)\rangle$ that are contingent on A_j depend on X_j, and
3. none of the securities $\langle A_k|\mathbf{pa}(A_k)\rangle$ such that $A_j \in \mathbf{pa}(A_k)$ that are conditional on A_j depend on X_j.

Items 2 and 3 ensure separability of payoffs from the available securities (we assume that any prior stakes are also separable). Then, invoking Proposition 7.4, $\mathrm{CI}_i^{\mathrm{RN}}[A_j, W_j, X_j]$ for all agents i and events j. As a result, D is an I-map of every $\Pr_i^{\mathrm{RN}}$ regardless of allocations or prices, including those at equilibrium. □

Proposition 7.1 and Corollaries 7.2 and 7.3 are immediately applicable. In particular, for agents with CARA, every decomposable I-market is operationally complete.

7.4.2 Impossibility Revisited

The examples in Chapter 5 where the LinOP (Example 5.1) and LogOP (Example 5.2) fail to preserve unanimous independencies have direct, negative implications for the possibility of compact I-markets for larger classes of agents or independencies.

Let two agents have beliefs over two events, as prescribed in Example 5.1 and pictured in Figure 5.1(a). Both agents agree that the events are independent. Suppose additionally that they both have GLU with equal wealth parameters. In a complete market (of $2^2 - 1 = 3$ linearly independent securities), the unique state prices (and the unique consensus risk-neutral probabilities) are the same as derived by the LinOP in the example, and do *not* reflect the independence. Thus an I-market consisting of only the two securities $\langle A_1 \rangle$ and $\langle A_2 \rangle$ is not operationally complete.

In Example 5.2 (Figure 5.1(d)), two agents agree that, among three primary events, A_1 and A_2 are independent, and A_3 depends on both. If both agents have CARA, then the unique state prices in a complete market equal the LogOP consensus probabilities, and do not preserve the independence between A_1 and A_2. Thus an I-market mirroring the agreed-upon polytree structure of Figure 5.1(d) (containing the six securities $\langle A_1 \rangle$, $\langle A_2 \rangle$, $\langle A_3 | A_1 A_2 \rangle$, $\langle A_3 | A_1 \bar{A}_2 \rangle$, $\langle A_3 | \bar{A}_1 A_2 \rangle$, and $\langle A_3 | \bar{A}_1 \bar{A}_2 \rangle$) is not operationally complete.

Chapter 8

Conclusions and Future Work

8.1 Conclusions

De Finetti's (1974) standard of rationality—the avoidance of guaranteed losses—seems the epitome of minimalism and innocuousness. That such a baseline assumption requires (and is assured by) behavior attuned to some underlying probability distribution is an at once astounding and persuasive argument for an agent to maintain subjective probabilistic beliefs.

And why should a group of agents tolerate arbitrage "leaks" any more readily than an individual? Sealing off all such leaks assures a benefit for least one agent, and at no cost to any other. The required sealant is the same as for an individual: outward rationality, or behavior indistinguishable from that guided by an aggregate probability distribution.

Whether motivated thusly, or by more practical concerns,[1] many researchers have embarked on the search for belief aggregation functions, or opinion pools. In this dissertation, I have developed and analyzed a market-based approach to forming aggregate beliefs, and have investigated the use of graphical models for representing the combined beliefs.

Figure 8.1 redisplays the same Venn diagram of Figure 1.1 in Chapter 1, filling in the main results of this dissertation under each of the three categories of contributions. Section 8.1.1 reviews the first category, covering motivations for adopting a market framework and contrasting market-based aggregation with traditional pooling functions. Section 8.1.2 discusses the second category of results: when and how can aggregate beliefs be represented compactly using graphical models? Section 8.1.3 recounts how these two

[1] The summarization of knowledge from multiple sources is a time-tested heuristic of scientific modeling, and appears to be the modus operandi of human and animal reasoners.

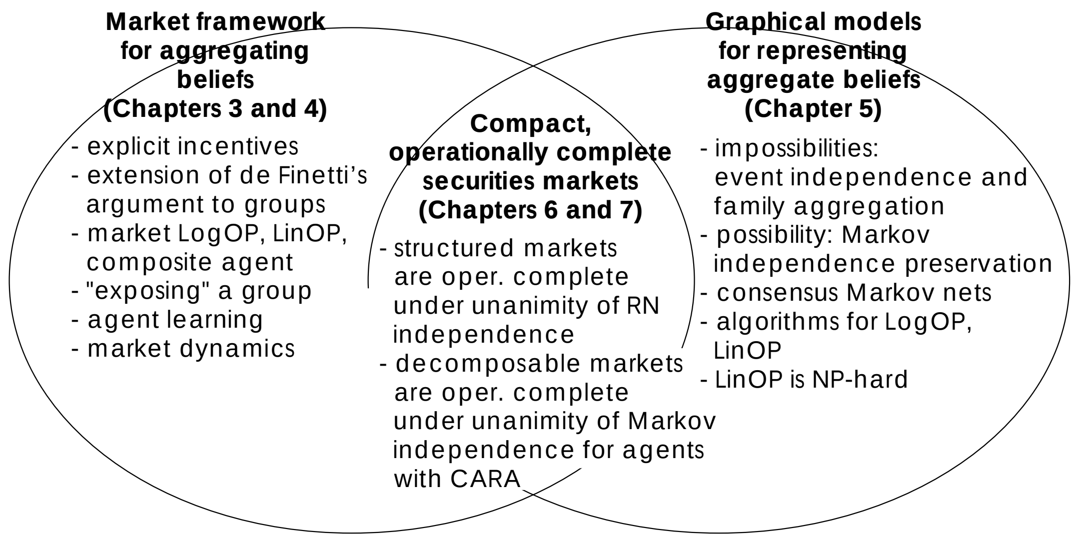


Figure 8.1: Results of this dissertation organized into the three categories of Figure 1.1.

sets of results converge to admit, in some circumstances, compact markets that still derive market probabilities for all states of nature $\omega \in \Omega$, but that are composed of exponentially fewer securities than would otherwise be necessary. Finally, Section 8.1.4 recharacterizes several of the contributions listed in Figure 8.1, by classifying them according whether they reveal new limitations on aggregation procedures or whether they expose new possibilities.

8.1.1 A Securities Market Framework

The absence of arbitrage is taken as given within the theory of securities markets.[2] Just as de Finetti's rational agent displays beliefs in the form of probabilities, so the equilibrium prices in a securities market exhibit all the properties of a legal probability distribution. Moreover, at equilibrium, all agents are in a state of consensus on risk-neutral probabilities (across available securities), and these probabilities equal the going prices. From this standpoint, then, security prices appear to have a natural interpretation as aggregate probabilities.

Markets also naturally support coordination among self-interested, decentralized agents. Each agent maximizes its own utility, and needs not rely on (or act as) a central coordinator. The mechanism affords limited privacy, as beliefs are not directly revealed. Lines of communication can be kept relatively sparse: each security can be auctioned at a distinct

[2]When arbitrage opportunities do momentarily arise, so the argument goes, some agent will quickly take advantage of the situation, thereby driving prices so as to close off the arbitrage.

site, and agents need communicate only with auctions that they are currently interested in.

Monetary incentives for participation and honesty are built in, and directly drive the aggregation. In contrast, standard opinion pools usually separate elicitation from aggregation. The burden of properly aligning incentives is placed squarely on the elicitation mechanism.

The reward structure of a securities market facilitates the dissemination of information. Well-informed agents are incented to trade in a manner that publicly reveals their knowledge. The cost to other agents for (indirectly) obtaining this information is at most that of opening and maintaining the relevant market.

Chapter 3 demonstrates that the two most basic aggregation functions—the linear and logarithmic opinion pools (LinOP, LogOP)—arise as special cases of this market framework, results that bear a resemblance to those derived by Wilson (1968) and Rubinstein (1974, 1975, 1976) in somewhat different contexts. Expert weights in the opinion pools correspond to normalized risk tolerances in the market model. While the former are notoriously ill-defined (French, 1985), the latter at least have a sharp meaning in decision-theoretic terms.

Chapter 4 casts the problem of updating based on prices as analogous to the supra Bayesian's problem of updating based on the beliefs of a single expert. I describe a scenario in which an agent's perception of its own precision relative to the market factors into the quantities of securities it demands, thereby affecting its contribution to prices. In this way, a notion of confidence-dependent weights emerges naturally when agents learn from prices.

When the market can be viewed as a rational composite agent (whose beliefs equal the prices), the interpretation of prices as consensual beliefs seems on most solid ground. Chapter 3 identifies two homogeneity constraints on agents (the same two that entail the LinOP and LogOP) that support the composite agent rationalization, again paralleling findings by Wilson (1968) and Rubinstein (1974). I defer further commentary on the existence and nonexistence of composite agents to Section 8.1.4.

8.1.2 A Graphical Modeling Perspective

Early formalisms for uncertain reasoning in artificial intelligence were largely ad hoc (e.g., certainty factors (Buchanan and Shortliffe, 1984)). Storing and manipulating joint probability distributions, though perhaps recognized as "correct", were seen as hopelessly intractable.

The advent of graphical modeling languages, in particular Bayesian networks (BNs), changed this view dramatically. The key to their success is the implicit representation of conditional independence as the absence of edges in a graph, matched with a corresponding factorization of the joint distribution.

Many applications in statistical modeling also benefit from the graphical perspective (Darroch et al., 1980). Characterizing the independencies among variables paints a useful and descriptive picture of the underlying qualitative workings of the modeled system. In fact, the fundamental importance of independence has led many researchers to presume that, whenever probability distributions are aggregated, unanimously held independencies should always be preserved. Section 8.1.4 covers the new impossibility and possibility results reported in this dissertation regarding the aggregation of graphical models.

8.1.3 Structured Markets

Complete securities markets are viewed in much the same light as were joint probability distributions—theoretically ideal, but practically unachievable. In fact, the number of securities required for completeness is precisely the dimensionality of a probability distribution over the full state space Ω.

Securities markets can be structured in direct analogy to BNs (Section 7.2). As with graphical models, if sufficient independencies are encoded in the structure, the size of the market can be exponentially reduced. Although structured markets are not complete in the traditional sense, Chapter 7 derives conditions under which they are nonetheless *operationally* complete. A sufficient condition is that, in equilibrium, all agents' risk-neutral independencies agree with those encoded in the market's structure. In this case, prices uniquely define a complete probability distribution over Ω, and allocations support a Pareto optimal allocation of risk among the given agents. It seems difficult to formulate an intuitive justification for supposing that a group of agents would be in agreement on risk-neutral independencies. Section 8.1.4 reviews the special case when agreement on true independencies is sufficient to yield operationally complete markets. That section also describes why agreement on true independencies is not sufficient in general.

8.1.4 Impossibilities and Possibilities

Positive progress in belief aggregation research, though certainly not lacking, is circumscribed by controversy and impossibilities. A proliferation of results in the 1980s, exemplified by groundbreaking contributions by Genest (1984a, 1984b, 1984c) and his coauthors (Genest et al., 1986; Genest and Wagner, 1987), do much to demarcate the impassable boundaries. Contributions in this dissertation as well fall on both sides of the impossibility fence. Section 8.1.4 catalogues those results that entail new limitations, Section 8.1.4 those that uncover new possibilities.

For the Pessimist...

A subset of results in this dissertation further confine and confound the search for reasonable aggregation procedures, by extending the impossibility theorems to new domains, and by raising new concerns.

A series of theorems (Lehrer and Wagner, 1983; Wagner, 1984) culminating in that of Genest and Wagner (1987) show that very weak and reasonable constraints on an aggregation function are enough to rule out independence preservation (i.e., the retention of all agreed-upon independencies within the aggregate distribution). But these theorems apply to functions that preserve all possible independencies between any events—even those not representable in a graphical model. A potential loophole remained that some reasonable function might preserve the independencies among primary events in a graphical model. Indeed, in Chapter 5, we see that the same conditions sufficient to rule out general independence preservation are *not* sufficient to rule out this weaker form. However, I show that, with the additional (uncontroversial) assumption of unanimity, the impossibility returns.

This result resurfaces in the study of structured securities markets in Chapter 7. The intuitive inclination to structure the market according to agreed-upon independencies proved fatally flawed. Prices in a securities market are essentially the output of an aggregation function—the one defined by market equilibrium—and are thus subject to all the general limitative theorems.

I derive a second impossibility theorem in Chapter 5 regarding the combination of BNs. A natural policy—that other authors have advocated or assumed—is to confine the aggregation locally, within each conditional probability table of the BN. I prove that any such local aggregation function necessarily fails to satisfy either unanimity or nondictatorship, two seemingly incontrovertible assumptions.

This theorem also finds application in context of securities markets. In Section 3.5, I describe a procedure of opening and withdrawing securities markets so that each revealed price is essentially the result of a local aggregation of the participating agents' beliefs. The market mechanism satisfies both unanimity (Proposition 3.5) and nondictatorship. Thus, apart from exceptional circumstances (e.g., all beliefs are identical, or the "market" is actually an individual) these prices do *not* conform to a coherent probability distribution, and the group is not outwardly rational.

Other results demonstrate that desirable operations, while not impossible, are instead (worse-case) intractable. Someone interested in computing the LinOP of several probability distributions, each represented as a BN, would not want to construct a consensus BN, as it would in general be fully connected. This suggests keeping the individual BNs separate, and computing the LinOP of any desired query at runtime. In Proposition 5.6, I prove that performing this computation is NP-hard, even if answering the same query is easy within each individual BN.

Even the positive results in Chapter 5 describing BN representations of the logarithmic opinion pool (LogOP), discussed in more depth below, are shaded by potential computational barriers. The consensus network structure must be made decomposable, a process that can increase the size of the representation exponentially. Similar computational concerns arise in Chapters 6 and 7 when the structure of the securities market is required to be decomposable. I also show in Chapter 7 that properly pricing securities and finding arbitrage opportunities within in a compact market is NP-hard.

For the Optimist...

On the other hand, some results in this dissertation can be characterized as possibility results. Each identifies a weakening of an impossibility theorem that exposes a nontrivial solution.

Market composite agents are not always possible, depending on the beliefs and utilities of the agents (Raiffa, 1968). As already discussed, however, certain assumptions regarding agents' utilities do yield composite agents whose beliefs equal the market's prices. When they exist, composite agents embody the standard of outward rationality at the group level that, for many, justifies the search for aggregate beliefs in the first place.

Full rationality requires not only maintenance of a probability distribution, but also proper Bayesian updates as evidence is observed. In Section 4.2, I extend the composite agent formalism to encompass belief updates. The setting is a multiperiod market, where a single security is traded repeatedly over time. For certain types of agents, their distribution of wealth at time t is a Beta distribution that precisely reflects the number of positive outcomes observed so far out of the t trials. The security's price tracks the expectation of this Beta distribution, and so the dynamics of price can be rationalized as the Bayesian updates of a composite agent.

Another possibility result arises by weakening Genest and Wagner's (1987) axiom for preserving independence, to require only the preservation of Markov independencies. Chapter 5 demonstrates that the LogOP does in fact maintain all agreed-upon Markov independencies. This suggests that, if the preservation of independence structure is important—and many authors argue that it is (Laddaga, 1977; Raiffa, 1968)—then the LogOP may be the most viable option. Markov independencies play an important role in the theory of graphical models, and are precisely the type representable in Markov networks (MNs) and decomposable BNs. I describe procedures for constructing MN and BN structures consistent with the LogOP consensus. I also delineate an algorithm for computing all of the conditional probability tables in a LogOP-consensus BN. This structured representation is potentially exponentially smaller than the standard "flat" representation.

The preservation of Markov independencies has a direct corollary in Chapter 7's investigation of structured securities markets. For a certain class of agents, true Markov independencies are always observable as risk-neutral independencies. Thus, if all agents are of this type, all beliefs agree with the independencies encoded in the market structure, and this structure is decomposable, then the market is operationally complete.

8.2 Future Work and Applications

8.2.1 Group Decision Making

A natural practical application of belief aggregation is as a sub-procedure within the more general context of *group decision-making*, where agents' beliefs *and* utilities are combined to enable inference of group decisions. Future plans include identifying and evaluating appropriate generalizations in pursuit of a market-based approach to group decision-making in situations involving both asymmetric uncertainty and heterogeneous preferences.

Savage (1954) and von Neumann and Morgenstern (1953) provide compelling axiomatic arguments that rational individuals should maintain both beliefs and utilities, and should maximize expected utility. Groups of agents are then dealt a difficult decision-making dilemma: either (1) adopt consensus beliefs and consensus utilities, or (2) violate these fundamental axioms of rationality. Magnifying the problem are a plethora of impossibility theorems associated with the former (Hylland and Zeckhauser, 1979; Mongin, 1995; Schervish et al., 1991; Seidenfeld et al., 1989; Seidenfeld and Schervish, 1991).

Notice that the market's composite agent is actually a group decision maker. Although I have concentrated on how the composite agent's beliefs (the prices) depend on the constituent agents' beliefs, Propositions 3.8 and 3.9 also define how the composite agent's utilities depend on the agents' utilities. Recall Hylland and Zeckhauser's (1979) impossibility theorem regarding group decision making (Section 2.1.6). The authors make what seems a natural assumption: combined beliefs should depend only on the agents' beliefs and combined utilities should depend only on the agents' utilities. This separation is in the spirit of most decision-theoretic treatments, including those of Savage and von Neumann and Morgenstern. Indeed, Hylland and Zeckhauser, in contemplating which axioms to weaken, characterize violation of "the principle of separate aggregation" as a "drastic alternative".

Yet the composite agent escapes this impossibility precisely because it *does* violate separation—its beliefs depend on both the agents' beliefs and their utilities. Other results support the reasonableness of discarding the separation axiom. Kadane and Winkler (1988) show that agents' outwardly observable beliefs are almost always confounded with their utilities anyway, and Nau (1995) develops a framework for solving a single agent's decision problem without requiring separation. It seems that weakening the separation criterion for group decision making is reasonable, and that the market framework offers a promising mechanism for doing so.

The standard of no arbitrage, or the avoidance of certain losses, pervades theories of both individual rationality (de Finetti, 1974; Savage, 1954) and financial markets.

> Many important results of financial economics are based squarely on the hypothesis of no arbitrage, and it serves as on of the most basic unifying principles of the study of financial markets. (Varian, 1987)

No arbitrage has also been established as the underlying principle behind correlated equilibria in noncooporative games (Nau and McCardle, 1990; Nau, 1992) and competitive equilibria in classical exchange economies (Nau and McCardle, 1991). No arbitrage is the closest thing we have to a unifying standard of rationality in both single-agent and multiagent settings, and will likely be at the heart of any reasonable group decision-making procedure. Moreover, market mechanisms seem naturally suited for enforcing no arbitrage at both the individual and group level.

Not all groups can be viewed as a composite agent (Raiffa, 1968), and, in the absence of complete markets, aggregate behavior may be distinguishable from individual behavior (Section 3.5). Characterizing the necessary and sufficient conditions for the existence of a composite agent is an important and open problem. Such characterizations must consider both the types of agents and the extent to which their aggregate behavior can be measured. If measurements are taken by observing the prices of securities, some relevant factors (explored in Sections 3.5 and 4.2) include the makeup of securities, whether securities can be added and removed, the availability of outside securities unbeknownst to the observer, and whether measurements are one-shot or repeated over time.

8.2.2 Securities Markets in Computational Resources

Researchers have proposed opening markets in computational resources—for example, CPU cycles (Regev and Nisan, 1998) or network bandwidth (MacKie-Mason and Varian, 1995)—in order to provide varying levels of service based on users' needs and resources.

The next logical step is to open *securities* markets in computational resources. For example, a security that pays off contingent on network failure provides both diagnostic and operational benefits: first, the going price is a proxy for the probability that the connection will fail—of value to system administrators and users; second, users can hedge against the unforeseen loss of a critical connection.

8.2.3 Securities Markets in Task Allocation Economies

Walsh and Wellman (1998, 1999) describe a market protocol for solving a class of task allocation problems. End consumers pay for the completion of certain tasks. Producers can complete and sell tasks to consumers or to other producers by purchasing needed supplies, or by paying yet other producers to complete relevant subtasks. Suppliers sell primary resources to producers. In this manner, agents form acyclic supply chains (representable as bipartite graphs) from suppliers through a potential hierarchy of producers to end consumers. Distinct supply chains are not always simultaneously viable, since agents contend for the same resources. Standard auction algorithms determine the prices and allocations of primary resources and tasks. The median performance of their protocol, tested on sets of randomly generated problems, is near optimal (Walsh and Wellman, 1999).

Their model is essentially one of task allocation under certainty. If any unforeseen changes do occur—for example, resources become unavailable, an agent does not fulfill its obligations, or the dependency among tasks is altered—then, in order to return to a good solution, the allocation must generally be recomputed. Upon such a reallocation, all agents that prefer the original solution probably deserve compensation. But at whose expense and to what degree? If the precipitating change was the loss or damage of a primary resource, no particular agent may be directly at fault.

A natural generalization of Walsh and Wellman's model would account more explicitly for uncertainty in the availability of primary resources. For example, suppose that, in addition to purchasing primary resources, producers and consumers purchase securities that pay off contingent on the availability of the resources. Dual benefits are again evident: prices reveal a consensus measure of the reliability of the resource supply, and producers and consumers can arrange for compensation in the event that a supplier fails to deliver on promised resources. This compensation may offset the potential cost of reallocation. If there are conditional independencies among resource failure events, structured securities markets might be employed.

8.2.4 Ensemble Learning of Probability Distributions

Learning BN structures and parameters to fit given data—and to predict unseen data—is a growing and successful subfield within machine learning. Another trend is so-called *ensemble* learning, in which the results of several learning algorithms are combined together to produce more accurate predictions (Dietterich, 1997). Bagging (Breiman, 1996) and boosting (Freund and Schapire, 1996) are two popular and effective ensemble learning algorithms. Ensemble algorithms typically employ weighted or unweighted voting schemes to aggregate predictions. As in the opinion pool literature, many competing methods for combination and weight assignment have been proposed, with none emerging as the clear winner.

The adaptive market framework explored in Section 4.2 can be thought of as an ensemble algorithm for learning probabilities, at least when agents are decreasingly risk averse. Learning occurs on two levels: first, within each period, agents learn from others by updating their beliefs based on prices; second, across periods, agents with better probabilistic models garner more wealth and hence greater influence on prices. Agents may specialize in subsets of Ω and still achieve success; the market naturally aggregates several such specialists to yield state prices across all $\omega \in \Omega$ (assuming enough securities). Section 4.2.1 establishes that learning is well behaved in the simple case of a single event and agents with GLU: the price converges toward the observed frequency of the event. Ensemble weights are proportional to agents' wealth parameters, and accurate agents are naturally rewarded over time at the expense of inaccurate ones. In an economy of agents trading in the conditional securities of a particular D-structured market, adaptive forces may drive ensemble learning of the CPTs of the corresponding BN. If agents can create and withdraw markets, BN structures might be learned as well.

Another promising avenue to investigate are the implications, if any, of the various impossibility theorems in belief aggregation (particularly those involving independence preservation) and social choice theory for ensemble learning algorithms.

8.2.5 NP-Markets: How to Get Everyone Else to Solve Your Problems

In May 1997, IBM's Deep Blue became the first computer to defeat a reigning world chess champion. Along with the \$700,000 winner's purse, the programming team won the \$100,000 Fredkin prize, an amount set aside in 1980 to stimulate computer chess research. A \$1.5 million prize still awaits the first computer champion of the Chinese board game Go.[3] Every year the Loebner prize[4], and \$2000, goes to the computer program that is judged most human-like; a \$100,000 prize is earmarked for the first computer to fool a judge into believing that it is human, a gold standard for artificial intelligence first popularized by Turing (1950). In January 1997, RSA Data Security sponsored a \$10,000 prize for the first person to decode a particular 56-bit DES-encrypted message.[5] A distributed collection of computers, coordinated across the Internet, accomplished the task, and collected the prize money, in June of the same year.[6]

By sponsoring contests, a funding agent can provide incentive for researchers to tackle the problems that it wants solved. However, significant rewards of this type are rare, and are certainly negligible compared to the number of challenging problems of interest. Few

[3] The Ing prize, sponsored by Acer Incorporated and the Ing Chang-Ki Wei-Chi (Go) Educational Foundation.

[4] http://www.loebner.net/Prizef/loebner-prize.html

[5] http://www.rsa.com/des/

[6] http://www.frii.com/ rcv/deschall.htm

individuals or groups can afford to back a meaningful prize and advertise it sufficiently. Others may simply not desire such publicity, especially if the problem itself is proprietary.

Hanson's proposed Idea Futures markets would support a more pervasive reward structure for scientific research. In practice, however, it is often difficult to precisely define each security's payoff-triggering event in a way that foresees all possible eventualities. As a result, an unbiased human judge is generally required for every security. When a security is vaguely defined or the judge is not trusted, agents are usually wary of trading in it, and its price may have little or no informational value.

There *is* a class of problems that are both interesting and difficult to solve, and yet any proposed solution can be easily and precisely verified: namely, the class of NP-complete problems. Imagine, for example, that a circuit design company has a library of particularly hard satisfiability (SAT) problems that it would like solved. The company publishes each problem on the web as a security that pays off \$1 if and only if the problem is proved satisfiable by a certain date. The company remains anonymous and the problem descriptions are randomized to cloak their underlying purpose, maintaining solution invariance. Agents from around the world devote excess CPU time to solving the problems.[7] If an agent finds a solution, it buys up large quantities of the security; similarly, if an agent proves a problem unsatisfiable, it sells *en masse*. Speculators trade based on their assessment of the probability that a solution will be found. The company itself, or others with a stake in the solutions, or even those simply interested in funding SAT-solver algorithms research, sell the security to subsidize the market. Many variations are possible, including securities that pay off only to the first agent to provide a solution, or markets in optimization problems that pay off for solutions better than some constant bound, or proportional to the quality of the solution. Such "NP-markets" would provide direct monetary incentives for the development of better algorithms for solving difficult computational problems.

[7] Brewer (1999) devises another interesting scheme to incent distributed agents to solve an NP-complete optimization problem—in this case, to compute a Pareto optimal allocation in a combinatorial auction. Traditionally, the auctioneer carries out the computation. Brewer suggests instead a procedure where bidders in the auction must settle for the current (possibly Pareto dominated) allocation unless someone demonstrates that a better solution, with smaller total consumer surplus, exists. Bidders whose utilities increase in the improved solution have incentive to report the solution if known. Bidders may also receive a percentage of the improvement when they report a better solution, with percentages increasing as the auction's clearing time approaches.

Bibliography

Andersson, M. R. and Sandholm, T. W. (1998). Leveled commitment contracts with myopic and strategic agents. In *Fifteenth National Conference on Artificial Intelligence*, pages 38–45.

Arrow, K. J. (1963). *Social Choice and Individual Values.* Yale University Press, second edition.

Arrow, K. J. (1964). The role of securities in the optimal allocation of risk-bearing. *Review of Economic Studies*, 31(2):91–96. Translated from *E*conométrie 1953.

Arrow, K. J. (1967). Values and collective decision-making. In Laslett, P. and Runciman, W. G., editors, *Philosophy, Politics and Society (third series)*, pages 215–232. Basil Blackwell Oxford.

Bartholdi, J. J. and Orlin, J. A. (1991). Single transferable vote resists strategic voting. *Social Choice and Welfare*, 8:341–354.

Bartholdi, J. J., Tovey, C. A., and Trick, M. A. (1989a). The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6:227–241.

Bartholdi, J. J., Tovey, C. A., and Trick, M. A. (1989b). Voting schemes for which it can be difficult to tell who won the election. *Social Choice and Welfare*, 6:157–165.

Bartholdi, J. J., Tovey, C. A., and Trick, M. A. (1992). How hard is it to control an election? *Mathematical and Computer Modelling*, 16(8–9):27–40.

Benediktsson, J. A. and Swain, P. H. (1992). Consensus theoretic classification methods. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(4):688–704.

Berger, J. O. (1985). *Statistical Decision Theory and Bayesian Analysis.* Springer-Verlag, New York, second edition.

Black, D. (1987). *The Theory of Committees and Elections.* Kluwer Academic, Norwell, MA.

Bonduelle, Y. (1987). *Aggregating expert opinions by resolving sources of disagreement.* PhD thesis, Stanford University.

Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(1).

Brewer, P. J. (1999). Decentralized computation procurement and computational robustness in a smart market. *Economic Theory*, 13:41–92.

Buchanan, B. G. and Shortliffe, E. H. (1984). *Rule-Based Expert Systems.* Addison-Wesley, Reading, MA.

Cheng, J. Q. and Wellman, M. P. (1998). The WALRAS algorithm: A convergent distributed implementation of general equilibrium outcomes. *Computational Economics*, 12:1–24.

Chyu, C. C. (1991a). *Computing Probabilities for Probabilistic Influence Diagrams.* PhD thesis, University of California at Berkeley.

Chyu, C. C. (1991b). Decomposable probabilistic influence diagrams. *Probability in the Engineering and Informational Sciences*, 5:229–243.

Clemen, R. T. and Winkler, R. L. (1993). Aggregation of point estimates: A flexible modeling approach. *Management Science*, 39(4):501–515.

Cooke, R. M. (1991). *Experts in Uncertainty: Opinion and Subjective Probability in Science.* Environmental Ethics and Science Policy Series. Oxford University Press, New York.

Cooper, G. (1990). The computational complexity of probabilistic inference using Bayes belief networks. *Artificial Intelligence*, 42:393–405.

Cormen, T. H., Leiserson, C. E., and Rivest, R. L. (1990). *Introduction to Algorithms.* MIT Press, Cambridge, MA.

Cox, R. T. (1946). Probability, frequency, and reasonable expectation. *American Journal of Physics*, 17:1–3.

Dagum, P. and Luby, M. (1993). Approximating probabilistic inference in Bayesian belief networks is NP-hard. *Artificial Intelligence*, 60:141–153.

Dalkey, N. C. (1972). An impossibility theorem for group probability functions. P-4862, The Rand Corporation, Santa Monica, CA.

Dalkey, N. C. (1975). Toward a theory of group estimation. In Linstone, H. A. and Turoff, M., editors, *The Delphi Method: Techniques and Applications*, pages 236–261. Addison-Wesley, Reading, MA.

Darroch, J. N., Lauritzen, S. L., and Speed, T. P. (1980). Markov fields and log-linear interaction models for contingency tables. *Annals of Statistics*, 8(3):522–539.

de Finetti, B. (1974). *Theory of Probability: A Critical Introductory Treatment*, volume 1. J. Wiley, New York.

Degroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121.

Degroot, M. H. and Mortera, J. (1991). Optimal linear opinion pools. *Management Science*, 37(5):546–558.

Delcher, A. L., Grove, A. J., Kasif, S., and Pearl, J. (1996). Logarithmic-time updates and queries in probabilistic networks. *Journal of Artificial Intelligence Research*, 4:37–59.

Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *Annals of Mathematical Statistics*, 38:325–339.

Dietterich, T. G. (1997). Machine learning research: Four current directions. *Artificial Intelligence Magazine*, 18(4):97–136.

Doyle, J. and Wellman, M. P. (1991). Impediments to universal preference-based default theories. *Artificial Intelligence*, 49:97–128.

Drèze, J. H. (1987). Market allocation under uncertainty. In Drèze, J. H., editor, *Essays on Economic Decisions under Uncertainty*, pages 119–143. Cambridge University Press, Cambridge.

Dubois, D. and Prade, H. (1988). *An Approach to Computerized Processing of Uncertainty.* Plenum Press, New York.

Eisenberg, E. and Gale, D. (1959). Consensus of subjective probabilities: The pari-mutuel method. *Annals of Mathematical Statistics*, 30:165–168.

Fagin, R., Halpern, J. Y., Moses, Y., and Vardi, M. Y. (1996). *Reasoning About Knowledge*. MIT Press, Cambridge, MA.

Faria, A. E. and Smith, J. Q. (1996). Conditional external bayesianity in decomposable influence diagrams. In *Bayesian Statistics 5*, pages 551–560.

Fishburn, P. C. (1970). Arrow's impossibility theorem: Concise proof and infinite voters. *Journal of Economic Theory*, 2(1):103–106.

Fishburn, P. C. (1987). *Interprofile Conditions and Impossibility*. Harwood Academic, NY.

Forsythe, R. and Lundholm, R. (1990). Information aggregation in an experimental market. *Econometrica*, 58(2):309–347.

Forsythe, R., Nelson, F., Neumann, G. R., and Wright, J. (1992). Anatomy of an experimental political stock market. *American Economic Review*, 82(5):1142–1161.

French, S. (1985). Group consensus probability distributions: A critical survey. *Bayesian Statistics*, 2:183–202.

Freund, Y. and Schapire, R. (1996). Experiments with a new boosting algorithm. In *Proceedings of the Thirteenth International Conference on Machine Learning*, pages 148–156, San Francisco, CA.

Geiger, D. (1990). *Graphoids: A Qualitative Framework for Probabilistic Inference*. PhD thesis, UCLA.

Genest, C. (1984a). A characterization theorem for externally Bayesian groups. *Annals of Statistics*, 12(3):1100–1105.

Genest, C. (1984b). A conflict between two axioms for combining subjective distributions. *Journal of the Royal Statistical Society*, 46(3):403–405.

Genest, C. (1984c). Pooling operators with the marginalization property. *Canadian Journal of Statistics*, 12(2):153–163.

Genest, C., McConway, K. J., and Schervish, M. J. (1986). Characterization of externally Bayesian pooling operators. *Annals of Statistics*, 14(2):487–501.

Genest, C. and Wagner, C. G. (1987). Further evidence against independence preservation in expert judgement synthesis. *Aequationes Mathematicae*, 32(1):74–86.

Genest, C. and Zidek, J. V. (1986). Combining probability distributions: A critique and an annotated bibliography. *Statistical Science*, 1(1):114–148.

Gibbard, A. (1973). Manipulation of voting schemes. *Econometrica*, 41(4):587–602.

Gibbard, A. (1977). Manipulation of schemes that mix voting with chance. *Econometrica*, 45(3):665–681.

Grossman, S. J. (1981). An introduction to the theory of rational expectations under asymmetric information. *Review of Economic Studies*, 48(4):541–559.

Hanson, R. (1991). Even adversarial agents should appear to agree. In *IJCAI-91 Workshop on Reasoning in Adversarial Domains*.

Hanson, R. (1998a). Consensus by identifying extremists. *Theory and Decision*, 44(3):293–301.

Hanson, R. (1998b). Decision markets. *IEEE Intelligent Systems*, 14(3):16–19.

Hanson, R. D. (1995). Could gambling save science? Encouraging an honest consensus. *Social Epistemology*, 9(1):3–33.

Hazelrigg, G. A. (1996). The implications of Arrow's impossibility theorem on approaches to optimal engineering design. *Journal of Mechanical Design*, 118(2):161–164.

Heckerman, D. (1991). *Probabilistic Similarity Networks*. MIT Press, Cambridge, MA.

Horvitz, E. (1997). Models of continual computation. In *Fourteenth National Conference on Artificial Intelligence*, pages 286–293, Cambridge, MA.

Horvitz, E. (1999). Thinking ahead: Continual computation policies for allocating idle and real-time resources to solve future challenges. In *Sixteenth International Joint Conference on Artificial Intelligence*, pages 1280–1286, Stockholm, Sweden.

Huang, C.-F. and Litzenberger, R. H. (1988). *Foundations for Financial Economics*. Elsevier Science, Amsterdam.

Hylland, A. and Zeckhauser, R. (1979). The impossibility of Bayesian group decision making with separate aggregation of beliefs and values. *Econometrica*, 47(6):1321–1336.

Jacobs, R. A. (1995). Methods for combining experts' probability assessments. *Neural Computation*, 7(5):867–888.

Jensen, F. V. (1996). *An Introduction to Bayesian Networks*. Springer, New York.

Jensen, F. V., Lauritzen, S. L., and Olesen, K. G. (1990). Bayesian updating in causal probabilistic networks by local computations. *Computational Statistics Quarterly*, 4:269–282.

Kadane, J. B. and Winkler, R. L. (1988). Separating probability elicitation from utilities. *Journal of the American Statistical Association*, 83(402):357–363.

Kasanen, E. and Trigeorgis, L. (1994). Market utility approach to investment valuation. *European Journal of Operational Research*, 74(2):294–309.

Keeney, R. L. and Raiffa, H. (1976). *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. John Wiley and Sons, New York.

Kloks, T. (1994). *Treewidth: Computations and Approximations*. Springer-Verlag, Berlin.

Laddaga, R. (1977). Lehrer and the consensus proposal. *Synthese*, 36:473–477.

Lauritzen, S. L. and Spiegelhalter, D. J. (1988). Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society, Series B*, 50:157–224.

Lehrer, K. and Wagner, C. G. (1983). Probability amalgamation and the independence issue: A reply to Laddaga. *Synthese*, 55(3):339–346.

Lenat, D. B. and Guha, R. V. (1990). *Building Large Knowledge-Based Systems: Representation and Inference in the CYC Project*. Addison-Wesley, Reading, MA.

Levy, W. B. and Delic, H. (1994). Maximum entropy aggregation of individual opinions. *IEEE Transactions on Systems, Man, and Cybernetics*, 24(4):606–613.

Lindley, D. V. (1985). Reconciliation of discrete probability distributions. In *Bayesian Statistics 2*, pages 375–390, Amsterdam. North-Holland.

Lindley, D. V. (1988). The use of probability statements. In *Accelerated Life Testing and Experts' Opinions in Reliability*, pages 25–57, Amsterdam. Elsevier.

Lucas, R. E. (1972). Expectations and the neutrality of money. *Journal of Economic Theory*, 4(2):103–24.

MacKie-Mason, J. and Varian, H. (1995). Pricing congestible network resources. *IEEE Journal on Selected Areas in Communications*, 13(7):1141–1149.

Mas-Colell, A., Whinston, M. D., and Green, J. R. (1995). *Microeconomic Theory*. Oxford University Press, New York.

Matzkevich, I. and Abramson, B. (1992). The topological fusion of Bayes nets. In *Eighth Conference on Uncertainty in Artificial Intelligence*, pages 152–158.

Matzkevich, I. and Abramson, B. (1993). Some complexity considerations in the combination of belief networks. In *Ninth Conference on Uncertainty in Artificial Intelligence*, pages 152–158.

McKelvey, R. D. and Page, T. (1986). Common knowledge, consensus, and aggregate information. *Econometrica*, 54(1):109–127.

McKelvey, R. D. and Page, T. (1990). Public and private information: An experimental study of information pooling. *Econometrica*, 58(6):1321–1339.

Minsky, M. (1986). *The Society of Mind*. Simon and Schuster, NY.

Mongin, P. (1995). Consistent Bayesian aggregation. *Journal of Economic Theory*, 66(2):313–351.

Morris, P. A. (1974). Decision analysis expert use. *Management Science*, 20(9):1233–1241.

Morris, P. A. (1977). Combining expert judgments: A Bayesian approach. *Management Science*, 23(7):679–693.

Morris, P. A. (1983). An axiomatic approach to expert resolution. *Management Science*, 29(1):24–32.

Myerson, R. B. and Satterthwaite, M. A. (1983). Efficient mechanisms for bilateral trading. *Journal of Economic Theory*, 29:265–281.

Myung, I. J., Ramamoorti, S., and Bailey, A. D. (1996). Maximum entropy aggregation of expert predictions. *Management Science*, 42(10):1420–1436.

Nau, R. F. (1992). Joint coherence in games of incomplete information. *Management Science*, 38(3):374–387.

Nau, R. F. (1995). Coherent decision analysis with inseparable probabilities and utilities. *Journal of Risk and Uncertainty*, 10(1):71–91.

Nau, R. F. and McCardle, K. F. (1990). Coherent behavior in noncooperative games. *Journal of Economic Theory*, 50(2):424–444.

Nau, R. F. and McCardle, K. F. (1991). Arbitrage, rationality, and equilibrium. *Theory and Decision*, 31:199–240.

Neapolitan, R. E. (1990). *Probabilistic Reasoning in Expert Systems: Theory and Algorithms*. John Wiley and Sons, New York.

Ng, K.-C. and Abramson, B. (1992). Consensus diagnosis—A simulation study. *IEEE Transactions on Systems, Man, and Cybernetics*, 22(5):916–928.

Ng, K.-C. and Abramson, B. (1994). Probabilistic multi-knowledge-base systems. *Applied Intelligence*, 4(2):219–236.

Nielsen, L. T., Brandenburger, A., Geanakoplos, J., McKelvey, R., and Page, T. (1990). Common knowledge of an aggregate of expectations. *Econometrica*, 58(5):1235–1239.

Norvig, T. (1967). Consensus of subjective probabilities: A convergence theorem. *Annals of Mathematical Statistics*, 38:221–225.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann.

Pearl, J. (1993). From Bayesian networks to causal networks. In *Adaptive Computing and Information Processing 1*, pages 165–194.

Pennock, D. M. (1997). Dags versus join trees: An argument for probabilistic inference using directed methods. Unpublished Technical Report.

Pennock, D. M. (1998). Logarithmic time parallel Bayesian inference. In *Fourteenth Conference on Uncertainty in Artificial Intelligence*, pages 431–438, Madison, WI, USA.

Pennock, D. M. and Horvitz, E. (1999). Analysis of the axiomatic foundations of collaborative filtering. In *Workshop on AI for Electronic Commerce at the Sixteenth National Conference on Artificial Intelligence*.

Pennock, D. M. and Wellman, M. P. (1996). Toward a market model for Bayesian inference. In *Twelfth Conference on Uncertainty in Artificial Intelligence*, pages 405–413, Portland, OR, USA.

Pennock, D. M. and Wellman, M. P. (1997). Representing aggregate belief through the competitive equilibrium of a securities market. In *Thirteenth Conference on Uncertainty in Artificial Intelligence*, pages 392–400, Providence, RI, USA.

Pennock, D. M. and Wellman, M. P. (1999). Graphical representations of consensus belief. In *Fifteenth Conference on Uncertainty in Artificial Intelligence (UAI–99)*, pages 531–540.

Plott, C. R. and Sunder, S. (1982). Efficiency of experimental security markets with insider information: An application of rational-expectations models. *Journal of Political Economy*, 90(4):663–98.

Plott, C. R. and Sunder, S. (1988). Rational expectations and the aggregation of diverse information in laboratory security markets. *Econometrica*, 56(5):1085–1118.

Plott, C. R., Wit, J., and Yang, W. C. (1997). Parimutuel betting markets as information aggregation devices: experimental results. Technical Report Social Science Working Paper 986, California Institute of Technology.

Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 32:122–136.

Raiffa, H. (1968). *Decision Analysis: Introductory Lectures on Choices under Uncertainty.* Addison-Wesley, Reading, MA.

Raiffa, H. and Schlaifer, R. (1961). *Applied Statistical Decision Theory.* Harvard University, Boston.

Regev, O. and Nisan, N. (1998). The popcorn market—an online market for computational resources. In *First International Conference on Information and Computation Economies*, Charleston, SC.

Rosenblueth, E. and Ordaz, M. (1992). Combination of expert opinions. *Journal of Scientific and Industrial Research*, 51:572–580.

Rubinstein, M. (1974). An aggregation theorem for securities markets. *Journal of Financial Economics*, 1(3):225–244.

Rubinstein, M. (1975). Securities market efficiency in an Arrow-Debreu economy. *American Economic Review*, 65(5):812–824.

Rubinstein, M. (1976). The strong case for the generalized logarithmic utility model as the premier model of financial markets. *Journal of Finance*, 31(2):551–571.

Russell, S. J. and Norvig, P. (1995). *Artificial Intelligence: A Modern Approach*. Prentice-Hall, New Jersey.

Saari, D. G. (1995). Chaotic exploration of aggregation paradoxes. *SIAM Review*, 37(1):37–52.

Sandholm, T. and Vulkan, N. (1999). Bargaining with deadlines. In *Sixteenth National Conference on Artificial Intelligence*, pages 44–51.

Sandholm, T. W. and Lesser, V. R. (1996). Advantages of a leveled commitment contracting protocol. In *Thirteenth National Conference on Artificial Intelligence*, pages 126–133.

Satterthwaite, M. A. (1975). Strategy-proofness and Arrow's conditions: existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10:187–217.

Savage, L. J. (1954). *The Foundations of Statistics*. Wiley, New York.

Schervish, M. J., Seidenfeld, T., and Kadane, J. B. (1991). Shared preferences and state-dependent utilities. *Management Science*, 37(12):1575–1589.

Seidenfeld, T., Kadane, J. B., and Schervish, M. J. (1989). On the shared preferences of two Bayesian decision makers. *Journal of Philosophy*, 86(5):225–244.

Seidenfeld, T. and Schervish, M. J. (1991). Two perspectives on consensus for (Bayesian) inference and decisions. *IEEE Transactions on Systems, Man, and Cybernetics*, 20(1):318–325.

Sen, A. (1970). The impossibility of a Paretian liberal. *Journal of Political Economy*, 78(1):152–157.

Sen, A. (1986). Social choice theory. In Arrow, K. J. and Intriligator, M. D., editors, *Handbook of Mathematical Economics, Volume III*, pages 1073–1181. Elsevier.

Shachter, R. D. (1986). Evaluating influence diagrams. *Operations Research*, 34(6):871–882.

Shachter, R. D. (1988). Probabilistic inference and influence diagrams. *Operations Research*, 36(4):589–604.

Shachter, R. D. (1990). Evidence absorption and propagation through evidence reversals. In Henrion, M., Shachter, R. D., Kanal, L. N., and Lemmer, J. F., editors, *Uncertainty in Artificial Intelligence*, volume 5, pages 173–190. North Holland.

Shachter, R. D., Andersen, S. K., and Poh, K. L. (1991). Directed reduction algorithms and decomposable graphs. In *Uncertainty in Artificial Intelligence*, volume 6, pages 197–208. North Holland, Amsterdam.

Shafer, G. (1976). *A Mathematical Theory of Evidence.* Princeton University Press, Princeton, NJ.

Shoven, J. B. and Whalley, J. (1992). *Applying General Equilibrium.* Cambridge University Press.

Smets, P. and Kennes, R. (1994). The transferable belief model. *Artificial Intelligence*, 66:191–234.

Smith, J. (1990). Statistical principles on graphs. In Oliver, R. and Smith, J., editors, *Influence Diagrams, Belief Nets and Decision Analysis*, page 89=120. Wiley, Chchester.

Spiegelhalter, D. J., Dawid, A. P., Lauritzen, S. L., and Cowell, R. G. (1993). Bayesian analysis in expert systems. *Statistical Science*, 8(3):219–203.

Spohn, W. (1988). Ordinal conditional functions: A dynamic theory of epistemic states. *Causation in Decision, Belief Change, and Statistics*, 2:105–134.

Turek, J. and Shasha, D. (1992). The many faces of consensus in distributed systems. *Computer*, 25(6):8–17.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59:433–560.

Varian, H. R. (1985). Divergence of opinion in complete markets: A note. *Journal of Finance*, 40(1):309–317.

Varian, H. R. (1987). The arbitrage principle in financial economics. *Journal of Economic Perspectives*, 1(2):55–72.

Varian, H. R. (1989). Differences of opinion in financial markets. In Stone, C. C., editor, *Financial Risk: Theory, Evidence and Implications*, pages 3–37. Kluwer Academic, Norwell, MA.

von Neumann, J. and Morgenstern, O. (1953). *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, NJ.

Wagner, C. (1984). Aggregating subjective probabilities: Some limitative theorems. *Notre Dame Journal of Formal Logic*, 25(3):233–240.

Walsh, W. E. and Wellman, M. P. (1998). A market protocol for decentralized task allocation. In *Third International Conference on Multiagent Systems*, pages 325–332.

Walsh, W. E. and Wellman, M. P. (1999). Efficiency and equilibrium in task allocation economies with hierarchical dependencies. In *Sixteenth Iternational Joint Conference on Artificial Intelligence*, pages 520–526.

Wellman, M. P. (1993). A market-oriented programming environment and its application to distributed multicommodity flow problems. *Journal of Artificial Intelligence Research*, 1:1–22.

Wellman, M. P. and Walsh, W. E. (1999). Distributed quiescence detection in multiagent negotiation. In *Workshop on Negotiation at the Sixteenth National Conference on Artificial Intelligence*, Orlando, FL, USA.

West, M. (1984). Bayesian aggregation. *Journal of the Royal Statistical Society. Series A. General*, 147(4):600–607.

West, M. and Crosse, J. (1992). Modeling probabilistic agent opinion. *Journal of the Royal Statistical Society. Series B. Methodological*, 54(1):285–299.

Whittaker, J. (1990). *Graphical Models in Applied Multivariate Statistics*. Wiley, Chichester, England; New York.

Wilson, R. (1968). The theory of syndicates. *Econometrica*, 36(1):119–132.

Winkler, R. L. (1968). The consensus of subjective probability distributions. *Management Science*, 15(2):B61–B75.

Winkler, R. L. (1981). Combining probability distributions from dependent sources. *Management Science*, 27(4):479–488.

Winkler, R. L. (1986). Expert resolution. *Management Science*, 32(3):298–303. Introduces commentary by Lindley, Schervish, Clemen, French, and Morris.

Wurman, P., Walsh, W., and Wellman, M. (1998a). Flexible double auctions for electronic commerce: Theory and implementation. *Decision Support Systems*, 24:17–27.

Wurman, P. R., Wellman, M. P., and Walsh, W. E. (1998b). The Michigan Internet AuctionBot: A configurable auction server for human and software agents. In *Second International Conference on Autonomous Agents*, pages 301–308, Minneapolis.

Xiang, Y. (1996). A probabilistic framework for cooperative multi-agent distributed interpretation and optimization of communication. *Artificial Intelligence*, 87(1–2):295–342.

Yule, G. U. (1903). Notes on the theory of association of attributes in statistics. *Biometrika*, 2:121–134.

Zadeh, L. A. (1979). A theory of approximate reasoning. *Machine Intelligence*, 9:149–194.