ON THE COST OF BEING UNCERTAIN

PAUL BOES

(576106)

Bachelor's Thesis submitted for acquiring the degree of

Bachelor of Science (B.Sc.) in Economics
at the
School of Business and Economics,
Humboldt-Universität zu Berlin

First Examiner: Prof. Georg Weizsäcker, PhD
Second Examiner: Prof. Dr. Ulrich Kamecke

July 12, 2018

ABSTRACT

Being uncertain about relevant information produces costs for economic agents in several ways, both directly and indirectly. In this thesis, I study how the direct costs — the costs associated not to any outcome of a lottery but to the participation in a lottery in the first place — of being uncertain should, or can, be reasonably modeled.

I first argue that a naive micro-economic approach to quantifying such costs — using the theory of insurance — has a number of problems. I then propose to use the *Shannon entropy* from information theory as an alternative measure of direct uncertainty costs and I motivate this proposal both formally, by introducing axioms that single out the Shannon entropy, and conceptually, by arguing that this quantity naturally measures the expected uncertainty costs for economic tasks that require an lagent to react differently to different events.

I then present two applications of these results: First, I apply them to transaction cost economics, where I argue, both theoretically and by giving empirical support, that uncertainty cost is a quantity that co-determines the decision of a firm whether to integrate part of its supply chain or not. Secondly, I apply them to decision theory, where I present an alternative to the expected utility hypothesis as a decision criterion, and show that this alternative naturally solves Allais', Rabin's and the Ellsberg Paradox.

Along the way, I also clarify the (often blurred) distinction between (non-Knightian) uncertainty and risk and show that the mathematical concept of *majorization* naturally quantifies the two notions and makes their relationship precise.

CONTENTS

1 INTRODUCTION

There are many types of uncertainty. In an economics setting, a manager might be uncertain about the consumer demand for the good which her company produces; Or she might be uncertain about the quality of some production factor that is provided by one of her suppliers; Or she might not know whether her employees actually work as hard as they say; Or whether her business partner is bluffing. In one way or another, these various ways of being uncertain produce costs and it would be both interesting and valuable to be able to model those costs. However, it seems silly at best to think that there is a single formalism that covers all these various types of uncertainty costs, since they arise via different mechanisms.

This thesis focuses on one such type of uncertainty cost. In particular, I will be concerned with the costs that arise *directly* from being in a state of uncertainty, or, in other words, those costs that would be absent in a situation that is the same as the original situation except that there is no uncertainty.

What are such direct costs and how are they relevant for economics? To illustrate the idea, consider a principal-agent type exchange between two parties in which one party (the agent) has some information A relevant to the exchange that is unknown to the other (the principal). In order to decide her strategy, the principal therefore has to act on the basis of a probability distribution p over the possible values of A, where this distribution represents the principal's best guess concerning the actual value of the information. Now, the way in which p usually enters models of decision making is that the principal decides her strategy based on maximizing her expected utility under p (e.g. [Sch99; Var14]). Now, if this was the only way in which p enters a model, then this would imply that all distributions that induce the same expected utility are equivalent from the point of view the principal and hence that she behaves in the same way when confronted with any of these distributions.

But this is an unrealistic implication! In the real world, different probability distributions come with different associated costs. Consider, for example, the case in which the principal is a manager and the agent a supplier of goods to the company run by the principal. Then A could be information about the type of the agent that determines the quality of the goods supplied. In this setting, it is easy to envision a scenario in which A can take three values - "lazy", "normal", "eager" - and in which the utility function of the principal is such that the expected utility of the supplier being normal with probability 1 equals the expected utility of the supplier being lazy and eager with probability 0.5 each. However, in the latter case the principal will have to act differently depending on whether the supplier turns out to be lazy or eager in order to *realize* her utility for each of the two respective types, and hence will on average waste some resources for preparatory measures towards an event that never occurs (she might for example insure herself against the supplier being lazy when in fact he is eager). In contrast, in the former case, there are no such costs, because the manager already has full information about the supplier's type. The resources that a rational economic agent will spend when faced with uncertain situations like the above are exactly what I call the *direct costs of uncertainty* and the aim of this thesis is to develop a formal and conceptual account for these costs.

The structure of the thesis is as follows: I will start off by presenting the way in which I believe standard micro-economic theory would go about modeling the above uncertainty costs and discuss its problems (Sec. 2). Part of this section will be a discussion about the relationship between the notions of risk and uncertainty, where I show that these notions are conceptually distinct but mathematically related via the notion of *majorization*. I will then present an alternative account that does not have those problems (Sec. 3). In particular, I will introduce axioms that single out the *Shannon entropy* [Sha48] as the only function that adequately quantifies the direct costs of uncertainty. I also discuss how to understand, in an economics context, why this function is particularly suitable to measure uncertainty cost. Then, I apply these results to two different fields of economics: First, I will apply them to the field of transaction cost economics, and argue, both theoretically and by giving empirical support, that uncertainty cost is a quantity that co-determines the decision of a firm whether to integrate part of its supply chain or not (Sec. 4). Secondly, I will apply the results of Section 3 to decision theory, where I use them to present an alternative to the expected utility hypothesis and show that this alternative naturally solves Allais', Rabin's and the Ellsberg Paradox (Sec. 5). I close with a summary and outlook (Sec. 6).

1.1 SETUP AND NOTATION

Before I start, let me first introduce the formalism that will be used throughout the remainder of the thesis. All that I want to say can be phrased in terms of lotteries.[1] Lotteries are an important tool in game, contract, and decision

1 For notational and conceptual convenience, I will keep the discussion to discrete random variables, however modulo some technical subtleties everything can be extended to continuous random variables.

theory, where they are used to model the behavior of agents under uncertainty or risk [Var14]. Formally, a lottery l is a finite and ordered set of outcomes $A = (a_1, \dots, a_n)$ together with associated probabilities of occurrence p_i such that $p_i \geqslant 0$ for all i and $\sum_i p_i = 1$. In other words, it corresponds to a random variable X_l such that $\mathrm{Prob}(X = a_i) = p_i$. For convenience, I will assume, without loss of generality, that $A \subset \mathbb{R}^n$, so that the outcomes a_i can be interpreted as payouts. Moreover, again w.l.o.g., I will assume that $a_1 \geqslant a_2 \geqslant \cdots \geqslant a_n$. Then, $L_n = \Delta_n \times \mathbb{R}^{n,\downarrow}$ is the set of lotteries with n outcomes. Here, Δ_n is the n-dimensional simplex, which describes the set of n-dimensional probability vectors $p^T = (p_1, \dots, p_n)$ and $\mathbb{R}^{n,\downarrow}$ is the space of non-increasingly ordered elements of $\mathbb{R}^n$. Hence, one can represent lotteries as tuples $l \equiv (p, A)$. Note that L_n is convex with respect to the natural addition operation inherited from its component vector spaces, so that, for any $\lambda \in [0,1]$ and $l, m \in L_n$, $\lambda l + (1-\lambda)m = (\lambda p_l + (1-\lambda)p_m, \lambda A_l + (1-\lambda)A_m) \in L_n$. However, L_n is not closed under more general affine sums, since neither $\mathbb{R}^{n,\downarrow}$ nor Δ_n are. Denote $\mathcal{L} = \bigcup_{n\in\mathbb{N}} L_n$ and $\Delta = \bigcup_{n\in\mathbb{N}} \Delta_n$. Finally, $\mathbb{E}(l) = \sum_i p_i a_i$ is the expectation value of a lottery and all logarithms are base 2.

Throughout the thesis, I will also discuss different orderings. An ordering over a set S is a relation $(\succ, S) \subseteq S \times S$, usually denotes by $(\succ, S)$ or simply $\succ$. For any such ordering and elements $l, m, k \in S$, $l \sim m$ denotes $l \succ m \wedge m \succ l$. An ordering is called a *pre-order* if it is *reflexive* ($l \succ l$, $\forall l$), and *transitive* ($l \succ k \wedge k \succ m \Rightarrow l \succ m$); It is called a *partial ordering* if it is a pre-order and *antisymmetric* ($l \sim m \Rightarrow l = m$); It is called a *total ordering* if it is a partial ordering and *complete* ($l \succ m \vee m \succ l$). Finally, the notion of an *order-preserving* function will be important. A function $g : \mathcal{L} \to \mathbb{R}$ is order-preserving for an ordering $\succ$ if $l \succ m \Rightarrow g(l) \geqslant g(m)$. If $-g$ is order-preserving, then g is called *order-reversing*.

2

HOW TO MODEL UNCERTAINTY COST — THE WRONG WAY

We want to know how we should quantify the direct costs of uncertainty. In terms of the above notation, this question can be phrased as the question: How should we model the direct costs of uncertainty for a decision maker faced with a given lottery l. That is, does there exist a function $f : \mathcal{L} \to \mathbb{R}$, or potentially a whole family of such functions, that adequately models these costs?

2.1 DEFINING UNCERTAINTY

To answer this question, I should first clarify what I mean by uncertainty. The direct uncertainty costs that I want to quantify are those that arise for an economic agent due to the fact that her knowledge about future events is limited. It therefore seems natural to say that one random variable (or lottery) is more uncertain than another if its outcomes are more difficult to predict than those of the other. To make this precise, let λ be a measure defined over the set $\mathcal{B}$ of Borel subsets of $\mathbb{R}^n$. For a random variable X taking values in $\mathbb{R}^n$, let

$$\mathcal{A}(X;p) = \{A \in \mathcal{B} : \mathrm{Prob}(X \in A) \geqslant p\}$$

and

$$\mu(X,p) = \inf_{A \in \mathcal{A}(X;p)} \lambda(A).$$

Whilst scary to read, these definitions are actually very simple: The Borel subsets are simply all subsets of $\mathbb{R}^n$ with a well-defined volume under λ (think the length of any line segment on $\mathbb{R}$ with respect to the uniform measure). The set $\mathcal{A}(X;p)$ is then the set of of all regions of $\mathbb{R}^n$ for which one can say that the outcome of X lies in that region with "confidence" probability p. So what $\mu(X,p)$ measures is the smallest region in $\mathbb{R}^n$ of which I can predict with confidence p that X will lie in it.

Definition 1 (Uncertainty) *A random variable* X *is* more uncertain *than another* Y*, written* $X \geqslant_U Y$*, if* $\mu(X,p) \geqslant \mu(Y,p)$ *for all* $p \in (0,1)$.

Clearly, the ordering $\geqslant_U$ induces an ordering on $\mathcal{L}$, defined as $l \succ_U m \Leftrightarrow X_l \leqslant_U X_m$. And it seems the minimum to require for any reasonable uncertainty cost function f that it respects this ordering, in the sense that $l \succ_U m \Rightarrow f(l) \leqslant f(m)$ (see Sec. 3.1).

2.2 UNCERTAINTY COST AS RISK PREMIUM

Having clarified the notion of uncertainty that we will be concerned with, let me first consider what a micro-economist would probably take to be the correct answer. I think that he would say that $f(l)$ is best modeled by the *risk premium of a full and actuarially fair insurance that a risk averse player* P *faced with the lottery* l *would choose*.

Let me unpack this statement. To begin with, the micro-economist would assume that P has a *(Bernoulli) utility function* $u : \mathbb{R} \to \mathbb{R}$. Next, he would moreover assume the expected utility hypothesis:

EXPECTED UTILITY HYPOTHESIS (EUH) Given a choice between lotteries, a decision maker will prefer the lottery that maximizes his expected utility under u.

Formally, the EUH says that there exists an ordering $(\succ_u, \mathcal{L})$ that represents P's preference between lotteries and that this ordering is given by

$$l \succ_u m \Leftrightarrow Eu(l) \geqslant Eu(m) \equiv \sum_{i=1} p_i^{(l)} u(a_i^{(l)}) \geqslant \sum_{i=1} p_i^{(m)} u(a_i^{(m)}),$$

where I have implicitly defined the function Eu as the expected utility of a lottery.

Now, note first that, given u, and assuming that the latter is strictly monotonically increasing — a reasonable assumption —, for any lottery l, there exists a unique number RP such that

$$u(\mathbb{E}(l) - RP) = Eu(l) \tag{1}$$

and we can find this number, since from the strict monotonicity of u it follows that the inverse u^{-1} is well defined. But the LHS of this equation can be interpreted as the expected utility of a trivial and unique[1] lottery $\hat{l} = (p^{(\hat{l})} = (1, 0, \ldots, 0), A^{(\hat{l})} = (\mathbb{E}(l) - RP, 0 \ldots, 0))$ with payout $\mathbb{E}(l) - RP$. By construction, $Eu(\hat{l}) = Eu(l)$. Hence, given EUH, it follows that $\hat{l} \sim_u l$, or in words, that P is indifferent between the two lotteries. But this means that

1 This is the first time the restriction to $\mathbb{R}^{n,\downarrow}$ becomes apparent, since otherwise $\hat{l}$ would be specified only up to permutation. This theme runs through the thesis: A pre-order on vector spaces will be turned into partial order by considering only the quotient space with respect to permutations. I will not usually discuss technical aspects such as this one in the main text.

P would be willing to pay RP to swap the lottery l that involves uncertainty against the lottery $\hat{l}$ that involves no uncertainty but whose certain payout is $\mathbb{E}(l) - RP$. Now, RP is called the *risk premium* and it is textbook theory that a fair insurance (one without expected returns) will offer such a swap at the price RP and that a P whose preferences follow the EUH will accept this offer [Var14]. Hence, since the only effect of this swap for P (remember, his expected utility remains unchanged) is to remove the uncertainty from the lottery, it seems reasonable to claim that $f(l) = RP(l)$, that is, that the cost of uncertainty should be modeled by the risk premium.

The final part of the unpacking of the above statement is to understand why, according to the micro-economist, our P needs to be risk averse. The reasoning is this: So far, we haven't ensured that RP is non-negative. But surely that should be the case, if it is to model the cost of uncertainty. Nobody should be able to make money off of (their own) uncertainty. Hence, we require that $RP(l) \geqslant 0$ for all lotteries l. But this yields

$$\mathbb{E}(l) \geqslant \hat{a} \Leftrightarrow u(\mathbb{E}(l)) \geqslant u(\hat{a}) = Eu(\hat{l}) = Eu(l) = \mathbb{E}(u(l)), \tag{2}$$

where we again used the strict monotonicity of u which implies $x \geqslant y \Leftrightarrow u(x) \geqslant u(y)$ with equality only for $x = y$. But it is the content of the very useful Jensen's inequality that the right part of (2), $u(\mathbb{E}(l)) \geqslant \mathbb{E}(u(l))$ is true if and only if u is concave. Moreover, having a concave Bernoulli utility function is the *definition* of being risk averse. Hence, to ensure positivity of uncertainty costs, P should be risk averse, and we now understand every word in the economist's claim.

2.3 PROBLEMS

This story goes down well and sounds very reasonable. Unfortunately, I think that it is highly unsatisfactory, for at least three reasons that I will discuss in this subsection.

Firstly, it seems to me that the risk premium approach conceptually overloads the utility function by requiring it to capture both the degree of risk aversion as well as the uncertainty cost that arises as a result of a given lottery. As an indication of this problem it suffices to note that surely a hedge fund, packed with risk affine guys in shirt sleeves, suffers from costs due to not being able to predict the future just like everybody else does. By modeling uncertainty costs via concave Bernoulli utility functions, one seems to remove the possibility of modeling uncertainty costs for risk affine players.

Secondly, for the risk premium to be more than just an academic measure of uncertainty cost, it would have to be the case that either all economic agents constantly insure themselves in every lottery they face, or there is reason to think that the costs of an insurance for a lottery are also the costs for a player in case she does not insure herself. But neither of these two is true. The costs of insuring oneself against every possible lottery are prohibitive, so only very few lotteries will actually be covered by insurances. Indeed, I think it's clear that the vast majority of realistic lotteries could not be covered by insurances at all, simply because there is no appealing market for insurances: Who would insure you against the manager of the other company being a choleric type, or against Trump's tweets? Moreover, there is no reason at all in the above story for why the risk premium should also be a good measure of the uncertainty cost in situations where no insurance applies. So at best the above risk premium approach could only apply to a very small subset of economic interactions.

Thirdly, and most importantly, the above approach mixes up the notions of risk and uncertainty. Luckily, the relationship between risk and uncertainty can be made conceptually and formally very clear, so that we can understand very clearly how and why the risk premium breaks down. However, to do so requires a bit of preparation.

2.3.1 *Uncertainty is not risk*

Given two lotteries, which one is riskier?[2] This question was answered to general satisfaction in a celebrated series of papers by Michael Rothschild and (the then young) Joseph Stiglitz in 1970 and 1971 [RS70; RS71]. They showed that three intuitive ways of comparing the riskiness of lotteries with the same expectation value boil down to one and the same thing. Let me first introduce them. Importantly, they all apply only to lotteries with the same expectation value. In terms of lotteries, the first way says that m should be considered riskier than l if m is "l plus noise". Formally, let X_l be the

2 In economics, the distinction between risk and uncertainty has a long tradition that goes back to [Kni21]. In this seminal book, Knight argued that much of economics is driven by inherently unquantifiable, at best estimate-driven knowledge. This type of knowledge is what Knight called "uncertainty" and juxtaposed it to the notion of "risk" which concerned quantifiable, probabilistic knowledge. Historically, Knight's distinction stands at the beginning of a splitting into schools of thought about how Knightian uncertainty should be dealt with in economics, a splitting that was roughly co-aligned with the Neoclassical and Keynesian schools [Köh17, Ch.2]. However, Knight's distinction is *not* what I'm concerned with here! This is clear both from the definition of uncertainty that I give as well as from the fact that I assume there always exist well-defined probabilities. Yet, the work on subjective probabilities that was given a rigorous formal footing in the work of Kolmogorov, de Finetti and others is usually associated to Knightian uncertainty. For the connection between such probabilities and my approach here, see footnote 3 in section 5.2.

random variable corresponding to some lottery l. Then we write $l \succ_N m$ if there exists a random variable Y with $\mathbb{E}(Y|X_l) = 0$ such that

$$X_m \underset{d}{=} X_l + Y,$$

where "$\underset{d}{=}$" means "has the same distribution as". The second way says that m should be considered riskier than l if every risk averse player prefers l to m. Thus, formally, we write $l \succ_R m$ if, for every concave function u,

$$Eu(l) \geqslant Eu(m).$$

The third way says that m should be considered riskier than l if m has "more weight in the tails" than l. To formalize this, Rothschild and Stiglitz use the notion of a *mean preserving spread* (MPS). While cumbersome to discuss formally, an MPS is simply any transformation of a lottery that preserves its mean but spreads across outcomes. This spreading necessarily adds weight to the tails of a distribution. Without making this formal, we write $l \succ_T m$ if there exists an MPS that takes l to m. Rothschild and Stiglitz then showed the following[3]:

Theorem 2 ([RS70]) *For all $l, m \in \mathcal{L}$, the following are equivalent:*

1. $l \succ_N m$
2. $l \succ_R m$
3. $l \succ_T m$

This theorem shows that all these three orderings coincide, giving very strong support to the claim that this ordering adequately captures the notion of riskiness. I will denote this ordering over lotteries as $\succ_{RS}$.

It is fair to say, then, that we have a pretty good idea what we mean when we say one lottery is riskier than another. Moreover, we can now also see that this ordering underlies, or at least goes hand in hand with, the notion of uncertainty in the risk premium approach to uncertainty costs that I presented above. This is because,

$$l \succ_{RS} m \Rightarrow RP(l) \leqslant RP(m), \tag{3}$$

or in other words, RP is an order-reversing function on $\succ_{RS}$. Using yet more words, it means that riskier lotteries according to the Rothschild-Stiglitz measure also produce higher risk premiums. To see why (3) is true, note that if $l \succ_{RS} m$, then by definition they share the same mean and by Thm. 2 every

3 The actual theorem was proven for continuous random variables, so this is really a special case of the general theorem.

risk averter prefers l to m. Combining this with (1) then yields the RHS of (3).

I can now formulate my precise criticism of the risk premium approach to uncertainty costs: $\succ_{RS}$ and $\succ_{U}$ do not coincide. In other words, there exist pairs of lotteries where the one with less uncertainty has higher associated risk premium. An example is given by

$$l = ((0.5, 0.5), (200, -200)), \quad m = ((0.2, 0.8), (400, -100)). \tag{4}$$

One can check that $l \succ_{RS} m$ while $m \succ_{U} l$. By (3), this implies that $\mathrm{RP}(l) < \mathrm{RP}(m)$, which contradicts our requirement that $f(m) \leqslant f(l)$ for any reasonable measure of uncertainty cost. Thus, the risk premium cannot be a valid measure of uncertainty cost.

2.3.2 *Majorization connects risk and uncertainty*

Before finally moving on to my own proposal, let me dwell a bit on the relationship between riskiness and uncertainty. This is because, from the above, it may seem as if riskiness and uncertainty are, ultimately, completely unrelated. This is not so, and there is a very elegant way of showing this. To do so, I will require the notion of *majorization*.

Definition 3 (Majorization) *Let v, w be two elements of a n-dimensional vector space $\mathcal{V}$ and denote as $v^{\downarrow}$ the permutation of the vector v whose entries are ordered non-increasingly, i.e. $v^{\downarrow}_1 \geqslant \cdots \geqslant v^{\downarrow}_n$, and similarly for w. We say that v* majorizes *w, and write $v \succ_M w$, if the following holds:*

$$\sum_{i=1}^{k} v^{\downarrow}_i \geqslant \sum_{i=1}^{k} w^{\downarrow}_i, \quad k \in \{1, \dots, n-1\}$$
$$\sum_{i=1}^{n} v^{\downarrow}_i = \sum_{i=1}^{n} w^{\downarrow}_i.$$

The point of this subsection is to show that majorization underlies both riskiness and uncertainty. While this may not be apparent from the definition, majorization naturally measures the spread of weight in a vector. It is a pre-order on $\mathcal{V}$ and a partial order on the ordered subspace of $\mathcal{V}$. It has found numerous applications in fields such as graph theory or matrix analysis [MOA11], but in fact has been "invented" in economics, where it was and is most often used to measure disparity and income inequality (e.g. [Lor05; Dal20; Atk70], see also [MOA11, Ch.13-F]).

What does majorization have to do with riskiness? Let

$$\tilde{l} = p \circ A = (p_1 a_1, p_2 a_2, \ldots, p_n a_n)$$

be the vector of the entry-wise (Hadamard) product of p and A. Then we have[4]

$$l \succ_{RS} m \Leftrightarrow \tilde{l} \succ_M \tilde{m}. \tag{5}$$

That is, the Rothschild-Stiglitz order of riskiness is nothing but majorization in disguise! Indeed, all of the formal results of Rothschild and Stiglitz in their [RS70] were at the time already known in the mathematics literature.[5]

Now, what does majorization have to do with uncertainty? This question is answered by the following proposition, proven in [MOA11, Prop. 17.E.11].

Proposition 1 $l \succ_U m \Leftrightarrow p^{(l)} \succ_M p^{(m)}$.

Thus, using majorization, we arrive at a conceptually and formally very clean picture of what the relationship between riskiness and uncertainty is: Riskiness is a majorization relation concerned with the spread of the *whole* lottery (by (5)), while uncertainty is a majorization relation concerned only with the spread of the *probability distribution* that forms part of the lottery (by the above proposition).

But this also implies that we can use the same mathematical tools to study these two notions. For instance, if there was a single function $g : \mathcal{V} \to \mathbb{R}$ such that $g(v) \leqslant g(w) \Leftrightarrow v \succ_M w$, then g would be the only function that could be used to quantify either riskiness or uncertainty. Unfortunately, there is no such function. There are, however, many order-reversing functions on $(\succ_M, \Delta)$. Such functions are called *Schur-concave* and the above discussion implies that every reasonable function of either riskiness or uncertainty should be Schur-concave. Which of these many functions should one use to measure uncertainty? In the next section, I will axiomatically pick out one Schur-concave function, the Shannon entropy, as a natural candidate.

---

4 Technically, for (5) to be true we require a slightly more general definition of majorization: The Rothschild-Stiglitz ordering did not make any assumption about the number of outcomes, so $\succ_{RS}$ relates lotteries all over $\mathcal{L}$. In contrast, $\succ_M$, as I defined it, only relates vectors in L_n. However, we could easily define majorization for vectors of different dimension by padding the smaller lottery with zeros in both p and A and in this case (5) holds.

5 $l \succ_R m \Leftrightarrow \tilde{l} \succ_M \tilde{m}$ was proven in [HLP29], $l \succ_N m \Leftrightarrow \tilde{l} \succ_M \tilde{m}$ in [Str65] and $l \succ_T m \Leftrightarrow \tilde{l} \succ_M \tilde{m}$ in [Mui03] and [HLP39].
In the literature, I have not come across any place where this connection between the work of Rothschild and Stiglitz and the relevant mathematics literature is made explicit. For example, already [Atk70] uses the connection without, however, referencing any of the relevant mathematical literature.

3 HOW TO MODEL UNCERTAINTY COST — A BETTER WAY

In this section, I argue that a function known as the *Shannon entropy* should be used to quantify direct uncertainty costs.

3.1 AXIOMS FOR UNCERTAINTY COST

At the beginning of the previous section, I have defined the notion of uncertainty whose associated costs I want to quantify. I have argued that every candidate uncertainty cost function f should reflect the corresponding "uncertainty" ordering $\succ_U$ over lotteries (and that RP does not do that). My first axiom says just this.

AXIOM 1 f should be order-reversing under $\succ_U$.

Axiom 1 says that more uncertain lotteries in the sense of Definition 1 should produce higher uncertainty cost — a reasonable requirement. By Proposition 1, this axiom implies that f should be a Schur-concave function of the probability distribution p associated to a given lottery only. Why so? Assume f was order-reversing and dependent on the outcomes A. Then there would have to exist two lotteries $l = (p, A)$ and $m = (p, B)$ with $A \neq B$ such that $f(l) < f(m)$. Now it is clear from the definition of an order-preserving or order-reversing function f on some ordering $\succ$ that, if $l \sim m$ with respect to this ordering, then also $f(l) = f(m)$. But in the above example, clearly $l \sim_U m$, so we require $f(l) = f(m)$, which yields a contradiction.

Now that we know that f depends only on the probability distribution of a lottery, we can ask what properties it should have. I propose the following ones:

AXIOM 2 (CONTINUITY) f should be continuous.

The idea behind this axiom is that changing the probability distribution only a little bit shouldn't lead to abrupt changes in the uncertainty cost.

To state the next axiom, denote as $\mathbf{1}_n = (1/n, \ldots, 1/n)$ the uniform probability vector in Δ_n.

AXIOM 3 (MAXIMALITY) $f(p \equiv \mathbf{1}_n)$ should be an increasing function in n.

This axiom reflects the fact that increasing the number of possible outcomes with equal probability should also increase the cost due to uncertainty.

The last axiom requires some setting up. Imagine a manager M with two suppliers S_1 and S_2. The manager faces uncertainty with respect to both these suppliers, however her uncertainty about them is correlated, meaning that once she learns something about S_1 she will in fact learn something about S_2. How should f model such a situation? Note first that it can be represented in terms of nested lotteries. Let l_1 denote the lottery that M faces with respect to S_1 and $l_{2|a_i}$ (with outcomes $b_{j|i}$ and respective probabilities $p_j^{(l_{2|a_i})}$) be the lottery that M faces with respect to S_2 once she has learned that the outcome of l_1 is a_i. Of course, we can describe this situation as one big "effective" lottery l with outcomes $a_i \cdot b_{j|i}$ and respective probabilities $p_1^{(l_1)} \cdot p_j^{(l_{2|a_i})}$. Now, a function $f : \mathcal{L} \to \mathbb{R}$ I call *additive with respect to nestings* if, for all such nested lotteries,

$$f(l) = f(l_1) + \sum_i p_i^{(l_1)} f(l_{2|a_i}). \tag{6}$$

AXIOM 4 (NESTING ADDITIVITY) f should be additive with respect to nestings.

While this axiom might seem at first glance somewhat ad hoc, I think that it actually ensures that f behaves just like we would expect uncertainty cost to behave when it comes to the combination of lotteries. For instance, if M learns nothing new about S_2 from learning about S_1, this would imply that $l_{2|a_i} \equiv l_2$ is the same for all outcomes a_i. In this case (6) implies that the total uncertainty cost should simply be the sum of the costs associated to each of the two lotteries l_1 and l_2.

I take it that all four of the above axioms are reasonable requirements on a function to adequately measure the cost of uncertainty. We can now use the following classic result by Shannon:[1]

Theorem 4 ([Sha48]) *Every $f : \Delta \to \mathbb{R}$ that satisfies Axioms 2,3 and 4 has the form*

$$f(p) = cH(p),$$

where c is a positive constant and

$$H(p) = -\sum_i p_i \log(p_i), \tag{7}$$

1 There exist characterizations of the Shannon entropy using other axiom sets, with differing degrees of intuitiveness [Csi08].

is the Shannon entropy.

Using this result and including Axiom 1, we obtain the following corollary:

Corollary 5 *Every* $f : \mathcal{L} \to \mathbb{R}$ *that satisfies Axioms 1 to 4 has the form*

$$f(l) = cH(p^{(l)}),$$

where c *is a positive constant.*

This result says that we should use a function called the Shannon entropy to measure uncertainty cost. Why? Because it's the *only* function that has the properties that we want any reasonable measure of uncertainty costs to have. But this might leave you with a bad aftertaste. Usually, we choose a function to represent something because we have an intuition for what that function does, and there is a conceptual link between what that function does and what we want to represent. Here, in contrast, the Shannon entropy just fell from the sky. But it is possible to give an alternative, conceptual motivation for why the Shannon entropy is a good measure of direct uncertainty costs? I will now present such a link.

3.2 WHAT THE SHANNON ENTROPY MEASURES

To understand, conceptually, how the Shannon entropy measures uncertainty cost, it makes sense to understand what the Shannon entropy measures in its original field of application, the theory of information.[2] What is it taken to measure there? If you look online, you will find many answers to this question, describing it as a measure of surprise or information. Here is a more precise statement about what it measures: Given a language with an alphabet A that consists of n letters a_i, let p_i be the probability that the letter a_i appears at some point in a some text.[3] Then $H(p)N$ is the shortest average length of a bit string (a string of only zeros and ones) that is required to encode messages of length N in that language so that a receiver could decode that message with probability close to one. To make sense of this, consider the following example: Let $A = \{a, c, g, t\}$ be a language consisting of four letters.[4] This language produces beautiful, poetic words like "$acgga$" or "$ccccc$". Now, if I wanted to send you a message in that language encoded in binary, I could choose the following encoding

2 In the following discussion, I assume that $c = 1$.

3 I assume this probability to be well-defined, for instance we could set p_i to equal the relative frequency of occurrence of a_i in all texts of this language known to me.

4 Indeed, using the above alphabet, we could measure the Shannon entropy of the genome for different kinds of species (recall the four nucleid acids). Fun fact: The Shannon entropy of the genome of the bacterium *Clostridium botulinum*, the source of botox, is about 1.85 [Kas18].

$$\begin{aligned} a &\to 00 & c &\to 01 \\ g &\to 10 & t &\to 11, \end{aligned}$$

in which every message of length N would have length 2N, since I use two bits to represent every letter. In fact, if in the language all four letters appear on average with the same frequency, then it is intuitively clear that this would also be the best encoding I could find. But what if the "a" and the "c" appear almost never in that language, say with probability $p_a = p_c = 2^{-5}$ (maybe one of them denotes questions, and asking questions is frowned upon down where they speak that language). Then the optimal encoding would instead be

$$\begin{aligned} a &\to 00000 & c &\to 11111 \\ g &\to 0 & t &\to 1. \end{aligned}$$

Following this rule, the average length of binary messages that I send to you will be much smaller than in the uniformly distributed case. For example, every N-letter word containing only "g" and "t" can be encoded in a bit-string of length N, so half as long as with the first encoding. And still, you could decode my messages correctly with probability close to one (not exactly one, because whenever a 00000 pops up, you have to guess whether I mean "a" or "ggggg"). But this is exactly what the Shannon entropy tells us as well: In the first case, where all four letters appear with the same probability, the Shannon entropy is $H((1/4, 1/4, 1/4, 1/4)) = 4(1/4 \log(1/4)) = 2$, while in the second case $H(\approx 1/2, \approx 1/2, \approx 0, \approx 0) \approx 2(1/2 \log(1/2)) = 1$.

Another way of seeing the above is the following. From (7), we can write the Shannon entropy as

$$H(p) = \mathbb{E}(X), \tag{8}$$

where X is a random variable with distribution $\mathrm{Prob}(X = x_i) = p_i$ and that takes values $x_i = -c \log p_i$. Now since all p_i lie between 0 and 1, smaller p_i produces a larger x_i. This sounds like in the above examples, where rare letters produced long messages. And indeed, x_i is exactly the number of bits that symbol a_i "costs" in the optimal encoding (like in the above example, where $p_a = 2^{-5}$ resulted in an encoding that uses 5 bits)!

But why am I saying all of this? What does this have to do with economics? The following: Zooming out, we learned that the Shannon entropy quantifies the average cost, in some relevant resource ("bits"), of solving a given task ("transmit message successfully in binary encoding"), under optimal allocation of resources in light of the probabilities with which events ("what is the message") in that task occur.[5] Moreover, there might be other costs to solving this task (for instance, one may have to pay the telephone provider a monthly rent), but what the Shannon entropy quantifies is *exactly the cost that is due to one's being uncertain, in the sense of Def. 1, about the event beforehand*. Further, it is clear that the Shannon entropy is only relevant for tasks in which the optimal solution actually depends on which event occurred. Call such tasks *event-dependent*. Now, I think that, whenever an economic problem can be understood as an event-dependent task, then an agent's optimal strategy to solving the task is very analogous to the optimal strategy in the message-sending game above, in the sense that this agent will allocate resources in such a way that she will require the least resources for the most likely outcomes, etc. This analogy between the tasks, then, is exactly the conceptual link between uncertainty cost and Shannon entropy that we sought earlier.

3.3 THE ENTROPIC UNCERTAINTY COST HYPOTHESIS

We have arrived at the idea that the Shannon entropy measures direct uncertainty cost in two ways: Firstly, by introducing axioms that show that the Shannon entropy is the only reasonable function to quantify the direct uncertainty costs that appear in lotteries; Secondly, by showing an analogy between event-dependent economic tasks and the message-sending task that gave a clear operational meaning to the Shannon entropy. Motivated by these insights, and under the assumption that economic agents solve event-dependent tasks under optimal allocation of resources, let me then formulate the above idea as a hypothesis for future reference:

ENTROPIC UNCERTAINTY COST HYPOTHESIS (EUCH) The Shannon entropy measures the direct uncertainty costs that accompany the solving of event-dependent economic tasks.

Assuming EUCH is true, what I have to do to motivate the use of the Shannon entropy in the context of a given economic problem is simply to show that this problem can be understood as an event-dependent task. I will do so in the applications below, but let me first compare this approach to the insurance approach from the last section.

5 Note that this way of understanding the Shannon entropy does not rely on the previous axioms at all. The connection between the two is simply that the optimal solution to the message-sending task *always* leads to an encoding that satisfies the axioms!

3.4 comparison with the insurance approach

At the end of the last section, where I discussed the insurance approach, I raised a number of problems that I had with that approach. Let me briefly explain why the current approach does not suffer from any of these problems: My first problem had to do with the fact that the insurance approach used the Bernoulli utility function to encode the costs due to uncertainty and that this was conceptually fishy because, for instance, it meant that risk affine players could not have uncertainty costs. The above approach does not suffer from this because the uncertainty cost is completely independent of the utility function. My second problem was that I considered the insurance approach insufficient because it applies only to the very small set of lotteries in which people would actually use an insurance. I think that the above considerations — the conceptual link between Shannon entropy and uncertainty cost — makes it clear that my alternative proposal is an adequate, or at least promising, way for modeling all those lotteries in which no insurance mechanism is used. My final problem was that the insurance approach equivocates uncertainty and risk. Well, I have excluded this possibility axiomatically in Axiom 1.

Now, that we have gained some understanding and intuition for the entropic approach, lets put it to use. This is what I will do in the remaining two chapters of this thesis: First, I apply my model of uncertainty costs to the "boundary of the firm"-problem from transaction cost economics. Then, I will suggest an alternative to the EUH and show that this alternative naturally solves at least three "paradoxes" that the EUH approach faces.

4 APPLICATION: BOUNDARY OF THE FIRM

In this section, I want to apply the model of uncertainty cost that I have introduced in the previous section to the field of transaction cost economics. In particular, I will address the "boundary of the firm"-problem that is one of the questions in the field. I will argue that uncertainty cost affects the agents' decisions in this problem and present empirical evidence for this prediction.

4.1 TRANSACTION COSTS AND THE STRUCTURE OF INSTITUTIONS

In 1937, Ronald Coase innocently asked "Why are there firms?". In his famous article, Coase argued that the prevalent theory of free market economy could not explain the existence of firms, because according to this theory all commercial activity should go through the market. This is because the latter's pareto-optimizing mechanism meant that whenever an entrepreneur would be in need of hiring somebody for a job, her most rational choice would be to find someone suitable on the labor market, contract them for the job and once both parties completed their part of the deal, the two would part and go their own ways. But as Coase rightly noticed, this is not what happens. Instead, many entrepreneurs choose to establish firms, where people get long-term contracts and offices, with Christmas parties, and so on. Clearly, the prevalent theory was missing something. This something Coase identified as the *transaction costs* that accompany the above process of finding a suitable someone on the labor market.

Transaction costs can be many things, such as the costs of setting up a contract (e.g. [Wil85]), the costs of obtaining information relevant to the employment process, the costs that arise from the existence of asymmetric information between parties (e.g. [Sch99]), or the costs that arise in the case of unforeseen events that are either not regulated by the contract or are difficult to make legally binding in the case of court cases (e.g. [Har95]). The idea then is that business-related activities can be carried out with significantly less transaction costs within a company than in the free market. There are many reasons for this difference in transaction costs. For instance, within a company the entrepreneur has more control over the actions of his employees, the costs of drafting and signing contracts are smaller, and the risk of unforeseen events is also smaller. Hence, taking into account the existence of transaction costs explains the existence of firms in a way that is compatible with the theory of efficient markets, in that the entrepreneur's overall utility under a firm structure exceeds that of organizing all her work via the market.

4.2 THE BOUNDARY OF THE FIRM

Indeed, the above reasoning can be taken to not only explain the existence of certain kinds of institutions in a market economy but also to explain their particular *structure*.[1] A classic problem in this regard concerns the boundary of the firm, and has been succinctly stated by Williamson as the question "[w]hat efficiency factors determine when a firm produces a good or service to its own needs rather than outsource" [Wil10, p.215]. In other words, the problem studies the question why sometimes firms expand their boundaries (for example by incorporating part of their supply chain) and sometimes they don't.

There are several existing approaches that highlight different efficiency factors in their models. An influential one that focuses on the notion of incomplete contracts is due to [Har95]: The basic idea behind this "property rights"-driven approach is that ownership over firms matters mainly because in the case of events that have not, or could not have, been covered by the contract between a firm and its supplier, it is the ownership structure that determines which party will be able to win any ensuing disagreements. Of course, such a scenario is not relevant if one assumes that contracts between parties are complete, meaning that all possible events pertaining to the economic relationship between the parties are covered by the contract between them. Hence, a key achievement of this approach is to weaken this latter assumption and explicitly model situation involving incomplete contracts.

Another approach was recently presented in [Alf+17]. Here, the central efficiency factor that determines the decision of a firm to integrate one of its suppliers is the ability of the firm to control the transition of the supplier from a generic output product to relationship-specific output product that maximizes the utility of the firm. See also [HR98; Sla06] and references therein for other approaches.

4.3 UNCERTAINTY COST AS EFFICIENCY FACTOR

You know where this is going. In this section I will propose that uncertainty cost should be added to the list of efficiency factors that affect the boundary

1 See also the work of Douglass C. North for attempts to apply essentially Coasian thinking to much broader socio-economic patterns. E.g. [Nor90; Nor05].

of a firm. The simple reason is that by taking ownership of another firm, many sources of uncertainty about that firm are removed, for obvious reasons. In particular, given a firm-supplier relationship, let l be a lottery that the firm manager M faces towards the supplier S. For example, this could concern the question whether S makes an effort to alter their production process to produce a good that is tailor-made to your firm. Or it could concern the question whether S is at the same time dealing with one of your competitors.

I think that it is clear that the "boundary of the firm"-problem can be understood as an event-dependent task, with the manager's task being "maximize the revenue of your company". Specifically, M will have to react differently to different outcomes of l. In particular, she will have to do different things *in order to realize the payout of the lottery* l! This is because the assumption that $A \subseteq \mathbb{R}$ was made mostly for mathematical convenience and because it seemed reasonable that one could replace the outcome of some cardinal-valued lottery (concerning, for example, the supplier's type) with the effective outcome for M (i. e. the achievable revenue given the supplier's type). But what this assumption wipes under the carpet is that different lottery outcomes require different procedures for the actual lottery outcome to be "transformed" into the effective, monetary value of the lottery (recall the example from the introduction)! And it is this fact that makes the "boundary of the firm"-problem an event-dependent task. Indeed, we can see the connection to the message-sending task immediately, because a company will probably prepare for the most likely supplier type and hence have an unusually high transformation cost in case the supplier turns out to be of a very unlikely type, just like in that task.[2] Once this is understood, we see that the EUCH applies and the uncertainty cost for M should be measured by the Shannon entropy.

Now, lets say that M has the opportunity to take ownership of S at the cost κ, meaning that his lottery will be updated to l'. Then the EUCH implies that, since the direct costs of uncertainty are given by $cH(p^{(l)})$, it makes sense for M to make use of this opportunity as soon as $H(p^{(l)}) - H(p^{(l')}) > \frac{\kappa}{c}$. This is because in this case the reduced uncertainty costs outweigh the costs κ of a takeover.

Now, the above may sound a bit like a truism. But in fact it lets us produce a non-trivial empirical prediction: In the very simple reasoning above,

2 Well, you might say, but the costs of transformation could simply be included into the payout of the lottery, by simply reducing the effective payout for outcomes that require cost-intensive procedures. However, this counter-argument neglects the fact that the player could decide to allocate resources in a way that is optimal *given* the probability distribution of the lottery. That is, a player with finite resources for running through procedures might allocate resources in such a way that the expected costs of realizing the gain from a lottery are minimized, just like in the message-sending task.

we can compare different hypothetical situations in which κ, c and the final entropy $H(p^{(l')})$ are fixed and the only thing that changes is the initial entropy $H(p^{(l)})$. In such a situation, the above reasoning implies that the amount of mergers, or integrations of companies, that occur in a country should positively correlates with the uncertainty about supplier's performance, as measured by the Shannon entropy. This is interesting because it seems to yield a first empirical sanity check for the theme of this section, namely that uncertainty costs co-determine the institutional structure in an economy. And indeed we can use existing data from [Alf+17] to support the correctness of this prediction! In that paper, the authors study a significantly more complicated model for the boundary of the firm that is meant to explain both integration and delegation decisions of firms. Using data from two very large databases, amongst others they estimate the following linear probability model:

$$\begin{aligned} \text{Integration}_{f,j,c,i} = &\gamma_0 + \gamma_1 \text{CVProductivity}_{i,c} \\ &+ \gamma_2 \text{MeanProductivity}_{i,c} \\ &+ \gamma_3 \mathbf{X}_f + \delta_i + \delta_f + \epsilon_{f,j,c,i} \end{aligned} \tag{9}$$

In this model, $\text{Integration}_{f,j,c,i}$ is a dummy variable that takes the value 1 if some firm f with primary activity in some business sector j (according to a classification of the database) and located in country c integrates an input i into their boundaries. $\text{CVProductivity}_{i,c}$ denotes the coefficient of variation (standard deviation/mean) of productivity of suppliers in input industry i located in country c. The value of this variable is determined using the sales per employee in thousands US Dollars as a proxy for supplier productivity. Further, $\text{MeanProductivity}_{i,c}$ is the mean for the above proxy, $\mathbf{X}_f$ is a vector of firm-level controls such as Employment, Age and Higher Education, and finally, δ_i, δ_f denote input industry fixed effects (FE) and firm FE respectively. In total, the regression involves 15 million firms located in 20 countries, restricting to industry-country (i, c) pairs with at least 50 independent supplier in that industry and country. The results for the regression are given in Fig. 1 below.

In this table, the columns denote different runs for the model with successively stronger controls (for example, fixed effect controls are stated for every kind of fixed effect in the lower part of the table). The upshot is that the value for γ_1, which is stated in the first row, is found to be highly significant and positive in every run. Moreover, the magnitude of the effect is also found to be significant, in that the regression shows that a one-standard-deviation increase in $\text{CVProductivity}_{i,c}$ increases the probability to integrate by 39 percent compared to the baseline probability of 1 percent. At the same

	(1)	(2)	(3)	(4)
CV Productivity$_{i,c}$	0.00066***	0.00063***	0.00063***	0.00062***
	(0.00013)	(0.00013)	(0.00013)	(0.00013)
Mean Productivity$_{i,c}$	-0.00000**	-0.00000**	-0.00000**	-0.00000**
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
log(Employment$_f$)			0.00636***	
			(0.00037)	
log(1+ Age$_f$)			0.00007	
			(0.00028)	
log(% Workforce with a College Degree$_p$)			-0.00090***	
			(0.00023)	
Country FE	yes	yes	yes	no
Input FE	yes	yes	yes	yes
Output FE	no	yes	yes	no
Firm FE	no	no	no	yes
R-squared	0.018	0.036	0.048	0.079
N	250,925	250,925	250,925	250,925

Figure 1: Results for the regression model (9), reproduced from [Alf+17]: The estimated coefficients are stated together with standard errors clustered at the input level in parentheses. $***, **$ and $*$ indicate statistical significance at the $1\%, 5\%$ and 10% levels respectively.

time, the mean productivity is found, again in all columns and again with high significance, to be irrelevant for the question of integration.

Now, in what way do these findings produce support for my prediction, that integration should positively correlate with the initial entropy $H(p)$ faced by a manager? To answer this, lets make the reasonable assumption that the distribution of $\text{CVProductivity}_{i,c}$ is Gaussian. Now, for a Gaussian distributions p with mean μ and standard deviation σ, we have that

$$H(p) \propto \log(2\pi e \sigma^2),$$

so that, by monotonicity of the logarithm, if the probability of integration positively correlates with the initial entropy, it should also positively correlate with the standard deviation, while the mean should be irrelevant (since the entropy is independent of μ). But this is exactly what the findings for $\text{CVProductivity}_{i,c}$ and $\text{MeanProductivity}_{i,c}$ say! Hence, the empirical data are consistent with the prediction that the uncertainty costs are a relevant efficiency factor that co-determine the boundary of the firm.

Of course, it would be interesting to repeat the above regression in such a way that logarithmic dependence on $\text{CVProductivity}_{i,c}$ is checked. This could be a way to discriminate between the support for the model of [Alf+17] (in which risk positively correlates with integration because it creates a larger "option value" for delegation and not because of uncertainty costs) and the ideas of this thesis. However, this lies beyond the scope of this essay.

5

APPLICATION: DECISION THEORY

In Section 2, the insurance-based approach for the quantification of uncertainty cost had, as its starting point, the EUH, a hypothesis about how players determine their preferences between lotteries. In this section, I want to present an alternative model for preferences between lotteries that incorporates the Shannon entropy.

5.1 AN ALTERNATIVE TO THE EUH

Let u be a Bernoulli utility function of some player P that is faced with an event-dependent task, so that EUCH applies. I do not make any assumptions on the properties of u other than that it always leads to well-defined values for $Eu(l)$. The model that I propose is one in which a weak version of the EUH holds in the sense that P *would* decide between lotteries based solely on expected utility, if uncertainty costs were absent. But alas, they are not, and so the utility of a given lottery is given by[1]

$$U(l) = Eu(l) - cH(p^{(l)}). \tag{10}$$

As we see, the idea is very simple: The utility of a lottery is its expected utility less the uncertainty costs associated to the lottery. The value of c depends on the setting and by empirically estimating it one can get an idea about just how strongly uncertainty costs affect utility. I now propose the following hypothesis:

EUH WITH UNCERTAINTY COST (EUHU) Given two lotteries l and m, a decision maker will prefer that lottery with the larger value of $U(l)$.

Denote by $\succ_H$ the ordering over lotteries that is induced by this utility function, i.e. $l \succ_H m \Leftrightarrow U(l) \geqslant U(m)$.

Let me now to a discussion of the difference between EUH and EUHU: To begin with, one may ask whether $\succ_H$ may not actually be modeled by EUH. That is, maybe there exists another Bernoulli utility function u' such that $l \succ_H m \Leftrightarrow Eu'(l) \geqslant Eu'(m)$. If this was the case, then EUHU wouldn't actually add anything very interesting to the debate. Luckily (for me), the answer

1 This quantity is similar to the free energy in thermodynamics, which is defined as $F(p) = \langle E \rangle_p - TH(p)$, where $\langle E \rangle$ is the average energy and T is the temperature of a heat bath modeling the environment.

to this question is negative. Intuitively, this follows from the fact that $U(l)$ is a non-linear function (since $H(p)$ is non-linear), while Eu' will always be linear and hence cannot reproduce the same ordering. Can we make this intuition precise? Yes, we can and the answer lies in the *von-Neumann-Morgenstern utility theorem*. To be able to state the theorem, we need two more properties for orderings. Call an ordering *continuous*, if

$$l \succ k \succ m \Rightarrow \exists \lambda \in [0,1] \text{ s.t. } \lambda l + (1-\lambda)m \sim k.$$

Call an ordering *independent* if

$$l \succ m \Rightarrow \lambda l + (1-\lambda)k \succ \lambda m + (1-\lambda)k, \quad \forall k \in \mathcal{L}, \lambda \in [0,1].$$

Independence represents the idea that preferences are robust to being mixed with alternatives. The theorem then states the following:

Theorem 6 ([NM53]) *Let $(\succ, \mathcal{L})$ denote an ordering. Then the following are equivalent:*

1. *$(\succ, \mathcal{L})$ is transitive, complete, continuous and independent,*
2. *There exists a $u : \mathbb{R} \to \mathbb{R}$ such that*

$$l \succ m \Leftrightarrow Eu(l) \geqslant Eu(m).$$

Now, it is easy to see that by definition $\succ_H$ is complete, transitive and continuous. *However it is not, in general independent*! How can we see this? Let me give two examples

Example 7 *Consider three lotteries $l, m, k \in L_2$ with the same expected utility $Eu(l) = Eu(m) = Eu(k)$. Since all lotteries involve only two outcomes, we have that*

$$U(l) = Eu(l) - H_2(p_1^{(l)}),$$

where I set $c = 1$ and $H_2(p) = -p\log p - (1-p)\log(1-p)$ is the binary entropy shown in Fig. 2. Now, under the above assumptions, $l \succ_H m \Leftrightarrow p_1^{(l)} \leqslant p_1^{(m)}$. To violate independence, assume that $p_1^{(l)} < p_1^{(m)} \leqslant 0.5 < p_1^{(k)}$. These values imply that $l \succ_H m$ and that there exists a value for $\lambda \in (0,1)$ such that

$$\tilde{p}_l \equiv \lambda p_1^{(l)} + (1-\lambda)p_1^{(k)} = 0.5$$
$$\tilde{p}_m \equiv \lambda p_1^{(m)} + (1-\lambda)p_1^{(k)} > 0.5$$

But this implies that $H_2(\tilde{p}_l) > H_2(\tilde{p}_m)$ and hence $m + (1-\lambda)k \succ_H \lambda l + (1-\lambda)k$, in violation of independence.

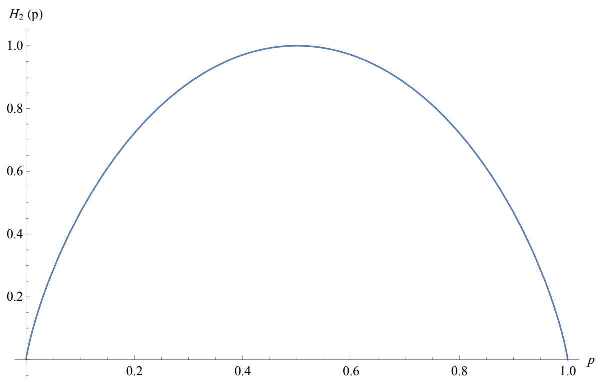

Figure 2: The binary Shannon entropy

What was going here? Mathematically, we could produce this violation only because of the non-linearity of $H_2(p)$ (specifically, the fact that $H_2(p)$ has a maximum in the middle). But conceptually we also have a clear picture: Mixing with an alternative lottery (in this case k) has a different effect on the uncertainty of the resulting lotteries. If, for example, l was a lottery with no uncertainty ($p^{(l)} = (1, 0)$) at all and m was maximally uncertain $p^{(m)} = (1/2, 1/2)$, then mixing with k would definitely lead to an *increase* in uncertainty costs for l while it would also definitely *reduce* the uncertainty costs for m. But this dependence is exactly what independence tries to rule out, and so it makes total sense that $\succ_H$ violates independence.

All of this may sound overly abstract. The second example is slightly more hands-on.

Example 8 *Consider a manager* M *of a company that produces precious earrings with production function* $\rho(x) = \sqrt{x}$ *and her metal supplier* S. M *has to place an order in metal but she is unsure whether the metal that* S *supplies will be good or bad. Now, in case the metal is good,* M *can sell the earrings at a price of* $2\$$ *per gram, but in case they're bad, she can only sell them for* $1\$$ *per gram. In both cases she buys the metal for* $0.5\$$ *a gram. So her utility function is*

$$u(a, x) = a\sqrt{x} - 0.5x$$

with $a \in \{1, 2\}$*. Let* p *be the probability that the metal is bad, so that* $l = ((1 - p, p), A = (2, 1))$ *is the lottery describing the situation. Plugging this into (10) (and setting* $c = 0.5$ *yields)*

$$\begin{aligned} U(l) &= \max_{x \in \mathbb{R}^+} \left[p(\sqrt{x} - 0.5x) + (1 - p)(2\sqrt{x} - 0.5x) \right] - 0.5H_2(p) \\ &= 2(1 - p) + \frac{p^2}{2} - 0.5H_2(p), \end{aligned}$$

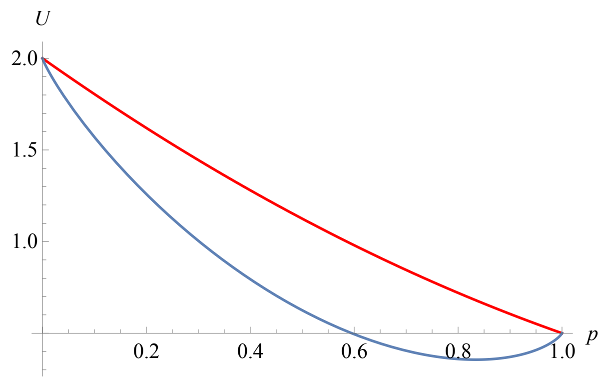

Figure 3: $U(l)$ with (blue) and without (red) uncertainty costs. The minimum at $p \approx 0.8$ breaks the independence property.

where the fact that the maximization is over a single variable x *reflects the fact that the order has to be placed before the quality is revealed and in the second step I have plugged in the solution to the simple optimization problem. In Fig. 3 I have plotted* $U(l)$ *as a function of* p *both with (blue) and without (red) the uncertainty cost term. We see that without the uncertainty cost the function satisfies independence (because of strict monotonicity) but adding uncertainty cost breaks this independence, by introducing a minimum, just like in the previous example.*

To summarize, EUH cannot model the predictions of EUHU and we understand very clearly why that is the case. Now, lets put EUHU to use and solve some riddles.

5.2 RESOLVING PARADOXES

In this subsection I show that the preference ordering $\succ_H$ does not suffer from Allais' paradox, Rabin's paradox or the Ellsberg paradox.

5.2.1 *Allais' paradox*

Allais' paradox [All53] is the standard example to illustrate the limits of the EUH. In fact, it was designed by Maurice Allais to do just that. It goes as follows: Consider the outcome set $A = (5, 1, 0)$ and the following four lotteries:

$$l_1 = ((0, 1, 0), A) \qquad l_2 = ((0.1, 0.89, 0.01), A)$$
$$l_3 = ((0.1, 0, 0.9), A) \qquad l_4 = ((0, 0.11, 0.89), A)$$

If you ask people for their preferences between the first two and second two respectively, then most of them will prefer l_1 over l_2 and l_3 over l_4 (see, for example, [Oli03] for an empirical study). But *no linear preference criterion* can reproduce this. And so EUH cannot do justice to those empirical findings. What would EUHU say? If we set, for example, $u(a) = \sqrt{a}$ and $c = 0.2$, then (10) yields

$$\begin{aligned} U(l_1) &= 1 & U(l_2) &\approx 0.95 \\ U(l_3) &\approx 0.36 & U(l_4) &\approx 0.26, \end{aligned}$$

so that we indeed reproduce the empirically observed ordering. Upon reflection, it seems in fact as if uncertainty cost is exactly what Allais' paradox is about! What goes on in the paradox is that in the first pair of lotteries, a significant amount of uncertainty is added to the lottery, so that the prospect of making significantly more money in l_2 is outweighed by the additional uncertainty cost. In contrast, in the second pair of lotteries, the uncertainty costs are very similar (since the distribution are almost the same) and so the preference of the higher payout dominates the decision.

I should note that this interpretation of Allais' paradox, as being concerned with uncertainty, is different from the one that rank-dependent utility theory, for example, produces [Qui93; TK92]. In such approaches, people's behavior is explained as an overweighting of unlikely and extreme (in terms of payout) events. In the EUHU approach, just like in the above theories, there is a high and non-linear cost associated to unlikely events (in the sense that the value $-\log p_i$ becomes large if p_i is very small), but in contrast to those theories this is independent of the payout.

5.2.2 *Rabin's Paradox*

Next, turn to another setting that EUH fails in — Rabin's Paradox [Rab00]. The message of this paradox is that in EUH, slight risk aversion about lotteries with relatively small payouts can be made to imply very high risk aversion about lotteries with relatively large payouts. Consider, for example, a player P with a Bernoulli utility function u which is such that, for every amount $w \in \mathbb{R}$,

$$\frac{1}{2}u(w-10) + \frac{1}{2}u(w+11) < u(w). \tag{11}$$

Clearly, P is risk averse, because for monotonically increasing u, the above implies that u is concave. And so if P follows the EUH, then given the choice, he would, for any amount w, prefer the lottery $l = ((1,0),(w,0))$ to the lottery $m = ((1/2,1/2),(w+10,w-10))$. Now, surely P's risk aversion has a limit. To find this limit, we can ask: What is the least amount $v \in \mathbb{R}$ such that P would prefer the lottery $k = ((1/2,1/2),(w+v,w-100))$ to the lottery

l. Rabin's paradox then is this: If the EUH is true, then given (11), there is *no* amount v such that P would prefer k to l. But this seems silly, because nobody could be that risk averse. Indeed, the attack against the EUH that lies behind this paradox is related to one of my criticisms against the insurance approach from Section 2, namely that the utility function u is, in a sense, abused if it is made to model all aspects of risk aversion (including uncertainty) alone.

How does EUHU deal with the situation? It does very easily. First, we want to produce a situation in which P has a utility function u such that, for any amount w,

$$l \succ_H m. \tag{12}$$

This is simple, just choose $u(a) = a$ and $c = 1$, so that $U(l) = w$ and $U(m) = w - \frac{1}{2}$, since $H(1/2, 1/2) = 1$ and (12) follows. At the same time, for any $v \geqslant 102$ we have that $k \succ_H l$ and nothing paradoxical is in view.

5.2.3 *Ellsberg Paradox*

Since we seem to be on a streak, let me consider one more paradox for the EUH — the so-called Ellsberg Paradox [Ell61]. It considers a setting in which a player is not only uncertain about the outcome of a lottery, but in fact uncertain about the distribution itself. It is generally taken to show that people exhibit an aversion against ambiguity and that the EUH does not reflect this aversion [GS89]. So how does it go? Consider an urn with 90 balls, 30 of which are red and the remaining 60 of which are either yellow or black. As in Allais' paradox, you are invited to choose your favorite among the following two pairs of lotteries.

l_A	l_B
If you draw a red ball, you get 100$.	If you draw a black ball, you get 100$.

l_C	l_D
If you draw a red or yellow ball, you get 100$.	If you draw a black or yellow ball, you get 100$.

If you are like most people, you will prefer l_A over l_B and l_D over l_C. The common explanation for this is that in both pairs of lotteries, the expected return is the same but in l_A and l_D, you know exactly what your odds of winning are, while in l_B and l_C, your odds depend on the actual distribution of black and yellow balls. EUH cannot reflect this fact and it says that you should either prefer l_A and l_C, or l_B and l_D.

Does EUHU deal with this paradox? It does, and as with the other two cases it does so very naturally I think. The key insight required is that the Shannon entropy decreases over coarse-graining. By this I mean the following: Let $p = (p_1, \dots, p_n)$ be some probability vector. Now, take any partition of the set $\{1, \dots, n\}$ into m subsets I_i and define the vector

$$\tilde{p} = (\sum_{j \in I_1} p_j, \sum_{j \in I_2} p_j, \dots, \sum_{j \in I_m} p_j)$$

This is a valid m-dimensional probability vector that is obtained from turning several entries of p into a single one of $\tilde{p}$. For this reason, any $\tilde{p}$ constructed in this way is called a *coarse-graining* of p. Now, the Shannon entropy is such that, for any p, $H(\tilde{p}) < H(p)$ for any non-trivial coarse-graining of p (this follows from the concavity of the logarithm). This makes sense for a measure of uncertainty, since surely one is less uncertain about the outcome of a coarse-grained distribution than one is about the original distribution. Now, how does this help with Ellsberg paradox? Since I don't want to get caught up in details, in Appendix A I show that, for the utility function $u(a) = a$, the above lotteries are such that $Eu(l_A) = Eu(l_B)$ and $Eu(l_C) = Eu(l_D)$, while $p^{(l_A)}$ is a coarse-graining of $p^{(l_B)}$ and $p^{(l_D)}$ is a coarse-graining of $p^{(l_C)}$. Plugging this into (10) then implies that, for any $c > 0$, $l_A \succ_H l_B$ and $l_D \succ_H l_C$, as required.[2]

5.3 UNCERTAINTY COST AS PSYCHOLOGICAL COST

At this point, you may think: Well, this is all nice, but what really is the connection between the Shannon entropy and the uncertainty cost in these paradoxes? Recall that I assumed at the beginning of this section that P faces an event-dependent task, so that EUCH applies. So your question really is: What are the event-dependent tasks in the above paradoxes and how strong is the analogy between them and the message-sending task from Sec. 3.2? My answer is that the task in all three cases is simply "Make the best decision"

2 Underlying this approach is a reductionist strategy that introduces a *Second Order Probability* (SPO) distribution over possible distributions of the yellow and black balls [CW92]. As will be clear to the reader of Appendix A, I make the assumption that this SPO is uniform and a critic might argue that this is an unwarranted recourse to Laplace's Principle of Indifference and that without motivating it I have to enter an infinite regress invoking higher order distributions. However, to this I can reply that the monotonicity of the Shannon entropy with respect to coarse-graining holds independently of this assumption and that *every* SPO (with and without the same expectation value) will allow me to solve the Ellsberg paradox. Hence, I do not need to enter an infinite regress and my assumption of uniformity was merely for convenience.

and that the analogy holds perfectly fine if one identifies "cognitive effort" or "psychological stress" as the relevant resource. If your task is to make the best decision then in solving this task you have to consider and weigh all possible events. In such situations unlikely events are particularly "costly" for a number of reasons, which include, but are not limited to, having to calculate with small, unintuitive numbers, or having to keep an event in mind despite it probably never happening anyway. All of these are costs associated to making the best decision, and indeed we can interpret the values $x_i = -\log p_i$ of the random variable X in (8) as the *psychological cost of being uncertain about an event that occurs with probability* p_i. [3]

3 In this thesis, I have not discussed where the probability distributions come from. But this is an important and subtle question: They might be "objective" or "frequentist" and represent, for example, the relative frequencies of lottery outcomes in past rounds; Or they might be "subjective" or "classical" and represent a player's limited knowledge of some hidden information [Köh17]. This distinction is implortant because it affects the meaningfulness and economic soundness of any approach attempting to quantify uncertainty cost. For example, the motivation for the Shannon entropy I presented in section 3.2 implicitly assumed the probabilities were of the first, frequentist type. In contrast, the "psychological stress"-motivation I present in this section goes naturally with a subjective probability. While I maintain that this motivation is meaningful, since lotteries can meaningfully be thought with subjective probability vectors, I should note here that there is a well-developed account of decision theory for subjective probabilities, including a subjective expected utility hypothesis (SEUH), that is based on the notion of *acts* rather than lotteries [Sav54; CW92]. It would be interesting to apply my arguments to this approach, and compare its non-linear approach with so-called *probabilistic sophistication* approaches to generalizing (SEUH) [MS90; KQ92]. However, I cannot do so here because of the limited scope of the thesis.

6 SUMMARY, OUTLOOK AND ACKNOWLEDGMENTS

SUMMARY

Let me summarize: In this thesis, I have developed a conceptual and formal account for direct uncertainty costs, which are costs that arise for economic agents from the very state of being uncertain. I have first argued that a naive micro-economic approach for quantifying such costs suffers from a number of problems. In discussing those problems I have also shown that there is a very simple way to understand how, both conceptually and formally, the notions of risk and uncertainty are related. I have then introduced axioms that single out the Shannon entropy as a suitable measure of uncertainty costs and also presented reasons to understand that what the Shannon entropy exactly measures are the minimal costs due to an agent having to react differently to different outcomes in an economic problem they face. I have then applied these results to two different economic fields: First, I have applied them to the field of transaction cost economics, and argued, both theoretically and by giving empirical support, that uncertainty cost is a quantity that co-determines the decision of a firm whether to integrate part of its supply chain or not. Finally, I have applied my results to the field of decision theory, where I used them to present EUHU, an alternative to the expected utility hypothesis, and shown that EUHU naturally solves Allais', Rabin's and the Ellsberg Paradox.

OUTLOOK

What are reasonable next steps for the future? At a formal level, I believe that the study of more general entropic quantities than the Shannon entropy in both decision theory and transaction cost economics could be of great interest. For instance, an entropic function that could be used to quantify the psychological stress in decision theoretic settings alternatively to the Shannon entropy is the so-called *Berg entropy* which is simply given as $H_B(p) = -\sum_i \log p_i$ and hence corresponds to the total stress, in the above sense, induced by a lottery, as opposed to the average stress as measured by the Shannon entropy. The same applies to the connection between riskiness, uncertainty and majorization. The latter is very well studied mathematically and it seems almost certain that many interesting results for economics could follow by checking those results.

At a conceptual level, I believe that the results of this thesis could be applicable to not only build models of inter-company relationships but also intra-company structure. That is, in a vein similar to the program of [Gar00], the internal organization of companies could be understood as an attempt to minimize the internal uncertainty costs.

At an empirical level, it would be interesting to use the datasets from [Alf+17] to test more rigorously my claim that the Shannon entropy positively correlates with integration decisions of companies.

ACKNOWLEDGEMENTS

I would like to thank Sebastian Schweighofer-Kodritsch for preparatory discussions, always being available for any questions I might have had and comments on an earlier draft. I would like to thank Babette Semmer and especially Prof. Robert Schmidt for extensive feedback on an earlier draft. I would also like to thank Prof. Urs Schweizer for clarifying feedback on a previous idea sketch for a topic that never made it into this thesis.

BIBLIOGRAPHY

[Alf+17] L. Alfaro, N. Bloom, P. Conconi, H. Fadinger, P. Legros, A. F. Newman, R. Sadun, and J. Van Reenen. *Come Together: Firm Boundaries and Delegation*. Harvard Business School Working Paper 18-051. 2017.

[All53] M. Allais. "Le Comportement de l'Homme Rationnel devant le Risque: Critique des Postulats et Axiomes de l'Ecole Americaine." In: *Econometrica* 21.4 (1953), pp. 503–546.

[Atk70] A. B. Atkinson. "On the measurement of inequality." In: *Journal of Economic Theory* 5 (1970), pp. 244–263.

[CW92] C. Camerer and M. Weber. "Recent Developments in Modeling Preferences: Uncertainty and Ambiguity." In: *Journal of Risk and Uncertainty* 5 (1992), pp. 325–370.

[Coa37] R. H. Coase. "The Nature of the Firm." In: *Economica* 4.16 (1937), pp. 386–405.

[Csi08] I. Csiszar. "Axiomatic Characterizations of Information Measures." In: *Entropy* 10 (2008), pp. 261–271.

[Dal20] H. Dalton. "The measurement of the inequality of incomes." In: *Economic Journal* 30 (1920), pp. 348–361.

[Ell61] D. Ellsberg. "Risk, Ambiguity, and the Savage Axioms*." In: *The Quarterly Journal of Economics* 75.4 (1961), pp. 643–669.

[Gar00] L. Garicano. "Hierarchies and the Organization of Knowledge in Production." In: *Journal of Political Economy* 108 (2000).

[GS89] I. Gilboa and D. Schmeidler. "Maxmin expected utility with non-unique prior." In: *Journal of Mathematical Economics* 18.2 (Apr. 1989), pp. 141–153.

[HLP29] G. H. Hardy, J. E. Littlewood, and G. Polya. "Some simple inequalities satisfied by convex function." In: *Messenger Math.* 58 (1929), pp. 145–152.

[HLP39] G. H. Hardy, J. E. Littlewood, and G. Polya. *Inequalities*. London and New York: Cambridge University Press, 1939.

[Har95] O. Hart. *Firms, Contracts, and Financial Structure*. Oxford: Clarendon Press, 1995.

[HR98] B. Holmstrom and J. Roberts. "The Boundaries of the Firm Revisited." In: *Journal of Economic Perspectives* 12.4 (Dec. 1998), pp. 73–94.

[Kas18] T. Kas. *Shannon Entropy and DNA (last access 24.6.2018)*. 2018. URL: `http://asserttrue.blogspot.com/2014/05/shannon-entropy-and-dna.html`.

[KQ92] D. Kelsey and J. Quiggin. "Theories of choice under ignorance and uncertainty." In: *Journal of Economic Surveys* 6.2 (1992), pp. 133–153.

[Kni21] F. H. Knight. *Risk, uncertainty and profit*. Charleston: Nabu Press, 1921.

[Köh17] J. Köhn. *Uncertainty in Economics*. Cham: Springer, 2017.

[Lor05] M. O. Lorenz. "Methods of measuring concentration of wealth." In: *Journal of the American Statistical Association* 9 (1905), pp. 209–219.

[MS90] M. Machina and D. Schmeidler. "A more robust definition of subjective probability." In: *UCSD, Discussion Paper* 90-29 (1990).

[MOA11] A. W. Marshall, I. Olkin, and B. C. Arnold. *Inequalities : theory of majorization and its applications*. Springer, 2011, p. 909.

[Mui03] R. G. Muirhead. "Some methods applicable to identities and inequalities of symmetric algebraic functions of n letters." In: *Proc. Edinburgh Math. Soc.* 21 (1903), pp. 144–57.

[NM53] J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton, NJ: Princeton University Press, 1953.

[Nor90] D. C. North. *Institutions, institutional change and economic performance*. Cambridge: Cambridge University Press, 1990.

[Nor05] D. C. North. *Understanding the process of economic change*. Princeton, NJ: Princeton University Press, 2005.

[Oli03] A. Oliver. "A quantitative and qualitative test of the Allais paradox using health outcomes." In: *Journal of Economic Psychology* 24.1 (2003), pp. 35–48.

[Qui93] J. Quiggin. *Generalized Expected Utility Theory*. Amsterdam: Springer Netherlands, 1993.

[Rab00] M. Rabin. "Risk Aversion and Expected-utility Theory: A Calibration Theorem." In: *Econometrica* 68 (2000), pp. 1281–1292.

[RS70] M. Rothschild and J. Stiglitz. "Increasing Risk: A Definition." In: *Journal of Economic Research* 2 (1970), pp. 225–243.

[RS71] M. Rothschild and J. Stiglitz. "Increasing Risk: Its Economic Consequences." In: *Journal of Economic Research* 3 (1971), pp. 66–84.

[Sav54] L. Savage. *The Foundations of Statistics*. New York: Dover Publications, 1954.

[Sch99] U. Schweizer. *Vertragstheorie*. Neue ökonomische Grundrisse. Tübingen: Mohr Siebeck, 1999.

[Sha48] C. Shannon. "A mathematical theory of communication." In: *Bell System Technical Journal* 27.3 (1948), pp. 379–423.

[Sla06] M. Slater. *The boundary of the firm*. Ed. by Andrew Campbell and David O. Faulkner. The Oxford Handbook of Strategy. Oxford: Oxford University Press, 2006.

[Str65] V. Strassen. "The existence of probability measures with given marginals." In: *Annals of Mathematical Statistics* 36 (1965), pp. 423–439.

[TK92] A. Tversky and D. Kahneman. "Advances in prospect theory: Cumulative representation of uncertainty." In: *Journal of Risk and Uncertainty* 5.4 (Oct. 1992), pp. 297–323.

[Var14] H. Varian. *Intermediate Microeconomics - A Modern Approach*. 9th ed. New Yorks: W. W. Norton & Company, 2014.

[Wil85] O. E. Williamson. *The economic institutions of capitalis*. China Social Sciences Publishing House, 1985.

[Wil10] O. E. Williamson. "Transaction Cost Economics: The Natural Progression." In: *Journal of Retailing* 86.3 (2010), pp. 215–226.

A

COARSE-GRAININGS OF LOTTERIES

In this appendix, I give the details for how the EUHU deals with the Ellsberg paradox. I will consider only the first pair in detail, since the situation with the second pair is exactly analogous. Consider again the first pair

l_A	l_B
If you draw a red ball, you get 100$.	If you draw a black ball, you get 100$.

We can straightforwardly write $l_A = ((1/3, 2/3), (100, 0))$. But what are the distribution $p \equiv p^{(l_B)}$ and outcome set $A \equiv A^{(l_B)}$ corresponding to lottery l_B? To tackle this question, note that there are 60 possible distributions of black and yellow balls in the paradox. In the first, there are no black balls and 60 yellow balls, in the second there is one black ball and 59 yellow balls, and so on. Now, for each of these distributions, there are two events, one in which one wins 100$ (that is, if the ball is black) and one in which one wins nothing (if the ball is not black). Hence, there are in total 120 events distinct events. Now, the probability of the event "m black balls and win" is $\frac{m}{60 \cdot 90}$, since with probability $\frac{1}{60}$ there are m black balls and given that there are m black balls, the chance of drawing one of them is $\frac{m}{90}$. By the same reasoning, the probability of the event "m black balls and lose" is $\frac{90-m}{60 \cdot 90}$. In total, this yields

$$l_B = (p = \frac{1}{60 \cdot 90}(\underbrace{0, 1, \dots, 60}_{\text{win-events}}, \underbrace{90, 89, \dots, 30}_{\text{lose-events}}),$$
$$A = (\underbrace{100, 100, \dots, 100}_{\text{60 times}}, \underbrace{0, \dots, 0}_{\text{60 times}})).$$

Now, using $\sum_{i=1}^{n} i = \frac{(n+1)n}{2}$ and $u(a) = a$, we can first calculate

$$Eu(l_B) = \sum_{m=1}^{60} \frac{100}{60 \cdot 90} m = \frac{61 \cdot 100}{90 \cdot 2} \approx \frac{100}{3},$$

which coincides with $Eu(l_A)$. Moreover, lets coarse-grain p by collecting all the win-events into one entry and the lose-events into a second one. This results in a distribution

$$\tilde{p} = \frac{1}{60 \cdot 90}\left(\sum_{m=1}^{60} m, \sum_{m=1}^{60}(90-m)\right) \approx (1/3, 2/3),$$

which shows that $p^{(l_A)}$ is a coarse-graining of $p^{(l_B)}$ as claimed.

By exactly the same reasoning we also arrive at the fact that $p^{(l_D)}$ is a coarse-graining of $p^{(l_C)}$.

DECLARATION OF AUTHORSHIP

Hiermit erkläre ich, Paul Boes, dass ich die vorliegende Arbeit noch nicht für andere Prüfungen eingereicht habe. Ich habe die Arbeit selbständig verfasst. Sämtliche Quellen einschließlich Internetquellen, die ich unverändert oder abgewandelt wiedergegeben habe, insbesondere Quellen für Texte, Grafiken, Tabellen und Bilder, habe ich als solche kenntlich gemacht. Ich bin mir darüber bewusst, dass bei Verstößen gegen diese Grundsätze ein Verfahren wegen Täuschungsversuchs bzw. Täuschung eingeleitet wird. Ich versichere zudem, dass ich noch keine Abschlussarbeit im Bereich der Wirtschaftswissenschaften eingereicht habe

I, Paul Boes, hereby declare that I have not previously submitted the present work for other examinations. I wrote this work independently. All sources, including sources from the Internet, that I have reproduced in either an unaltered or modified form (particularly sources for texts, graphs, tables and images), have been acknowledged by me as such. I hereby confirm that I have authored this Bachelor's thesis independently and without use of others than the indicated sources. All passages which are literally or in general matter taken out of publications or other sources are marked as such. I understand that violations of these principles will result in proceedings regarding deception or attempted deception. I further ascertain that I have not submitted a Bachelor's or Master's Thesis in the field of economics before.

Berlin, July 12, 2018