Decision Making under Extreme Uncertainty

Michael Schwarz*
Harvard University
Department of Economics

May 1999

Abstract

The current paper derives restrictions on an agent's behavior from invariance principles, in particular, from the requirement that an agent's choices be independent of the units of measurement. It is shown that invariance principles are sufficient for approximating an agent's behavior in environments where little information is available. Strong restrictions on admissible shape of tails of the posterior distribution of an unknown variable are derived. For the case of independently and identically distributed draws, closed form solution for the family of tails of posterior distribution is obtained. It is shown that even if an agent assumes that the variable in question is drawn from a finite parametric family of distributions with exponential-like tails, the posterior distribution obtained by integrating out the unknown parameters has very fat tails. Many results obtained in this paper do not rely on expected utility axioms and thus could be combined with either expected or non-expected utility theories. Decision problems arising in the areas ranging from industrial R&D planning to risk management provide motivation and potential applications for the theory developed herein.

*I am grateful to Kenneth Arrow, Jeremy Bulow, Darrell Duffie, Edward Lazear, Ilya Segal, and Robert Wilson. Also, I would like to thank E. Dahan, L. Mirny, and M. Yildiz. Send correspondence to Michael Schwarz, Stanford University Graduate School of Business, Stanford, CA 94309, Phone (650)725-6875, Email: pschwarz@leland.stanford.edu.

1. Introduction

The current paper investigates decision making in environments of extreme uncertainty. We use invariance principles to derive restrictions on an agent's behavior. According to invariance principles, an agent treats the units of measurement as an uninformative label. For instance, a broker can perform calculations in dollars or in cents, but resulting decision to buy or sell should not depend on the choice of the measurement scale. The class of decision problems considered in the current work can be illustrated by the following example. In 1986, scientists at IBM Zurich research laboratories discovered high temperature superconductors. A numerical measure of a superconductor performance is the maximum current that it can sustain. The earliest high-temperature superconductors deliver performance below the threshold of practical usability. Furthermore, a number of projects are complementary to the advancement in high-temperature superconductor technology if the performance of high-temperature superconductors falls in some range. In the interests of reducing time to market, such projects may be carried out before semiconductor technology achieves the required performance level. Consequently, planning R&D activities in areas ranging from accelerator design to the development of high-speed trains requires a forecast of the performance that high temperature superconductors will deliver in the future. In this case, the relative likelihood of performance realizations exceeding some threshold (i.e., the tails of the distribution) may be of particular interest. The uncertainty in such circumstances appears extreme. There are no stationary trends, no data points close to the relevant values of a variable and no theory to guide the forecast. This provides an example of an environment approximating an information vacuum.

We investigate decision making in environments where little information is available. The issues considered here pertain to a broad range of problems. For instance, we make statements about the shape of tails of the distribution that an agent assigns to a positive random variable.[1] The shape of the tails is important in a number of contexts. For example, Weitzman (1979) and Morgan and Manning (1985) study optimal search; the best search strategy critically depends on the distribution of extreme values.[2] Also, the distribution of outliers of asset returns has been studied by Mandelbrot (1963), Fama (1965) and more recently by Kearns

[1] The tails of the size distributions of firms, cities, bankruptcies and many other are of interest for economists. See for instance Zipf (1950), Stanley et al. (1995).

[2] Dahan and Mendelson (1998) study optimal R&D prototyping strategies; their results are sensitive to assumptions about the shape of tails of the distribution.

and Pagan (1997). A comprehensive theoretical framework for statistical study of the extreme values can be found in Gumbel (1958) and Galambos (1978).[3]

The results and the approach of this paper are very different from the above-mentioned literature. Extreme value theory is based on classical statistics. In contrast, our arguments have a Bayesian flavor.

We derive restrictions on an agent's behavior from invariance principles. This makes our approach similar to the one used for deriving non-informative priors in Bayesian statistics.[4] Invariance principles require that an agent's actions and beliefs be independent of the units of measurement that are being used; such a restriction may be interpreted as a basic rationality requirement. The invariance restrictions are stated in Section 2.

The invariance approach was introduced into statistics by Jeffreys (1932).[5] The existing applications of invariance to decision science are in the context of expected utility theory. In contrast, most of the results reported in Part I are obtained exclusively from invariance considerations and do not rely on expected utility or Savage axioms. Consequently, our results can be combined with expected or non-expected utility theories. This is important, because in the cases where choices involving rare events or decision problems under extreme uncertainty are concerned applicability of expected utility theory is debatable.[6] At the same time, the invariance principle is particularly relevant when information is scarce.

Even when substantial data relevant for estimating a random variable is available, the data may be of limited value for estimating the tails of the distribution. Whenever a decision depends on an agent's beliefs about the behavior of the tails of the distribution, we will refer to such decision problems as decision making under information scarcity. In Savage's framework, decision making under information scarcity is characterized by the tails of the agent's prior. We show that invariance restrictions imply that even if some relevant data is available to an agent, her beliefs about the tails of the distribution (i.e. information scarcity region) are well approximated by the Jeffreys' ignorance prior. Jeffreys (1961)

[3] Embrechts, Kluppelberg and Mikosch (1997) is an up to date monograph on extreme value theory.

[4] See for example, Berger (1985).

[5] Many laws of physics can be obtained form invariance postulates in areas ranging from electromagnetism to relativity theory. In economics, invariance ideas were used to deepen our understanding of bargaining (Nash (1950)). There are numerous applications of invariance in econometrics; see for example Dagenais and Dufour (1991) and Stein (1959).

[6] The deviations from expected utility theory are addressed from different angles by Ellsberg (1961), Machina (1987), Tversky, Slovic and Kahneman (1990) and many other.

and the subsequent authors interpreted the Jeffreys' ignorance prior as the prior describing an agent's beliefs about a parameter in the absence of any relevant information (i.e., beliefs in an information vacuum).

Summary of Results

This paper consists of two independent parts investigating decision making under information scarcity from different angles. In Part I utility maximization is not assumed. There, we describe the restrictions on an agent's strategy arising from invariance principles. In contrast, Part II considers a Bayesian utility maximizer and characterizes beliefs consistent with invariance restrictions (see Section 5).

While the approaches taken in Part I and Part II are very different, both lead to the same predictions about an agent's behavior. The following example highlights some of the most important results obtained in this paper. Consider a risk-neutral agent facing a decision problem with a payoff function

$$p(x,\alpha) = \begin{cases} -|\ln x - \ln \alpha| & \text{if } x \in [a,b] \\ 0 & \text{if } x \notin [a,b] \end{cases} . \tag{1.1}$$

Here α denotes the agent's action, and x is the realization of a positive random variable X. Intuitively, the agent's action, α, can be interpreted as a "guess" or the agent's assessment of the realization of a random variable. Obviously, the agent's strategy in such decision problem depends on the beliefs that she has about the conditional probability distribution of X on an interval $[a,b]$. Theorem 4.1 shows that even if an agent has data useful for estimating X, an agent's action in a decision problem with payoff function given by Equation 1.1 is approximated by $\sqrt{ab}$ if the interval $[a,b]$ is far in the tail region. It is interesting that invariance restrictions alone are sufficient to pin down the agent's choices in some decision problems.[7]

In the second part of the paper, we combine expected utility theory and invariance based restrictions on an agent's behavior. This allows us to identify invariant beliefs, i.e., a set of Bayesian beliefs for which invariance axioms are satisfied, where beliefs are defined as a finite-parametric family of density functions and a prior distribution of the parameter.

Under the conditions of Theorem 5.3, the tails of the posterior distribution are approximated by $\frac{c}{x(\ln x)^K}$ where c and K are constants, we refer to this as the

[7]Note that the payoff function in this example belongs to a narrow class of invariant payoff functions, these are introduced in Definition 3.6. The restrictions implied by invariance axioms are the strongest for decision problems characterized by invariant payoff functions.

Tail-distribution.[8,9] That is to say, our findings suggest that within our axiomatic framework an agent's beliefs about the distribution of outliers are approximated by the Tail-distribution. Note the similarity between the Tail-distribution and the Jeffreys' [ignorance] prior; both are characterized by very fat tails. From Proposition B.4, it follows that in the example given earlier expected utility maximization with respect to a Tail-distributed posterior will result in the limit in selection of the same action as the action of an agent who maximizes utility with respect to the Jeffreys' prior.

Part I

2. Setup

This section provides an intuitive explanation of concepts and assumptions we use. (A more formal but less transparent exposition is given in the appendix.)

We will consider decision problems where the payoff depends on the realization of a positive random variable unknown at the time when the action is chosen. Such decision problems can be described by a quadruplet $d = (X, p(x, \alpha), A, I)$ where

X is a positive random variable,

x is the realization of X,

[8] Note that in a Bayesian framework, if a variable is drawn from a distribution that belongs to some family of probability densities, the probability distribution that an agent assigns to this variable need not belong to the family of densities from which the variable is drawn. In fact, we show that an agent who has invariant beliefs will arrive to a posterior distribution of the random variable in question that has no finite moments. This is true even if the agent knows that the variable is drawn from a probability density that belongs to a family of densities such that all moments are finite for each density (this is because, the parameter of the distribution is unknown to the agent, consequently after integrating over distributions with thin tails one may obtain a distribution with fat tails). This point is essential for establishing applicability area of our results and applicability area of empirical studies of extreme values mentioned earlier. We investigate an agent's beliefs about the tails of the distributions and not the distribution of observed outcomes, consequently, the present theory can not be either supported or refuted based on statistical analysis of outliers of economic variables. We make statements about an agent's beliefs that can be tested by observing an agent's actions that are revealing of her beliefs about the relative likelihood of extreme values. For instance, the pricing of far out of the money options contains information about an agent's beliefs relating tails of the distribution of stock returns.

[9] The Jeffreys' ignorance prior is not a proper distribution (i.e. it does not integrate to a constant). In contrast, the integral of the Tail-distribution converges as long as $k > 1$.

$p : R_+ \times R_+ \to R$ is a payoff function,

A is an action space,

$\alpha \in A$ is an action chosen before the value of X is revealed,

I is the information set containing all information relevant for estimating the unknown variable X.

We will consider decision problems where the information set consists of values $x_1, x_2, ...x_N$ of positive real valued variables $X_1, X_2, ..., X_N$. In this case the information set can be written as $I = I_{x_1 x_2 ... x_N}^{X_1 X_2 ... X_N}$. We will also assume that action space A is the set of real numbers.

An agent's strategy in a decision problem is a measure on the action space, and a pure strategy is an element of the action space.

We impose invariance based restrictions on an agent's strategy. Below is the simplified version of Axiom 0, introduced in the appendix.

A0. Consider decision problems $d = (X, p(x,\alpha), A, I = I_{x_1 x_2 ... x_N}^{X_1 X_2 ... X_N})$ *and* $d' = (X, p'(x,\alpha), A, I = I_{x_1 x_2 ... x_N}^{X_1 X_2 ... X_N})$, *action* $\alpha \in R_{++}$ *and* $p'(x,\alpha) = p(x, h(\alpha))$, *where* $h(\alpha)$ *is an invertible function. If an agent's strategy in the first decision problem is* α^* *then her strategy in the second decision problem is given by* α' *such that* $h(\alpha') = \alpha^*$.

A0 requires that the strategy is invariant with respect to relabeling of an action space. Any expected utility maximizer automatically satisfies *A0*.

In order to formulate further invariance restrictions on an agent's behavior the concepts of dimensional vectors and decision equivalent variables need to be introduced.

Definition 2.1. *We will say that two vectors* $(X, X_1, ..., X_N)$ *and* $(Y, Y_1, ..., Y_N)$ *are decision equivalent if for any payoff function* $p(.,.)$ *the agent chooses identical strategies in decision problems* $(X, p(x,\alpha), A, I = I_{x_1 x_2 ... x_N}^{X_1 X_2 ... X_N})$ *and* $(Y, p(x,\alpha), A, I = I_{x_1 x_2 ... x_N}^{Y Y_1 ... Y_N})$.

Since equivalence is defined in terms of strategies of an agent, variables can only be equivalent with respect to a particular agent.

Now let us discuss the notion of a dimensional variable. A dimensional variable represents a quantity measured as a ratio with some arbitrarily selected reference quantity. Only ratios of dimensional variables are meaningful. For instance, length, price etc. are dimensional variables, because they are defined as ratios with some arbitrary unit of measurement.

We can define a dimensional vector $(X, X_1, ..., X_N)$ as a vector consisting of variables measured in the same units such as meters, dollars, kilos etc. (a more formal definition of dimensional vectors is given in the appendix).

Let us define dimensionality compatible transformations g as a function that leaves ratios unchanged with respect to changes in the units of measurement, i.e., for any $x_1, x_2, t > 0$,

$$g(x_1)/g(x_2) = g(\lambda x_1)/g(\lambda x_2).$$

It is straightforward that a function $g(x)$ belongs to the set of dimensionality compatible transformations G if and only if $g(x)$ is given by $g(x) = tx^k$. Consequently, we assume that if $\overrightarrow{X} = (X, X_1, ..., X_N)$ is a dimensional vector, then $\overrightarrow{Y} = (Y, Y_1, ..., Y_N)$, where $\overrightarrow{Y} = (tX^k, ..., tX_N^k)$ is a dimensional vector for any $t > 0$ and $k \in R/\{0\}$. We denote this vector $\overrightarrow{Y} = t(\overrightarrow{X})^k$.

We would like to impose an invariance axiom on an agent's behavior by requiring that an agent's beliefs are purely data driven and the inference rules are the same regardless if the variable in question is the future speed of a microprocessor or the time necessary for performing a benchmark computation. In other words, we assume that the name of the unit of measurement is an uninformative label. More formally, invariance assumption can be stated as follows.

A1. Provided $\overrightarrow{X}$ *is a dimensional vector and* $\overrightarrow{Y} = t(\overrightarrow{X})^k$, $t > 0$, $k \in R$ $k \neq 0$, *then* $\overrightarrow{Y}$ *and* $\overrightarrow{X}$ *are decision equivalent.*

To illustrate the intuition behind this restriction, consider a decision problem where $\overrightarrow{X}$ contains a time series of exchange rate between currencies A and B. Then $\overrightarrow{Y} = (\overrightarrow{X})^{-1}$ contains exchange rates between currencies B and A. According to *A1,* $\overrightarrow{Y}$ and $\overrightarrow{X}$ are decision equivalent i.e. the currency labels are not informative. Restriction *A1* also expresses the requirement that an agent's strategy does not depend on the unit of measurement. (In other words a broker can perform calculations in dollars or in cents, but resulting decision to buy or sell should not depend on the choice of scale of measurement.)

3. Decision Making in an Information Vacuum

If a payoff function is characterized by certain symmetry properties then invariance restrictions stated in the previous section imply strong predictions about an agent's behavior. In this section we identify a class of decision problems where an agent's behavior is determined solely by invariance considerations.

Let us consider a decision problem $d = (X, p^{\varepsilon,a,b}, A, I = I^{X_1 X_2 \ldots X_N}_{x_1 x_2 \ldots x_N})$, where
$X : \Omega \to R^1_{++}$ and $X_i : \Omega \to R^1_{++}$ at each $i \in \{1, \ldots N\}$,
$\overrightarrow{X} = (X, X_1, \ldots, X_N)$ is a dimensional vector,
$A = R^1_{++}$,
and the payoff function $p^{\varepsilon,a,b} : R^1_{++} \times R^1_{++} \to R^1$ is defined by

$$p^{\varepsilon,a,b}(x, \alpha) = \begin{cases} \varepsilon \pi(x, \alpha) & \text{if } x \in [a, b] \\ 0 & \text{if } x \notin [a, b] \end{cases} . \tag{3.1}$$

Here ε, a, b are positive parameters and $\pi(x, \alpha) : R^1_{++} \times R^1_{++} \to R^1$.

We will start with investigating an agent's behavior in an information vacuum, where no data is available. In this case we will denote the information set by $I = \emptyset$.

Definition 3.1. *Define a payoff relevant range as the set of values of X such that an agent's action influences the payoff. (c is in the payoff relevant range if there are two actions α_1 and α_2 such that $p(c,\alpha_1) \neq p(c,\alpha_2)$.)*

The payoff in decision problem d is described by Equation 3.1. The payoff relevant range in decision problem d is $[a, b]$ and the payoffs are proportional to parameter ε.

When the strategy of an agent in decision problem d is pure, the action taken by an agent is denoted by $\alpha = f(a, b, \varepsilon)$. We extend $f(a, b, \varepsilon)$ to R^2_{++} by setting $f(a, b, \varepsilon) \equiv f(b, a, \varepsilon)$.

Let us define an assessment function that approximates the strategy that an agent would choose if the payoffs in all states of the world were multiplied by a small number. To emphasize the connection between the assessment function and the agent's strategy we abuse notation by denoting the assessment function as $f(a, b)$. (If an agent were risk neutral expected utility maximizer, than the assessment function would coincide with the agent's strategy.)

Definition 3.2. *We define the assessment function by setting*
$f(a, b) = \lim_{\varepsilon \to 0} f(a, b, \varepsilon)$ *at each* $(a, b) \in R^2_{++}$.

Note that existence of an assessment function is not guaranteed. An obvious, nevertheless important property of the assessment function is that it remains unchanged when the payoff function of the decision problem is multiplied by a constant.

Considering a limiting case where payoffs converge to zero allows us to separate "beliefs" from risk and uncertainty aversion. The assessment function can be viewed as a partial characterization of an agent's "beliefs".

Let us introduce definitions of a few general symmetry properties, that we will use in further analysis.

Definition 3.3. *We will refer to homogeneous functions of degree one as* 0-*symmetric (i.e. a function is* 0-*symmetric if* $f(ta, tb) = tf(a, b)$*). For* $k \in R\backslash\{0\}$ *a function is* k-*symmetric if* $f(a^k, b^k) = (f(a, b))^k$.

To gain intuitive understanding of k-symmetry consider the following example. If a and b represent the upper and lower bounds on the size of a square and a^2, b^2 are the upper and lower bounds on the area of the square 2-symmetry insures that the assessment of the area is consistent with the assessment of the side of the square. Analogously, if a and b represent distances from the city center, 0-symmetry of an assessment function insures that the assessment is independent from the units of measurement.

Definition 3.4. *A function is symmetric if it is* k-*symmetric for* $\forall\ k \geq 0$ *and for* $k = -1$.

It is easy to check that a symmetric function is k-symmetric also for all negative k.

Definition 3.5. *A function* $\pi : R^2_{++} \to R$ *is* 0-*invariant if for some* $r : R_{++} \to R$ *the equation*

$$\pi(\lambda x, \lambda \alpha) = r(\lambda) \cdot \pi(x, \alpha) \tag{3.2}$$

holds for any $\lambda, x, \alpha \in R_{++}$.

Definition 3.6. *A function* $\pi(.,.)$ *is* t-*invariant if for some* $r : R_{++} \to R$ *the equation*

$$\pi((x)^t, (\alpha)^t) = r(t) \cdot \pi(x, \alpha) \tag{3.3}$$

holds for any $x, \alpha \in R_{++}$ *(here* $t \in R\backslash\{0\}$*). A function is invariant if it is* t-*invariant for* $\forall t \in R^1$.

Definition 3.7. *A decision problem* $d = (X, p^{\varepsilon,a,b}, A, I = I^{X_1X_2...X_N}_{x_1x_2...x_N})$ *is t-invariant if its payoff function is given by Equation 3.1 and the corresponding function* $\pi(x, \alpha)$ *is t-invariant.*

Proposition 3.8. *Consider a t-invariant decision problem* $d = (X, p^{\varepsilon,a,b}, A, I = \emptyset)$. *It follows from A0 and A1 that, if an assessment function* $f(a, b)$ *of a decision problem exists, then* $f(a, b)$ *is t-symmetric.*

All proofs are in the Appendix.

Proposition 3.9. *As long as there exists an assessment function of an invariant decision problem* $d = (X, p^{\varepsilon,a,b}, A, I = \emptyset)$, *it is given by* $f(a, b) = \sqrt{ab}$.

It is easy to verify that a function $-|\ln x - \ln \alpha|$ is an example of an invariant payoff function. A complete characterization of invariant payoff functions is given in Schwarz (1998).

Let us discuss the intuition behind Proposition 3.9. Suppose an agent knows that a variable is dimensional and, aside from that, she has no information relevant for estimating the variable. In this case her actions in decision problems where payoff is contingent on different dimensional variables are the same, i.e., in this case the name of a variable is merely an uninformative label. (Effectively, provided that an agent has never heard of either tugric or dugric her strategy for selecting a guess from an interval [1,4] is the same regardless if she is guessing the value of a "tugric" or the length of a "dugric".) This imposes a sever restriction on an agent's choices. For instance, if an agent is asked to "guess" exchange rate between currencies A and B conditional on the rate being between a and b, a "guess" of $(a + b)/2$ is not "reasonable" because, if this decision problem is reformulated in terms of exchange rate between B and A the range becomes $1/b$ to $1/a$ and the "guess" in the mirror decision problem $(1/b+1/a)/2$ is not a reciprocal of the guess in the original problem. A remarkable property of invariant decision problems is that a strategy in the "image game" must be reciprocal of the strategy in the original game.

Surprisingly, in an information vacuum the invariance consideration along are sufficient to uniquely pin done the strategy of an agent in some decision problems. We showed that if the payoff relevant range in an invariant decision problem is given by $[a, b]$, then an agent's strategy in such a decision problem is approximated by the geometric mean given by[10] $\sqrt{ab}$. Combining the results of this section with expected utility axioms one can show that the prior of an agent in an information vacuum corresponds to the Jeffreys' prior[11] $1/x$, the proof is given in Schwarz (1998).

4. Decision Making in Environments Where Information is Available but Scarce

Up to this point we were only concerned with decision making in an information vacuum. While necessity to make decisions under uncertainty is common place, decision making in an information vacuum may appear to be an unrealistic abstraction. Nevertheless, it is a useful benchmark case because it provides an insight into a wide range of situation where an agent's prior knowledge is very limited. This section establishes that behavior of an agent in environments with "little information" is approximated by an agent's behavior in an information vacuum. The above statement will be made precise later in the paper.

Let us consider a class of decision problems where the information about a payoff relevant variable is available but scarce. We will show that in the limit as information becomes increasingly scarce an agent's strategy approaches her strategy in an information vacuum. Let us analyze an invariant decision problem $d = (X, p^{\varepsilon,a,b}, A, I)$. Unlike the case of an information vacuum considered earlier, we no longer assume that the information set in decision problem d is empty. The information set in d, denoted by $I^{X_1 \ldots X_N}_{x_1 \ldots x_N}$, contains realizations $(x_1 x_2 \ldots x_N)$ of variables $X_1, X_2, \ldots, X_N$ $(X_i : \Omega \to R_{++})$ that are jointly distributed with X, where $\overrightarrow{X} = (X, X_1, X_2, \ldots, X_N)$ is a dimensional vector. Let us assume that the data in the information set fall in the range $[x_{\min}, x_{\max}]$. Information scarcity increases as the payoff relevant range becomes farther away from the data range $[x_{\min}, x_{\max}]$.

If the information set in d were empty then, according to Proposition 3.9 the assessment function would be $\sqrt{ab}$. We will show that in decision problem d characterized by a non empty information set the assessment function converges

[10] We assume that an agent selects a pure strategy in invariant decision problems, which is normally the case for an expected utility maximizer. In order to broaden the compatibility of invariance axioms with non expected utility theories a case of mixed strategies may be considered as well. Schwarz (1998) shows that all restrictions on a pure strategy obtained in this section also apply to the median of a mixed strategy.

[11] Distribution $1/x$ was first proposed as a prior for an unknown parameter in Jeffreys (1931).

to $\sqrt{ab}$ as the payoff relevant range becomes farther away from the data range $[x_{\min}, x_{\max}]$. Consequently, no matter how much data is available, an agent's "beliefs" regarding the values of X very far from the data range can be approximated by beliefs in an information vacuum, where beliefs are manifested by a strategy in invariant decision problems. This claim is made formal by the following proposition.

Theorem 4.1. *Consider an invariant decision problem* $d = (X, p^{\varepsilon,a,b}, A, I = I^{X_1...X_N}_{x_1...x_N})$. *Let us assume that assessment function* $f(x_1...x_N, a, b)$ *exists and continuous, and the limits*

$$\varphi(x_1...x_N, a, b) = \lim_{t\to\infty}(f(x_1...x_N, at, bt)/t),$$

$$\gamma(x_1...x_N, a, b) = \lim_{k\to\infty}\lim_{t\to\infty}\frac{(f(x_1...x_N, (ta)^k, (tb)^k))^{1/k}}{t}$$

exist. Then $\gamma(x_1...x_N, a, b) = \sqrt{ab}$.

Corollary 4.2. *Assume the hypotheses of the Theorem 4.1 are satisfied. If* $\lim_{k\to-\infty}\lim_{t\to 0}(\frac{f(x_1...x_N,(ta)^k,(tb)^k)}{t^k})^{1/k} \equiv \widetilde{\gamma(a,b)}$ *exists, then* $\widetilde{\gamma(a,b)} = \sqrt{ab}$.

The above statement extends the result of Proposition 4.1 for the case where payoff relevant range tends to zero. A comparison of Proposition 3.9 to Theorem 4.1 and its corollary suggests that even when some relevant data is available the beliefs of an agent regarding relative likelihood of outliers are well approximated by beliefs of an agent in an information vacuum.

Now let us derive the restrictions on a prior probability distribution of an agent that behaves in accordance with Theorem 4.1. Let us consider an assessment function $f(x_1...x_N, a, b)$ under conditions of Theorem 4.1. Let us fix $(x_1...x_N)$. Assume that there exists a function $\rho(x)$ such that the assessment function $f(a, b) = f(x_1...x_N, a, b)$ is a maximizer of $\int_a^b \pi(x, \alpha)\rho(x)dx$, where $\pi(x, \alpha)$ is a function of the form $\pi(x, \alpha) = -|\ln x - \ln\alpha|$.

Proposition 4.3. *Assume that the function* $x^s\rho(x)$ *is monotonic with respect to* x *for sufficiently large* x *and* $s \in [1-\delta, 1[\smile]1, 1+\delta]$. *Then for any* $\varepsilon > 0$ *we have*

$x^{1+\varepsilon}\rho(x) \to \infty$ *and* $x^{1-\varepsilon}\rho(x) \to 0$ *as* $x \to \infty$.[12,13]

In the Savage's framework, Proposition 4.3 is a statement regarding the shape of the tails of an agent's prior probability distribution. Provided the conditions of the proposition above are satisfied the best power function estimate of the tails of the distribution is $\frac{1}{x}$. More formally $\frac{1}{x^{1+\varepsilon}}$ converges to zero faster than $\rho(x)$ for any $\varepsilon > 0$. This does not imply that probability distribution of a dimensional variable has to be improper, for instance $\frac{1}{x(\ln x)^2}$ is "best approximated" by $\frac{1}{x}$; however, $\int_c^{+\infty} \frac{1}{x(\ln x)^2} dx < \infty$. In the following section we derive an agent's Bayesian beliefs consistent with invariance. We will explicitly obtain the probability distribution corresponding to the tails of the distribution and we will see that it is indeed consistent with Proposition 4.3.

Part II

5. Comparison with Bayesian Statistics

The results of the earlier sections are of very general nature because their derivation is based solely on invariance restrictions on an agent's behavior. Previous sections do not rely on the axioms of expected utility maximization or existence of a prior probability distribution.[14] We did not have to specify a particular updating procedure in order to obtain our results.

In this section we identify a class of Bayesian beliefs consistent with invariance restriction *A1* introduced in Section 2. Let us consider a decision problem d similar to that of Part I. This decision problem is contingent on a dimensional vector $(X, X_1, ..., X_N)$. The values $x_1...x_N$ of variables $X_1, ..., X_N$ are contained in the agent's information set, and the realization of random variable X is unknown.

[12] If $\frac{\rho(x)}{x^{\varepsilon}}$ does not approach zero as x goes to infinity then $\int_{-\infty}^{+\infty} \rho(x)dx = \infty$. However, the statement that $\frac{\rho(x)}{x^{\varepsilon}} \to 0$ is not entirely superfluous because in this proposition we do not require $\rho(x)$ to be a proper probability density i.e. the conditions of the proposition do not require or imply that $\int_{-\infty}^{+\infty} \rho(x)dx < \infty$.

[13] In Proposition 4.3 we required that the function $x^s\rho(x)$ is monotonic with respect to x for sufficiently large x, possibly, the theorem can be proven without imposing this requirement.

[14] In Part I, we did not use Savage axioms (Savage,1954) or expected utility axioms (von Neumann and Morgenstern, 1944). In fact, in the first part of the paper an agent's strategy was not viewed as a solution to a maximization problem.

Assume that the payoff is equal to $p(x, \alpha)$ if $x \in [a, b]$ and zero otherwise, $\alpha \in R_{++}$ is an action of an agent.

Then, the strategy of an expected utility maximizer is the solution of the following maximization problem:

$$\max \int_a^b F^{[a,b]}_{X|X_1...X_N}(x|x_1...x_N)U(p(x, \alpha))dx.$$

Here $U(.)$ represents an agent's utility for money and $F^{[a,b]}_{X|X_1...X_N}(x|x_1...x_N)$ is a probability distribution of X conditional on $x \in [a, b]$. Let us assume that X and $X_1, ..., X_N$ are independently drawn from the same probability distribution, i.e., there exists a probability density function $\rho(.)$, $\int_0^\infty \rho(x)dx = 1$ such that the joint distribution of $(x, x_1, ..., x_N)$ is $\rho(x)\rho(x_1)...\rho(x_N)$. Beliefs of a Bayesian agent are characterized by a family of densities $M = \{\rho(x; \theta)|\theta \in \Theta\}$ and a prior $W(\theta)$. We will denote beliefs by $B = (M,\ W(\theta))$. Suppose that

$$\int_\Theta \rho(x; \theta)\rho(x_1; \theta)...\rho(x_N; \theta)W(\theta)d\theta < \infty,$$

then a standard Bayesian updating procedure can be applied. The resulting posterior distribution is proportional to

$$F_{X|X_1...X_N;B}(x|x_1...x_N) = \int_\Theta \rho(x; \theta)\rho(x_1; \theta)...\rho(x_N; \theta)W(\theta)d\theta. \tag{5.1}$$

Definition 5.1. *We will say that beliefs $B = (\{\rho(x; \theta)|\theta \in \Theta\}, W(\theta))$ are equivalent to beliefs $B' = (\{\rho'(x; \gamma)|\gamma \in \Gamma\}, W'(\gamma))$ if $\forall$ $(x_1...x_N)$ and $\forall$ $a, b \in R_{++}$ substituting B and B' into Bayes' rule yields*

$$F_{X|X_1...X_N;B}(x|x_1...x_N) = \chi F_{X|X_1...X_N;B'}(x|x_1...x_N), \tag{5.2}$$

where χ is a constant factor.

In other words, beliefs are equivalent if they correspond to the same probability distribution.[15]

[15] As long as a decision maker is concerned, only the probability distribution conditional on the payoff relevant range is of interest. This distribution is given by

$$F^{[a,b]}_{X|X_1...X_N;B}(x|x_1...x_N) = \frac{F_{X|X_1...X_N}(x|x_1...x_N)}{\int_a^b F_{X|X_1...X_N}(x|x_1...x_N)dx} \tag{5.3}$$

Consequently, multiplying $F_{X|X_1...X_N}(x|x_1, ...x_N)$ by a constant does not change the agent's strategy. Also note that since only probability distributions conditional on an interval are relevant here, we might allow $F_{X|X_1...X_N}(x|x_1, ...x_N)$ to represent an infinite measure.

Now, let us consider the requirement on an agent's beliefs implied by the invariance principle. Let $\overrightarrow{X} = (X, X_1, ..., X_N)$ be a dimensional vector. Then, the invariance restriction *A1* implies that $\overrightarrow{X}$ is decision equivalent[16] to $\overrightarrow{X'} = (X', X'_1, ..., X'_N) = (tX^k, tX_1^k, ..., tX_N^k)$ for any $t > 0$, $k \neq 0$. For a Bayesian utility maximizer this condition is satisfied when beliefs B corresponding to dimensional vector $(X, X_1, ..., X_N)$ are equivalent to beliefs B' corresponding to a dimensional vector $(X', X'_1, ..., X'_N)$. Note that $\overrightarrow{X'}$ is a function of $\overrightarrow{X}$. Consequently, we can express beliefs B' in terms of beliefs B about $\overrightarrow{X}$. More formally, we can define an action of group G consisting of transformations $x \mapsto tx^k$ on beliefs B. It is easy to verify that the appropriate transformation of beliefs corresponding to an action of group G is given by $B' = (\{\rho'(x;\theta)|\theta \in \Theta\}, W(\theta))$, where $\rho'(x;\theta) = \frac{1}{kt^{1/k}} x^{\frac{1-k}{k}} \rho((\frac{x}{t})^{1/k};\theta)$. Now we are ready to define invariant beliefs.

Definition 5.2. *Beliefs $B = (\{\rho(x;\theta)|\theta \in \Theta\}, W(\theta))$ are invariant if $\forall\ t > 0$, $k \neq 0$ these beliefs are equivalent to $B' = (\{\rho'(x;\theta)|\theta \in \Theta\}, W(\theta))$, where $\rho'(x;\theta) = \frac{1}{kt^{1/k}} x^{\frac{1-k}{k}} \rho((\frac{x}{t})^{1/k};\theta)$.*

By construction, invariant beliefs are the beliefs that satisfy restriction *A1.*

Now let us identify the set of invariant beliefs. Let us fix a function $h(\xi, \nu)$, where $\xi \in R$, $\nu \in V$ and V is a measure space. Then we can define a family of probability distributions with densities $\rho(x;\mu,\sigma,\nu) = \frac{1}{x\sigma} h(\frac{\ln x - \mu}{\sigma};\nu)$. We will assume that $\int_{-\infty}^{+\infty} \xi^i h(\xi,\nu) d\xi < c$ and $h(\xi,\nu) < c$, where $i \in R_+$ and c is a constant independent of ν.

Theorem 5.3. *Consider a set of beliefs $B^* = (M^*, W^*)$ where*

$$M^* = \{\frac{1}{x\sigma} h(\frac{\ln x - \mu}{\sigma};\nu) | (\mu \in R^1, \sigma \in R^1_+, \nu \in V\}, \tag{5.4}$$
$$W^*(\mu,\sigma,\nu) = T(\nu)/\sigma^s$$

and $\int_V T(\nu) d\nu < \infty$. Then the following is true:
(1) beliefs B^* are invariant
(2) the asymptotic behavior of $F_{X|X_1...X_N;B^*}(x|x_1...x_N)$ as $x \to \infty$ is described by the formula $\frac{const}{x(\ln x)^{N+s}}$ where N is the number of observations and s is the parameter in the expression for $W^*(\mu,\sigma,\nu)$.

[16] Decision equivalence is defined in Section 2. If $\overrightarrow{X'}$ and $\overrightarrow{X}$ are decision equivalent, then an agent's behavior is the same in decision problems $d = (X, p(.,.), A, I_{x_1 x_2 ... x_N}^{X_1 X_2 ... X_N})$ and $d' = (X', p(.,.), A, I_{x_1 x_2 ... x_N}^{X'_1 X'_2 ... X'_N})$.

(Equivalently we can say

$$F_{X|X_1...X_N;B^*}(x|x_1...x_N) = \frac{c(x, x_1, ...x_N)}{x(\ln x)^{N+s}},$$

and $\lim_{x\to\infty} c(xx_1...x_N) = const > 0$, where $const$ is independent of $(x_1...x_N)$.)

Proposition 5.3 established that beliefs given by Equation 5.4 are invariant. Now let us formulate the regularity conditions necessary for proving the converse. First note that our definition of beliefs presumes that the family of density functions depends on a finite number of parameters. Beliefs can alternatively be defined as a measure on a set of densities. The measure theoretic definition of beliefs is more general because it does not require that the measure representing beliefs is concentrated on a family of densities depending on a finite number of parameters.[17] Also note that there is a unique measure on the space of density functions corresponding to beliefs. Let m denote the measure on the set of densities $\{\rho\}$ corresponding to beliefs. We require that the measure on the set of densities induced by beliefs is well behaved in the sense of the definition below. A cylinder set is defined as

$$\mathcal{M}^{x_1,...x_k}_{a_1,b_1,...a_k,b_k} = \{\rho|\rho(x_1) \in [a_1, b_1], ..., \rho(x_k) \in [a_k, b_k])\}. \quad (5.5)$$

Let us denote

$$F_n(x_1...x_n) = \int_{\{\rho\}} \rho(x_1)...\rho(x_n)dm. \quad (5.6)$$

Definition 5.4. *Beliefs B are well behaved if corresponding measure m on set of densities $\{\rho\}$ has the following properties: (1) The $\sigma-$algebra of measurable sets is generated by cylindrical sets (i.e. sets $\mathcal{M}^{x_1,...x_k}_{a_1,b_1,...a_k,b_k}$) and (2) $F_k(x, ...x) \leq A^k k!$ for fixed x.*

Theorem 5.5. *For any well behaved invariant beliefs there exist equivalent beliefs that are characterized by Equation 5.4.*[18]

[17] See Eaton (1983) for a general measure theoretic treatment of invariance-related issues.

[18] Currently I am in the process of rewising the proof of this theorem and making it consise. The sketch of the proof is available from the author upon request. The claim of the theorem is established by applying the results from the classical problem of moments.

6. Concluding Remarks

The results presented in Section 5 characterize the set of invariant beliefs. The beliefs obtained in Part II are consistent with the predictions about an agent's behavior obtained in Part I.[19] For example, Proposition B.4 uses the beliefs identified in Part II of this paper in order to study the invariant decision problem that was considered in Part I. According to Proposition B.4 the limit of the assessment function is $\sqrt{ab}$, the same as it is in Part I. Also, the shape of tails of the posterior distribution established in Theorem 5.3 of Part II is consistent with Proposition 4.3 of Part I.

Note that according to Equation 5.4, the agent's prior is given by $W^*(\mu, \sigma, \nu) = T(\nu)/\sigma^s$; this prior is improper (an improper prior is represented by a density function that does not integrate to a constant). Representing complete ignorance by an improper prior has both theoretical and intuitive appeal. For instance, one may argue that in the absence of any relevant information an agent cannot make judgments about the expected value of a variable. Representing an agent's beliefs regarding a positive parameter by an improper prior ensures that an agent does not assign a positive number to the expected value of the parameter. Nevertheless, updating an improper prior generally results in a proper posterior distribution; consequently, improper priors can be used in Bayesian econometrics.[20]

Also, our characterization of invariant beliefs can be related to the debate on legitimacy of the common prior assumption. (The hypothesis that agents who have identical information should have identical beliefs was advocated by Harsanyi (1967), for a recent perspective on the legitimacy of common prior assumption see Aumann (1998) and Gul (1998).) Of course, if agents with identical information were to make identical inferences, finding the universal inference rule would be a valid and important scientific question. The proposed characterization of invariance beliefs can be viewed as an attempt to find the common prior or a set of priors that a rational agent may use. This characterization also seems to suggest

[19] Additional results establishing consistency of Part I and Part II are reported in Schwarz (1998). Of course, the "remarkable coincidence" between results obtained in two parts of the paper is due to the fact that invariance restriction was imposed in both parts of the paper. Nevertheless, consistency between results of Part I and Part II is not automatic because regularity conditions imposed on agent behavior in the two parts are substantially different.

[20] There are numerous theoretical and empirical studies and monographs utilizing improper prior for estimation problems and obtaining proper posteriors by updating improper priors; see for example Erickson (1989) and Jaynes (1995).

that there is a relatively restricted set of priors equally suitable for use by a rational agent. Within the framework of some utility theories, the behavior of an agent under uncertainty is characterized by multiple priors (Gilboa and Schmeidler (1989)). In this case, uncertainty aversion can be represented by assuming that an agent maximizes her expected utility using most pessimistic of her priors. Potentially, there is a natural match between choice theories involving multiple priors and our result that identifies a limited set of "reasonable" priors.

References

[1] Aumann, R. (1988) "Common Priors: A Reply to Gul," *Econometrica*, Vol. 66, No. 4, pp. 929-38.

[2] Berger J. O. (1985) *Statistical Decision Theory and Bayesian Analysis* 2d edition.

[3] Dahan, E. and H. Mendelson (1998), "Optimal Parallel and Sequential Prototyping in Product Design", Unpublished Manuscript, MIT.

[4] Dagenais, M. and J. Dufour (1991) "Invariance, Nonlinear Models, and Asymptotic Tests" *Econometrica*, Vol. 59, No. 6, pp. 1601-1615.

[5] Eaton, M. (1983) *Multivariate statistics*, Wiley.

[6] Ellsberg, D. (1961) "Risk, Ambiguity, and the Savage Axioms," *The Quarterly Journal of Economics*, Vol. 75, No. 4, pp. 643-669.

[7] Embrechts, P., C. Kluppelberg and T. Mikosch (1997) *Modelling Extremal Events*, Springer.

[8] Erickson, T. (1989) "Proper Posteriors from Improper Priors for an Unidentified Errors-in-Variables Model," *Econometrica*, Vol. 57, No. 6, pp. 1299-1316.

[9] Fama, E. F. (1965), "The Behaviour of Stock Market Prices," *Journal of Business* Vol. 38, pp. 34-105.

[10] Galambos, J. (1978) *The Asymptotic Theory of Extreme order Statistics.* John Wiley and Sons.

[11] Gul, F. (1988) "A Comment on Aumann's Bayesian View," *Econometrica*, v66, n4, pp. 923-27.

[12] Gumbel, E. J. (1958) *Statistics of Extremes.* Columbia University Press.

[13] Harsanyi, J. C. (1967-68) "Games with Incomplete Information Played by Bayesian Players," *Management Science*, 14, pp. 159-183, 320-334, 486-502.

[14] Kearns, P. and A. Pagan (1997) "Estimating the Density Tail Index for Financial Time Series," *Review of Economics and Statistics*, v79 n2, pp. 171-75.

[15] Nash, J. (1950) "The Bargaining Problem," *Econometrica*, Vol. 18, No. 2., pp. 155-162.

[16] Jaynes, E. T. (1995) Probability Theory as Extended Logic, Unpublished Manuscript, University Of Washington.

[17] Jeffreys, H. (1932) "On the Theory of Errors and Least Squares", *Proc. Roy. Soc.*, 138, pp.48-55.

[18] Jeffreys, H. (1961), *Theory of Probability* (3rd ed.). Oxford, UK.: Claredon Press.

[19] Machina, M. (1987) "Choice Under Uncertainty: Problems Solved and Unsolved," *The Journal of Economic Perspectives*, Vol. 1, No. 1, pp. 121-154.

[20] Mandelbrot, M. (1963) "The variation of certain speculative prices," *J. Business* Vol. 36, pp. 394-419.

[21] Savage, L. J. (1954) *The foundations of Statistics*, Wiley, New York.

[22] Schwarz, M. (1998) *Anticipating Rare Events*, manuscript (Ph.D. Dissertation in draft form), Stanford University.

[23] Stanley, M., H. R., et al. (1995) "Zipf Plots and the Size Distribution of Firms," *Economics Letters*, Vol. 49 No. 4, pp. 453-57.

[24] Stein, C. (1959) "An Example of wide discrepancy between fiducial and confidence interval," *Ann. Math. Statist.* 30 pp. 877-880.

[25] Tversky, A., P. Slovic and D. Kahneman (1990) "The Causes of Preference Reversal," *The American Economic Review*, Vol. 80, No. 1., pp. 204-217.

[26] Weitzman, M., L. (1979) "Optimal Search for the Best Alternative," *Econometrica*, Vol. 47, No. 3, pp.641-654.

[27] Zellner, A. (1971) *An introduction to Bayesian Inference in Economics*, Wiley.

[28] Zipf, G. (1950) "A Note on Brand-Names and Related Economic Phenomena," *Econometrica*, Vol. 18, No. 3. pp. 260-263.

A. Formal Definitions

Concepts like variable, decision problem, dimensional vector etc. were defined with various degree of rigor in the body of the paper. This appendix offers more formal definitions of these concepts.

Definition A.1. *A variable is a function on the space of states of the world Ω taking values in some space $B, V : \Omega \to B$.*

Definition A.2. *A decision problem is characterized by a payoff function, an action space, an information set, a payoff space and the space of the states of the world Ω. Here the payoff function is a function from $\Omega \times A$ into the payoff space Π, $\pi : \Omega \times A \to \Pi$. The information set is a subset of Ω, $I \subset \Omega$.*

An information vacuum is symbolically represented by an empty information set i.e. in an information vacuum any state of the world is possible. The information set contains all states of the world that are known to be impossible. A is the action space. Both the space of states of the world and the space of payoffs are the same for all decision problems, consequently, we can denote a decision problem as a triple $d = (\pi, A, I)$.

Let E be a one dimensional vector space. If $e \in E, e \neq 0$,we can define $\mathcal{E}$ as a set of elements of E having the form te where $t \in R_+$. For a fixed e we can construct an isomorphism of E and R^1; then the space $\mathcal{E}$ corresponds to positive semi axis.

We defined above a decision problem where the payoff depends directly on the state of the world and the information set is a subset of the states of the world. Let us define a decision problem where the payoff is contingent on the value of a variable $X : \Omega \to \mathcal{E}$.

Definition A.3. *The payoff is contingent on the value of variable X if the payoff of a decision problem is given by $\pi : \Omega \times A \to \Pi$ and $\exists\ p : \mathcal{E} \times A \to \Pi$ such that for $\forall \alpha \in A$ and $\omega \in \Omega$, $p(X(\omega), \alpha) = \pi(\omega, \alpha)$.*

For an $X-$contingent payoff function π a variable X partitions the space Ω into equivalence classes with respect to the payoff function. We can describe an $X-$contingent decision problem by a quadruplet $d = (X, p, A, I)$. For conciseness we refer to both payoff function π and "an X-contingent payoff function" p as payoff functions.

Similarly, we can say that an information set is contingent on a vector containing the values $x_1 x_2 ... x_N$ of variables $(X_1 X_2 ... X_N)$. An information set denoted by $I^{X_1 X_2 ... X_N}_{x_1 x_2 ... x_N} \subset \Omega$ is given by

$$I^{X_1X_2...X_N}_{x_1x_2...x_N} = \{\omega \in \Omega | X_1(\omega) \neq x_1 or X_2(\omega) \neq x_2 ... or X_N(\omega) \neq x_N\}$$

.

Definition A.4. *If a decision problem is contingent on a variable X and the information set is contingent on the values of variables $(X_1X_2...X_N)$ we will say that a decision problem is contingent on a vector of variables $\overrightarrow{X} = (XX_1X_2...X_N)$. A decision problem contingent on a vector $\overrightarrow{X}$ is denoted by $d = (X, p, A, I^{X_1X_2...X_N}_{x_1x_2...x_N})$.*

Note that any decision problem where payoff is contingent on a variable vector $\overrightarrow{X'}$ can be represented as a decision problem where payoff is contingent on a vector $\overrightarrow{X}$ that takes values in R^1_+. Thus, without loss of generality we can consider only dimensional variables taking values in R^1_+.

Definition A.5. *A strategy in a decision problem $d = (\pi, A, I)$ is a measure μ_d on A taking values in space $\mathcal{E}$.*

Definition A.6. *An agent is a function from the space of decision problems D into strategies on D i.e. For $\forall d \in D$, μ_d denotes a strategy.*

In this framework an information set is "external" to an agent i.e. the information set is a part of a decision problem and not an attribute of an agent. When we say that an agent knows certain data, or makes decisions in the information vacuum we refer to an information set of the decision problem that an agent is facing.

Definition A.7. *A strategy is pure if it is represented by a measure concentrated on one element of A.*

Axiom 0: Consider two decision problems $d = (\pi, A, I)$, $d' = (\pi', A', I)$, such that $\exists$ a bijective measurable map $h : A \to A'$ and for $\forall\ \alpha\ \forall\ \omega$ the following is true $\pi(\omega, \alpha) = \pi'(\omega, h(\alpha))$. Denote an agents strategies in d and d' as measures μ_d and $\mu_{d'}$. Then the strategy of an agent are connected by the following formula: for any measurable set $N \subset A$, and its image under bijection $N' = h(N) \subset A'$ we have $\mu_d(N) = \mu_{d'}(N')$.

This axiom states that for every strategy in d there exists a strategy in d' such that payoffs from these strategies are the same, and vice versa, given an agent's strategy μ_d in d her strategy $\mu_{d'}$ in d' is such that the payoffs under μ_d and $\mu_{d'}$ are the same in every state of the world.

Definition A.8. *Two variables vectors are decision equivalent if for* $\forall\ A$*,* $\forall\ p$ *the strategies in* $(X, p, A, I_{x_1x_2...x_N}^{X_1X_2...X_N})$ *and* $(Y, p, A, I_{y_1y_2...y_N}^{Y_1Y_2...Y_N})$ *are the same.*

When two variables are equivalent relative to $I = \emptyset$ we will say that these variables are equivalent. Note that variables vectors equivalent relative to one agent need not be equivalent relative to another agent.

Definition: A function g from $\mathcal{E}$ into $\mathcal{E}$ is dimensionality compatible if for $\forall\ e_1$ and $e_2 \in \mathcal{E}$ and $e_2 \neq 0$ and for $\forall\ t > 0, t \in R$, the following is true

$$\frac{g(te_1)}{g(te_2)} = \frac{g(e_1)}{g(e_2)} \tag{A.1}$$

The set of all dimensionality compatible functions is denoted $\mathcal{G}$. Applying a dimensionality compatible transformation to a variable vector $\overrightarrow{X}$ we obtain a vector $g\overrightarrow{X} \equiv (gX, gX_1gX_2...gX_N)$.

Definition A.9. *A vector of variables* $(XX_1...X_N)$ *is dimensional if for any decision problem* $(X, p, A, I_{x_1x_2...x_N}^{X_1X_2...X_N})$ *contingent on a vector* $(XX_1...X_N)$ *and for* $\forall g_1 \in \mathcal{G}$ *and* $g_2 \in \mathcal{G}$ *decision problem* $d' = (g_1X, p, A, I_{x_1x_2...x_N}^{g_1X_1g_1X_2...g_1X_N})$ *is equivalent to a decision problem* $d'' = (g_2X, p, A, I_{x_1x_2...x_N}^{g_2X_1g_2X_2...g_2X_N})$*.*[21]

From Equation A.1, it is straightforward to show that $g \in \mathcal{G}$ if and only if $\exists\ k \in R$ $e \in E, e \neq 0$ such that for any X the following holds $gX = (\frac{X}{e})^k$. Consequently, the following assumption is embedded in the definition of a dimensional vector.

A1. A dimensional vector $\overrightarrow{X}$ is decision equivalent to $\overrightarrow{Y} = t(\overrightarrow{X})^k$ for $\forall\ t > 0$, $\forall\ k \in R\backslash\{0\}$ (where $t(\overrightarrow{X})^k$ denotes $(tX^ktX_1^k...tX_N^k)$).

[21] According to this definition a dimensional vector variable is a subjective category, because decision problems are equivalent relative to a particular agent, and there is no guarantee that quantities that are dimensional in a physical sense, such as price, length, exchange rate etc., correspond to dimensional variables in a sense of definition proposed here. However, all physical dimensional quantities are dimensional in the sense of our definition of dimensional vectors if for any two physical dimensional quantities represented by variables an agent's strategy is purely data driven i.e. the strategy is chosen by a fixed formula based on the values of observations contained in the information set.

B. Proofs

Proof of Proposition 3.8. Since $\overrightarrow{X} = (X)$ is a dimensional vector then from A1 follows that Xt and X^t are decision equivalent to X. Let $\pi(.,.)$ be a t-invariant payoff function (the connection between $\pi(.,.)$ and $p^{\varepsilon,a,b}(.,.)$ is given by Equation 3.1). Consider two decision problems $d = (X, p^{\varepsilon,a,b}, A, I = \emptyset)$ and $d' = (Y, p^{\varepsilon,a^t,b^t}, A, I = \emptyset)$, where $Y = X^t$. The payoff relevant range in decision problem d is $[a, b]$, the payoff relevant range in decision problem d' is $[a^t, b^t]$. The values of assessment function in these decision problems are $\alpha_1 = f(a, b)$ and $\alpha_2 = f(a^t, b^t)$ respectively. If we substitute $y = x^t$ into the second decision problem the payoff function remains the same up to a multiplication by a constant: $\pi(x^t, \alpha) = g(t)\pi(x, \alpha^{1/t})$. Note that the assessment function remains unchanged when the payoff function is multiplied by a constant. Consequently, applying A0 we get $\alpha_2^{1/t} = f(a, b) \Rightarrow f(a^t, b^t) = f(a, b)^t$. The proof for 0-symmetric payoff functions is analogous.□

The following Corollary follows immediately from Proposition 3.8.

Corollary B.1. *Provided the conditions of Proposition 3.8 are satisfied the assessment function of an invariant game obeys the following conditions:*

1) $f(ta, tb) = tf(a, b)$ for $\forall t \in R^1_{++}$
2)$f(a^k, b^k) = (f(a, b))^k$ for $\forall k \in R^1$.

The proof of Proposition 3.9 follows immediately from Corollary B.1 and the following lemma.

Lemma B.2. *If a function $f(a, b)$ is k-symmetric for every $k \geq 0$ and $a \leq b$, then $f(a, b) = a^{(1-\alpha)}b^\alpha$ for some α. If a function $f(a, b)$ is symmetric, then $f(a, b) = \sqrt{ab}$.*

Proof of Lemma:
Define

$$g(x, y) = \ln(f(e^x, e^y)). \tag{B.1}$$

From 0-symmetry follows that

$$g(x + \alpha, y + \alpha) = g(x, y) + \alpha \tag{B.2}$$

Indeed, $g(x + \alpha, y + \alpha)$=$\ln(f(e^{x+\alpha}, e^{y+\alpha}))$=$\ln(f(e^x, e^y))$+$\alpha$ for every α
Let $\alpha = -y$. Then,

$$g(x,y) = g(x-y,0) - y = h(x-y) - y \text{ where } h(x-y) = g(x-y,0). \quad \text{(B.3)}$$

Now let us show that t-symmetry implies

$$g(tx,ty) = tg(x,y). \quad \text{(B.4)}$$

Indeed, $g(tx,ty)$=$\ln(f(e^{xt},e^{yt})) = \ln((f(e^x,e^y))^t) = tg(x,y)$.

Combining Equations B.3 and B.4 we obtain

$$g(tx,ty) = h((x-y)t) + yt = t(h(x-y)+y) \Longrightarrow h(t) = t \cdot constant_1 \quad \text{(B.5)}$$

Substituting Equation B.5 into Equation B.3 we obtain $g(x,y) = x \cdot k + y(1-k)$. Substituting the result back into Equation B.1 we obtain

$f(a,b) = a^{(1-\alpha)}b^{\alpha}$ for $a < b$.

From the above equation and $f(a,b) = f(b,a)$ we obtain

$$f(a,b) = (\min\{a,b\})^{(1-\alpha)}(\max\{a,b\})^{\alpha}.$$

If $f(.,.)$ satisfies (-1)-symmetry then $f(a,b) = a^{(1-\alpha)}b^{\alpha} = f(b^{-1},a^{-1})^{-1} = a^{\alpha}b^{(1-\alpha)}$, thus $f(a,b) = \sqrt{ab}$.□

We will use the result of Lemma B.2 several times later in the paper[22].

Proof of proposition 3.9. From Corollary B.1 and the above lemma follows the claim of Proposition 3.9.□

Proof of Theorem 4.1. Let us emphasize that the limit in the definition of $\gamma(x_1...x_N,a,b)$ is considered as double limit. However, it follows from our conditions that it can be replaced with any of repeated limits.

First we note that an assessment function of an invariant game contingent on a dimensional vector is k-symmetric for any k, i.e., $(f(x_1...x_N,a,b))^k = f(x_1^k...x_N^k,a^k,b^k)$ and for $k=0$, $f(tx_1...tx_N,ta,tb) = tf(x_1...x_N,a,b)$. (The proof of k-symmetry of $f(x_1...x_N,a,b)$ is analogous to that in Proposition 3.8.) The following lemma completes the proof of Proposition 4.1.

[22] The geometric mean is also a natural focal point in a bargaining game where an outcome is a point on an interval $[a,b]$ and the ideal point of one agent is a, while b is the ideal point for of the other agent.. For instance, a possible application of assessment function of the form $\sqrt{ab}$ is in reconciliation of contradictions in a contract. Suppose a contract between a client and a builder specifies that a room should be 5 m by 5 m i.e. 25 m^2, and somewhere else in the same contract the same room is specified to have area 30 m^2; in this case “meeting in the middle” would yield different results depending if meters or meters squared are used to find the middle, while the geometric average insures unit invariance.

Lemma B.3. *Consider a continuous function* $f(x_1...x_N, a, b)$ *such that:*
1)$f(tx_1...tx_N, ta, tb) = tf(x_1...x_N, a, b)$ *and* $f(x_1^k...x_N^k, a^k, b^k) = (f(x_1...x_N, a, b))^k$
2)*the limits* $\varphi(x_1...x_N, a, b) = \lim_{t\to\infty} \frac{f(x_1...x_N, at, bt)}{t}$ *and*

$$\gamma(x_1...x_N, a, b) = \lim_{k\to\infty\, t\to\infty} \frac{(f(x_1...x_N, (ta)^k, (tb)^k))^{1/k}}{t} \tag{B.6}$$

exist. Then, the following is true:

A. $\gamma(x_1...x_N, a, b) = \sqrt{ab}$,

B. There exists $\psi^*(a, b)$ *such that* $\forall$ $(x_1, ...x_N) \in R^N_{++}$, $\psi(x_1...x_N, a, b) = \psi^*(a, b)$, *where* $\psi(x_1...x_N, a, b)$ *denotes* $\lim_{k\to\infty}(f(x_1...x_N, a^k, b^k)^{1/k}$.

Proof. To prove part B of Lemma, we notice that

$$\psi(x_1...x_N, a, b) = \lim_{k\to\infty}(f(x_1...x_N, a^k, b^k)^{1/k}\{= \lim_{k\to\infty}(f((x_1^{1/k})^k...(x_N^{1/k})^k, a^k, b^k))^{1/k}\} =$$

$$\lim_{k\to\infty}(f((x_1^{1/k})...(x_N^{1/k}), a, b) = f(1, ..., 1, a, b).$$

(We used continuity of the function $f(.,.)$ and the fact that $x^{1/k}$ tends to one as k tends to infinity.) We obtain B with $\psi^*(a, b) = f(1, ..., 1, a, b)$.

To prove part A of the lemma we notice that the double limit in Equation B.6 is equal to repeated limit. Taking the limit with respect to k, first we obtain

$$\gamma(x_1...x_N, a, b) = \lim_{t\to\infty}\lim_{k\to\infty} \frac{(f(x_1...x_N, (ta)^k, (tb)^k))^{1/k}}{t} = \lim_{t\to\infty} \frac{\psi^*(ta, tb)}{t}. \tag{B.7}$$

Consequently, $\gamma(x_1...x_N, a, b)$ is independent of x_i, thus writing $\gamma(a, b) = \gamma(x_1...x_N, a, b)$, we obtain

$$\gamma(\tau a, \tau b) = \lim_{t\to\infty} \frac{\psi^*(\tau ta, \tau tb)}{t} = \tau \lim_{t\to\infty} \frac{\psi^*(\tau ta, \tau tb)}{\tau t} = \tau\gamma(a, b).$$

Now we can calculate the repeated limit for γ by first taking the limit for $t \to \infty$ and then $k \to \infty$. We obtain

$\gamma(x_1...x_N, a, b) = \lim_{k\to\infty}\lim_{t\to\infty}(\frac{f(x_1...x_N, (ta)^k, (tb)^k)}{t^k})^{1/k} = \lim_{k\to\infty}(\varphi(x_1...x_N, a^k, b^k))^{1/k}$

Thus, $\gamma(a^l, b^l) = \gamma(x_1^l...x_N^l, a^l, b^l) = \gamma(x_1...x_N, a^l, b^l) = \lim_{k\to\infty}(\varphi(x_1...x_N, a^{lk}, b^{lk}))^{1/k} = \lim_{k\to\infty}((\varphi(x_1...x_N, a^{lk}, b^{lk}))^{1/lk})^l = (\lim_{k\to\infty}(\varphi(x_1...x_N, a^{lk}, b^{lk}))^{1/lk})^l = (\gamma(a, b))^l$

We have shown that $\gamma(a, b)$ is k-symmetric for any $k \geq 0$, thus we can apply Lemma B.2 to obtain $\gamma(a, b) = a^\alpha b^{1-\alpha}$ for $a < b$. Assessment function $f(x_1...x_N, a, b)$

is symmetric with respect to interchanging a, b so is $\gamma(a,b)$,also $f(x_1...x_N, a, b)$ and $\gamma(a,b)$ are -1-symmetric. Therefore, $\gamma(a,b) = a^{\alpha}b^{1-\alpha} = a^{1-\alpha}b^{\alpha} = \sqrt{ab}\square$

The statement of Corollary 4.2 can be derived easily from Theorem 4.1 or proved independently along the same lines (we omit the proof).

Proof of Proposition 4.3:

The idea of the proof is as follows. It is easy to check that if under the conditions of this proposition $\rho(x)$ were asymptotically equal to x^{-s} then necessarily s would be equal to one. The other key idea is that the assessment function increases when $\rho(x)$ is replaced with $\rho(x)x^{\varepsilon}, \varepsilon > 0$.

Define $g(x;s) \equiv x^s\rho(x)$. According to the conditions of the proposition either $\frac{dg(x;s)}{dx} \geq 0$ for all $x > x^*$or $\frac{dg(x;s)}{dx} \leq 0$ for all $x > x^*$. Consequently, there are three possible cases

1) $\exists\delta > 0$ and $x^* > 0$ such that $\forall s \in [1-\delta, 1+\delta]$ and all $x > x^*$, $\frac{dg(x;s)}{dx} \leq 0$.

2) $\exists\delta > 0$ and $x^* > 0$ such that for $\forall s \in [1-\delta, 1+\delta]$ and all $x > x^*$, $\frac{dg(x;s)}{dx} \geq 0$.

3)$\exists\delta > 0$ and $x^* > 0$ such that for $\forall s \in [1-\delta, 1[$ and all $x > x^*$, $\frac{dg(x;s)}{dx} \leq 0$ and for $\forall s \in]1, 1+\delta]$ and all $x > x^*$, $\frac{dg(x;s)}{dx} \geq 0$. (It is easy to show that this list of cases is complete.) Let us prove that cases 1) and 2) can be ruled out.

Step 1. Denote $f^{\rho}(a,b) \equiv f^{\rho}(x_1...x_N, a, b)$, where $f^{\rho}(x_1...x_N, a, b)$ is an assessment function corresponding to $\rho(x) = g(x)s(x)$ i.e.

$$f^{\rho}(a,b) = \arg\max_{\alpha} \int_a^b \pi(x,\alpha)\rho(x)dx$$

and

$$f^{s}(a,b) = \arg\max_{\alpha} \int_a^b \pi(x,\alpha)s(x)dx,$$

where $s(x) = \frac{\rho(x)}{g(x;s)} = x^{-s}$. Let us show that if $\exists\, x^*$ such that for any $x > x^*$, $g'(x) \leq 0$ and $s'(x) \leq 0$ then $f^{\rho}(a,b) \leq f^{s}(a,b)$ for any $a, b > x^*$.

Consider an invariant payoff function $\pi(x,\alpha) = -|\ln\frac{x}{\alpha}|$. Applying first order conditions we obtain

$$\int_a^{f^{\rho}(a,b)} g(x)s(x)dx = \int_{f^{\rho}(a,b)}^b g(x)s(x)dx \qquad \text{(B.8)}$$

and

$$\int_a^{f^s(a,b)} s(x)dx = \int_{f^s(a,b)}^b s(x)dx \tag{B.9}$$

Let us suppose $a > x^*$, $b > x^*$ and $g'(x) \leq 0$ for $x \in [a, b]$ then

$$\int_a^{f^s(a,b)} g(x)s(x)dx \geq g(f^s(a,b)) \int_a^{f^s(a,b)} s(x)dx$$

and

$$\int_{f^s(a,b)}^b g(x)s(x)dx \leq g(f^s(a,b)) \int_{f^s(a,b)}^b s(x)dx.$$

Consequently we have,

$$\int_a^{f^s(a,b)} g(x)s(x)dx \geq \int_{f^s(a,b)}^b g(x)s(x)dx$$

which implies $f^s(a,b) \geq f^\rho(a,b)$. Thus,

$$\lim_{t\to\infty} \lim_{k\to\infty} \frac{(f^s((ta)^k,(tb)^k))^{1/k}}{t} \geq \lim_{t\to\infty} \lim_{k\to\infty} \frac{(f^\rho((ta)^k,(tb)^k))^{1/k}}{t}$$

Step 2. Denote $\gamma^s(a,b) \equiv \lim_{t\to\infty} \lim_{k\to\infty} \frac{(f^s((ta)^k,(tb)^k))^{1/k}}{t}$ and $\gamma^\rho(a,b) \equiv \lim_{t\to\infty} \lim_{k\to\infty} \frac{(f^\rho((ta)^k,(tb)^k))^{1/k}}{t}$. According to Proposition 4.1 $\gamma^\rho(a,b) = \sqrt{ab}$. In step 1 we established that $\gamma^s(a,b) \geq \gamma^\rho(a,b)$ if $g'(x) \leq 0$. Let us compute $\gamma^s(a,b)$ for $s(x) = \frac{1}{x^s}$. Substituting $s(x)$ into B.9 we get

$$\int_{a^k}^{\alpha^k} x^{-s}dx = \int_{\alpha^k}^{b^k} x^{-s}dx.$$

Solving for α and taking the limit we obtain $\alpha(k) = (\frac{b^{k(-s+1)}+a^{k(-s+1)}}{2})^{1/k(-s+1)}$ for $s \neq 1$ and $\ln\alpha = \frac{\ln b + \ln a}{2}$, $\alpha = \sqrt{ab}$ for $s=1$. For $s > 1$ $\gamma^s(a,b) = a$, for $s < 1$ $\gamma^s(a,b) = b$ and for $s = 1$ $\gamma^s(a,b) = \sqrt{ab}$. Note that $\gamma^s(a,b) \geq \gamma^\rho(a,b)$. In case 1) there are some $s > 1$ such that $g'(x;s) \leq 0$, hence case 1) can be ruled out. Using an analogous argument case 2) can be ruled out as well. Thus, we have shown that $\exists\delta > 0$ and $x^* > 0$ such that for each $s \in [1-\delta, 1[$, $\frac{dg(x;s)}{dx} \leq 0$ and for each $s \in]1, 1+\delta]$, $\frac{dg(x;s)}{dx} \geq 0$. However, by construction $g(x;s) = x^{s-1}g(x;1)$. Consequently, for any $\varepsilon > 0$ the following is true: $x^\varepsilon \cdot g(x;1)_{x\to+\infty} \to +\infty$ and $\frac{g(x;1)}{x^\varepsilon} \to 0$.□

Proof of Theorem 5.3:

1) First denote the action of a group G on R^1_+ as g^X_{tk}, $g^X_{tk}x = tx^k$. An action of this group on the space of probability density functions is denoted

$$g^\rho_{tk}, g^\rho_{tk}r(x) = r((\frac{x}{t})^{1/k})\frac{1}{kt^{1/k}}x^{\frac{1-k}{k}}.$$

(If $r(x)$ is a density of probability distribution of variable X, then $(g^\rho_{tk}r)(x)$ or simply $g^\rho_{tk}r(x)$ is a density of variable tX^k.) We will say that a set M is invariant, or closed under an action of the group G if for $\forall\ \rho(x) \in M$ for $\forall\ t \in R_{++}$, $k \in R$, $g^\rho_{tk}\rho(x) \in M$. Let us show that M^* is closed under G. Indeed, $(g^\rho_{tk}\rho)(x) = \frac{1}{\sigma(x/t)^{1/k}}h(\frac{\ln(\frac{x}{t})^{1/k}-\mu}{\sigma};\nu)\frac{1}{kt^{1/k}}x^{\frac{1-k}{k}}$ and therefore for any $\rho(x;\mu,\sigma,\nu) = \frac{1}{\sigma x}h(\frac{\ln x-\mu}{\sigma};\nu)$, the following is true $g^\rho_{tk}\rho(x;\mu,\sigma,\nu) \in M^*$ and $g^\rho_{tk}\rho(x;\mu,\sigma,\nu) = \rho(x;\mu k+\ln t, k\sigma,\nu)$. Also note that we can define an action of the group on the parameter space as $g^\theta_{tk}(\mu,\sigma,\nu) = (\mu k+\ln t, k\sigma,\nu)$.

Combining Equation 5.1 and Equation 5.2 we obtain

$$\int_\Theta \rho(x;\theta)\rho(x_1;\theta)...\rho(x_N;\theta)W(\theta)d\theta = k\int_\Theta g^\rho_{tk}\rho(x;\theta)g^\rho_{tk}\rho(x_1;\theta)...g^\rho_{tk}\rho(x_N;\theta)W(\theta)d\theta. \tag{B.10}$$

We have to check that the above equality holds for

$$\rho(x;\mu,\sigma,\nu) = \frac{1}{\sigma x}h(\frac{\ln x-\mu}{\sigma};\nu) \in M^*.$$

The right hand side of Equation B.10 is proportional to

$\int_{-\infty}^{+\infty}\int_0^{+\infty}\int_V g^\rho_{tk}\rho(x;\mu,\sigma,\nu)g^\rho_{tk}\rho(x_1;\mu,\sigma,\nu)...g^\rho_{tk}\rho(x_N;\mu,\sigma,\nu)\frac{T(\nu)}{\sigma^s}d\theta =$

$$\int_{-\infty}^{+\infty}\int_0^{+\infty}\int_V \rho(x;\mu k+\ln t, k\sigma,\nu)\rho(x_1;\mu k+\ln t, k\sigma,\nu)...\rho(x_N;\mu k+\ln t, k\sigma,\nu)\frac{T(\nu)}{\sigma^s}d\mu d\sigma d\nu \tag{B.11}$$

Now we can change variables taking $\mu' = \mu k+\ln t$, $\sigma' = k\sigma$ and $\nu' = \nu$. The Jacobean of this transformation is

$J = \det\begin{vmatrix} 1/k & 0 & 0 \\ 0 & 1/k & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1/k^2$. After transforming the variables in Equation B.11 we obtain

$$\int_{-\infty}^{+\infty}\int_0^{+\infty}\int_V \rho(x;\mu',\sigma',\nu')\rho(x_1;\mu',\sigma',\nu')...\rho(x_N;\mu',\sigma',\nu')\frac{k^{s-2}T(\nu')}{\sigma^s}d\mu' d\sigma' d\nu'.$$

This expression coincides with the left-hand side of Equation B.10 (up to a factor).

Proof of 2) Let us consider

$F_{X|X_1...X_N}(x|x_1...x_N) =$

$$\int_{-\infty}^{+\infty}\int_{0}^{+\infty}\int_{V}\frac{1}{\sigma^{N+1}x^{N+1}}h(\frac{\ln x-\mu}{\sigma};\nu)h(\frac{\ln x_1-\mu}{\sigma};\nu)...h(\frac{\ln x_N-\mu}{\sigma};\nu)\frac{T(\nu)}{\sigma^s}d\mu d\sigma d\nu. \tag{B.12}$$

Introducing notations $y = \ln x$, $y_i = \ln x_i$ we obtain from this expression

$$F_Y(y) = \int_V\int_{-\infty}^{+\infty}\int_{0}^{+\infty} h(\frac{y-\mu}{\sigma};\nu)h(\frac{y_1-\mu}{\sigma};\nu)...h(\frac{y_N-\mu}{\sigma};\nu)\frac{T(\nu)}{\sigma^{s+N+1}}d\mu d\sigma d\nu, \tag{B.13}$$

where c is a constant factor. Let us introduce notations $y = zK$, $y_i = z_iK$, $\mu = mK$, $\sigma = lK$. Then

$$F_Z(z) = \int_V\int_{-\infty}^{+\infty}\int_{0}^{+\infty} h(\frac{z-m}{l};\nu)h(\frac{z_1-m}{l};\nu)...h(\frac{z_N-m}{l};\nu)\frac{T(\nu)}{K^{s+N-1}\cdot l^{S+N+1}}dmdld\nu \tag{B.14}$$

Now we change variables in this integral replacing m, l, ν with u, v, ν, where $u = \frac{z-m}{l}$, $v = \frac{z_1-m}{l}$. Expressing everything in terms of new variables we compute the Jacobian of the transformation

$$J = \det\begin{vmatrix} -\frac{z-z_1}{(u-v)^2} & \frac{z-z_1}{(u-v)^2} & 0 \\ \frac{z_1(u-v)-(uz_1-vz)}{(u-v)^2} & \frac{-z(u-v)+(uz_1-vz)}{(u-v)^2} & 0 \\ 0 & 0 & 1 \end{vmatrix} = \frac{(z-z_1)^2}{(u-v)^3}.$$

We see that

$F_Z(z) = \frac{1}{((z-z_1)K)^{s+N-1}}\times$

$$\int_V\int_{-\infty}^{+\infty}\int_{v}^{+\infty} h(u;\nu)h(v;\nu)\Pi_{i=2}^{N}h(\frac{z_iu-z_iv-uz_1+vz}{z-z_1};\nu)(u-v)^{s+N-2}T(\nu)dudvd\nu \tag{B.15}$$

Now we will make use of the fact that $z_i/z \longrightarrow 0$. According to Lebesgue dominated convergence theorem we can take a limit inside an integral if there exists a summable function that dominates the integrand.

The function $h(v, \nu)$ is bounded, consequently for any u and v

$$h(u;\nu)h(v;\nu)\Pi_{i=2}^{N}h(\frac{z_i u - z_i v - uz_1 + vz}{z - z_1};\nu)(u{-}v)^{s+N-2}T(\nu) < h(u;\nu)h(v;\nu)(u{-}v)^{s+N-2}T(\nu)A$$

where A is a constant. Also, $\int_{-\infty}^{+\infty}\int_{v}^{+\infty} h(u;\nu)h(v;\nu)(u-v)^{s+N-2}dudv$ is finite because it is bounded by $\sum_{i=1}^{s+N-2}\sum_{j=1}^{s+N-2} Q_i Q_j \omega_{ij}$, where ω_{ij} are some coefficients and $Q_i{=}\int_{-\infty}^{+\infty} h(v;\nu)v^i dv < \infty$, Q_i's are moments of $h(u, v)$ that are finite by assumption of the theorem.

Taking the limit inside the integral in Equation B.15 we obtain that $F_Z(z){\cdot}(zK)^{s+N-1}$ tends to a constant as $z \longrightarrow \infty$. Consequently, the asymptotic behavior of F is given by

$$F_Y(y)_{y\to\infty} \to \frac{const}{y^{s+N-1}}.$$

Finally, $y = \ln x$, thus $F_X(x)_{x\to\infty} \to \frac{const}{x(\ln x)^{s+N-1}}$. □

Proposition B.4. *Denote an agent's strategy by $f(a, b, x_1...x_N)$ where $x_1...x_N$ is the data available to the agent in an invariant decision problem characterized by a payoff function of the form*

$$p(x,\alpha) = \begin{cases} -|\ln x - \ln\alpha| & x \in [a,b] \\ 0 & otherwise \end{cases}.$$

Assume that all the conditions of Theorem 5.3 hold and that either $a < b < 1$ or $1 < a < b$. Then the $\lim_{t\to\infty}(f(x_1...x_N, at, bt)/t)$ exists and it equals to $\sqrt{ab}$.[23]

Proof of Proposition B.4. Let us find

$$\lim_{t\to\infty}(f(x_1...x_N, at, bt)/t)$$

corresponding to a distribution given by

$$F_X(x) = \frac{c(x, x_1, ...x_N)}{x(\ln x)^{N+s}}.$$

[23] The claim of this proposition probably remains true if some of the assumptions imposed in Proposition 5.3 are dropped. For instance, the assumption that all moments of function $h(.,.)$ exist is probably redundant in the context of this theorem, however, in the interests of simplicity we formulate and prove Proposition B.4 under assumptions imposed in Proposition 5.3.

Since $\lim_{x\to\infty} c(x, x_1...x_N) = const$, where $const$ is independent of $(x_1...x_N)$, the

$$\lim_{t\to\infty}(f(x_1...x_N, at, bt)/t)$$

is the same for densities $F_X(x)$ and $F'_X(x) = \frac{1}{x(\ln x)^{N+s}}$. The first order condition corresponding to the invariant payoff function and the probability density function $F_X(x)$ is given by

$$\int_{at}^{\alpha(t)t} \frac{1}{x(\ln x)^{N+s}} dx = \int_{\alpha(t)t}^{bt} \frac{1}{x(\ln x)^{N+s}} dx.$$

Differentiating both sides of the first order condition with respect to t yields

$$2(\ln \alpha t)^J - (\ln at)^J = (\ln bt)^J, \text{ where } J = -N - s + 1.$$

Rearranging terms and taking the limit we obtain

$$\lim_{t\to\infty} \ln \alpha(t) = \lim\left(\frac{((\ln bt)^J + \ln(at)^J)^{1/J}}{2^{1/J}} - \ln t\right) =$$

$$\lim((1 - \frac{s(\ln b + \ln a)}{2\ln t})^{1/J} \ln t - \ln t) = \frac{\ln b - \ln a}{2}.$$

Thus, $\alpha = f(a, b) = \sqrt{ab}\square$

According to Theorem 5.3 the tails of the distribution are proportional to $\frac{1}{x|\ln x|^C}$. Note that among all power functions, this distribution is best approximated by $\frac{1}{x}$ and this is consistent with the statement of Proposition 4.3, which was derived from general invariance considerations. Also, Proposition B.4 is consistent with the claim of Proposition B.6; while the assumptions behind these two propositions are different, both agree that

$$\gamma(x_1...x_N, a, b) = \lim_{k\to\infty} \lim_{t\to\infty} \frac{(f(x_1...x_N, (ta)^k, (tb)^k))^{1/k}}{t} = \sqrt{ab}. \qquad \text{(B.16)}$$

Alternatively, one can prove that $\gamma(x_1...x_N, a, b) = \sqrt{ab}$, utilizing Proposition B.6 and showing that necessary limits exist and continuous.