A THEORY OF ATTRIBUTION

Barry E. Feldman[*]

MAY 29, 2007
VERSION 1.1

ABSTRACT

Attribution of economic joint effects is achieved with a random order model of their relative importance. *Random order consistency* and elementary axioms uniquely identify *linear* and *proportional marginal attribution*. These are the Shapley (1953) and proportional (Feldman (1999, 2002) and Ortmann (2000)) values of the dual of the implied cooperative game. Random order consistency does not use a reduced game. *Restricted potentials* facilitate identification of proportional value derivatives and coalition formation results. Attributions of econometric model performance, using data from Fair (1978), show stability across models. Proportional marginal attribution (PMA) is found to correctly identify factor relative importance and to have a role in model construction. A portfolio attribution example illuminates basic issues regarding utility attribution and demonstrates investment applications. PMA is also shown to mitigate concerns (*e.g.*, Thomas (1977)) regarding strategic behavior induced by linear cost attribution.

JEL Classifications: C10, C52, C71, D00, G11, M31 and M41.

Keywords: Attribution, coalition formation, consistency, cost allocation, joint costs, joint effects, proportional value, random order model, relative importance, restricted potential, Shapley value and variance decomposition.

What is the relative importance of factors influencing teenage pregnancy or global warming? How should corporate costs or profits be allocated? And what is the relative importance of assets in an investment portfolio?

This paper presents an economic theory of attribution capable of addressing these questions that is based on the general question: "What is the probability that any order of model factors is correctly ordered by their relative importance?" Elementary axioms identify two probability distributions over factor orders. In *proportional marginal attribution* (PMA), a factor's expected contribution share is equal the probability it is most important. In *linear attribution* (LA) all orders are equally likely.

This work is based on and contributes to cooperative game theory, a mathematical approach to the study of bargaining. Cooperative game theory provides models for the representation of the joint effects inherent in bargaining and methods for their attribution. The simplest n-person bargaining game provides a sturdy but limited model of bargaining, yet a perfect structure for the attribution problems studied here.

von Neumann (1928/1952) first defines the n-person cooperative game. The Shapley (1953) value is today the central solution concept. A versatile axiomatically defined mathematical structure, it is also the equilibrium outcome of noncooperative bargaining games such as Gull (1988) and Hart and Mas-Colell (1996). It and its generalization to games with a continuum of players, the Aumann-Shapley (1974) value, have been applied to diverse problems, including cost allocation (*e.g.*, Shubik (1962) and Young (1985)), general equilibrium (*e.g.*, Shapley (1964)), and price setting (*e.g.*, Billera and Heath (1982)). Roth (1988) and Hart (2006) provide reviews.

Shapley (1953) shows his value represents the expected marginal contributions of players to a growing coalition when all orders of "arrival" are equally likely. Weber (1988) characterizes the set of all linear random order values. A significant literature has developed (*e.g.*, Khmelnitskya (1999) and Segal (2003)). Linearity, however, prevents the worths of coalitions from influencing the random order arrival process.

Consistency in cooperative games has been based on special games that "reduce" players from an initial game. If a solution concept is *consistent*, allocations to the remaining players are unchanged. Hart and Mas-Colell (1988) define a reduced game and use it to characterize the Shapley value with consistency. Feldman (1999, 2002) shows parallel results for the proportional value using the same reduced game.

Random order consistency requires that an *inert* player or factor, one having no joint effects with others, not disturb the probabilities of suborders in which it is *not* included. The random order approach to consistency does not require a reduced game. This facilitates practical application. Random order consistency is shown to provide a simple and powerful expectations-based approach to cooperative value.

The proportional value (Ortmann (2000) and Feldman (1999, 2002)) satisfies a proportional gain rule instead of the balanced (equal) contributions rule of Myerson (1980) and is defined by a ratio potential rather than the linear Hart and Mas-Colell (1988) potential. Feldman (2002) shows the proportional value induces a random

order process influenced by coalitional worths and that it is the equilibrium outcome of a noncooperative bargaining game where players' probabilities of proposing are proportional to their expected payoff. Khmelnitskya and Driessen (2001) and Calvo and Driessen (2003) study generalizations.

Section 1 of this paper presents a microeconomic model of relative importance based on set marginal contributions in a random order model. Random order consistency and other axioms are used to directly identify a probability distribution over orders. The Shapley and proportional values are identified in a common framework, without assuming linearity or proportionality, though the change of a single axiom.

Section 2 defines *restricted* potentials and obtains proportional value derivatives and coalition formation results. The proportional value is shown not converse monotonic. A family of values including the Shapley and proportional values is shown.

$$* \qquad * \qquad *$$

Attribution involves a single monotonic utility function. Players in the corresponding bargaining problem must be assumed to have identical linear utility functions. In attribution, factors and factor sets replace players and coalitions. The worth of a factor set is based on the utility resulting from the presence of its factors or the restriction of utility maximization to the variation of its factors. Three examples demonstrate the potential of this theory of attribution.

Statistical tests ignore joint effects by design. Linear attribution of OLS R^2 as a measure of relative importance is proposed by Hoffman (1960), Lindeman, *et al.* (1980), Kruskal (1987) and Chevan and Sutherland (1991). McNally (2000) initiates its use in ecology studies. Soofi, *et al.* (2000) develop a model of linear attribution in categorical analysis. Linear attribution of model R^2 allows a factor to appear important due only to correlations with factors in the true model. Feldman (2005a) first proposes an *exclusion* axiom which requires that such a factor receive zero attribution share, and with other axioms, characterizes proportional marginal attribution of model R^2 (there called *proportional marginal variance decomposition* or *PMVD*). Grömping (2006, 2007) provides attribution tools and studies OLS LA and PMA.

Section 3 studies econometric attribution. The model objective function is considered to be the utility function of an analyst. Random order consistency is found compatible with econometric attribution. Attributions are found stable and comparable across a variety models based on data from Fair (1977). PMA identifies factor relative importance and model relationships missed by joint significance tests.

Section 4 considers utility attribution in microeconomic models more generally. Random order consistency is found to be an acceptable axiom whenever an attribution problem can be composed with an independent problem. A portfolio optimization example illustrates utility attribution in a typical microeconomic context. The example provides methods for evaluating the relative importance of assets in an investor's portfolio and the per-dollar contributions of assets to investor utility.

Section 5 considers cost allocation as an attribution problem. Random order consistency is found to be an appropriate axiom for cost attribution. An example illustrates the relative vulnerability of LA to strategic definition of cost centers.

The conclusion summarizes, compares random order and reduced game consistency and considers the common assumption that proportional solutions are translation dependent in the light of results of developed here. Extended proofs follow.

1 Basic results

1.1 General framework

Transferrable utility (TU) games represent the *worth* of a coalition by a single number. Let $N = \{1, 2, \ldots, n\}$ be the *players* or *factors*. A standard TU game v is a map $v : 2^N \to \mathbb{R}_+ \cup 0$ from all $S \subseteq N$ to the nonnegative real numbers, with $v(\varnothing) \equiv 0$. The $S \subseteq N$ are *coalitions* or *factor sets*. Let $\bar{i} = \{i\}$ and $\overline{12} = \{1, 2\}$. Set subtraction of $\bar{i}$ from S is written $S \setminus \bar{i}$. Games are assumed *weakly monotonic*: $v(S \cup \bar{i}) \geq v(S)$.

In an attribution problem, $v(S)$ is the expected utility of a single decision maker. If S represents binary characteristics, then $v(S)$ is the utility resulting from the joint presence of the factors.[1] Factor set S may instead provide the decision maker with decision space Δ_S, with $\sigma_S \in \Delta_S$ being a possible choice. Then $v(S)$ is the maximum

$$v(S) = \max_{\sigma_S \in \Delta_S} U(\sigma_S) - U(\varnothing), \tag{1}$$

where the *null utility* $U(\varnothing)$ represents the utility obtained when there are no optimization choices. The null utility is the *basis point* of Feldman (2005b). It is clearly needed with exponential and other utility functions that can generate negative utility. Its importance in other circumstances will become apparent.

The weak monotonicity of v corresponds to the assumption that maximization over a greater number of factors cannot result in a decline in expected utility.

The *dual game* to v represents the marginal contributions of all coalitions or factor sets. The dual worth $w(S)$ is $v(N)$ minus the worth of the complement of S:

$$w(S) = v(N) - v(N \setminus S). \tag{2}$$

1.2 The random order model

Attribution is based on a random order model of relative importance. A discrete marginality replaces the more typical economic focus on continuous marginal conditions. In this model factors "arrive," one at time, to join a growing factor set. Factors

---

[1]A *monotonic cover game*, where $v^*(S) = max_{R \subset S} v(R)$, can be used to ensure monotonicity.

could be indexed by arrival order. However, due to the central role of joint marginal contributions, it is simpler to consider the "last to arrive" as the first in an order. The last to arrive can instead be thought of as "first to leave."

An *order* $r = (r_1, r_2, \ldots, r_n)$ is a permutation of the factors (or players) of N. The set of all orders of N is $\mathcal{R}(N)$. Define $S_k^r = \{r_1, \ldots, r_k\}$ to be the set (coalition) of the first k factors (players) in the order r. These are the *last* k factors to arrive in order r. If S has k factors, S is *included* in r, written $S \in r$, if and only if $S = S_k^r$. Thus $w(S_i^r)$ is the joint marginal contribution of the last i factors to enter. Let $r(i)$ be the position of i in order r, so that $r_{r(i)} = i$. When factors are ordered according to relative importance, r_1 is *least* important and r_n is *most* important.

Let r^S be an order of the factors of a set $S \subsetneq N$ and consider an $r \in \mathcal{R}(N)$. Assume that for any i and j in S such that $r^S(i) > r^S(j)$, it is also true that $r(i) > r(j)$. The order of the factors in r^S is preserved in r, so r^S is a *suborder* of r, written $r^S \subset r$. Also, r is a *superorder* of r^S, written $r \supset r^S$.

Let $m^P(r)$ be an n-vector of *positional marginal* contributions defined by r. Element i, $m_i^P(r)$, represents the marginal contribution of the factor in position i in r. This quantity is defined as follows with respect to the dual game:

$$m_i^P(r) = w(S_i^r) - w(S_{i-1}^r), \qquad\qquad i = 1, 2, \ldots, n, \qquad\qquad (3)$$

where $S_0^r = \varnothing$ and $w(\varnothing) = 0$. Let $m^S(r)$ be the n-vector of *set marginal* contributions:

$$m_i^S(r) = w(S_i^r) = v(N) - v(N \setminus S_i^r(r)), \qquad\qquad i = 1, 2, \ldots, n. \qquad\qquad (4)$$

The probability $p(r)$ that any order $r \in \mathcal{R}(N)$ is the realized order of arrivals is defined by a likelihood function $L(r)$ with $L(r) \geq 0$ for all r:

$$p(r^*) = \frac{L(r^*)}{\displaystyle\sum_{r \in \mathcal{R}(N)} L(r)}. \qquad\qquad (5)$$

The expected utility contribution of factor i with respect to distribution p is then

$$\phi_i(v) = E_p[m_{r(i)}^P(r)] = \sum_{r \in \mathcal{R}(N)} p(r)\, m_{r(i)}^P(r). \qquad\qquad (6)$$

1.3   Axioms and the fundamental theorem

Axioms are used to identify probability distributions over orderings and not the value function itself. Such a distribution may be understood to represent the probability that any order is correctly ordered by the relative importance of its factors.

A factor z is *inert* if and only if it adds exactly its own non-zero worth to every coalition: $v(\bar{z}) > 0$ and, for any $S \subseteq N \setminus \bar{z}$, $v(S \cup \bar{z}) = v(S) + v(\bar{z})$. Let v^* be the game created by adding a set of inert factors Z, so that $N^* = N \cup Z$.

Axiom 1.1 RANDOM ORDER CONSISTENCY: $p\left(r \,|\, v\right) = \displaystyle\sum_{\substack{r^* \in \mathcal{R}(N^*) \\ r^* \supset r}} p\left(r^* \,|\, v^*\right).$

Random order consistency requires that an inert factor or factor set have no effect on the probability of any suborder of the remaining factors. Consider an order r in $\mathcal{R}(N)$. Then the sum of the probabilities of all orders $r^* \in \mathcal{R}(N^*)$ such that $r \subset r^*$ relative to the game v^* must equal the probability of r relative to v.

If the axiom holds for a likelihood L in games of cardinalities m and n, then L is random order consistent between these cardinalities. If L is random order consistent, then it is random order consistent between all cardinalities m and n, $2 \le m < n$.

Compared to Hart and Mas-Colell (1988) reduced game consistency, random order consistency requires the existence of inert factors, but does not require a reduced game or directly constrain value allocation when inert factors are not present.

Axiom 1.2 EXCLUSION: *If $m_1^S(r) = 0$ when $r(z) = 1$ and $m_1^S(r^*) > 0$ when $r^*(j) = 1$, for all $j \neq z$, then for any r^* such that $r^*(z) > 1$, $p\left(r^*\right) = 0$.*

Assume z makes zero *final marginal contribution* (arriving last in any order r) and all other factors make positive final marginal contributions. Exclusion then requires that only orders where z arrives last are assigned positive probability.

The basic idea of exclusion is that a factor making zero final marginal contribution should receive no attribution share. This outcome, however, is not demanded when there are two or more such factors (*cf.*, Feldman (2002), Section 6.1).

Exclusion will considered in the limit as marginal contributions grow small. Let v_0 be strictly monotonic. Define a sequence of games $v_1, v_2, \ldots, v_\infty$ where $v_t(S) = v_0(S)$ for $S \neq N \setminus \bar{z}$ and $\lim_{t \to \infty} v_t(N \setminus \bar{z}) = v_0(N)$. If $r(z) = 1$ then $\lim_{t \to \infty} m_1^S(r \,|\, v_t) = 0$. Exclusion requires that if $r^*(z) > 1$, then $\lim_{t \to \infty} p\left(r^* \,|\, v_t\right) = 0$.

Axiom 1.3 SEPARABILITY: $L(r) = \sum l_i(m_i^S(r))$ *or* $L(r) = \prod l_i(m_i^S(r))$.

Separability requires the likelihood function be a sum or product of subliklihoods. Separability limits complexity but does not imply separability of p or ϕ.

Axiom 1.4 ANONYMITY: *If $m^S(r^*) = m^S(r)$, then $L(r^*) = L(r)$.*

Anonymity requires $L(r)$ depend only on the marginal contributions of the factor sets included in r.

Axiom 1.5 INCLUSION: *If $m^S_{r(i)}(r) > 0$, then there is an r^* such that $p(r^*) > 0$ and $m^P_{r^*(i)}(r^*) > 0$.*

Inclusion requires that if a factor makes a positive positional contribution in any order r that there must be an order r^* with positive probability in which the factor also makes a positive positional contribution. It could be that $r^* = r$.

The following result is proved in the next two subsections.

Theorem 1.1 (The Fundamental Theorem of Attribution) *Let v be a game representing an attribution or bargaining problem, and let w be its dual. Require that the value attributed to factors or players equal the their expected marginal contribution resulting from a probability distribution over orders in the random order model of relative importance.*

(i) *The Shapley and proportional values of w, corresponding to linear and proportional marginal attribution, are the only random order consistent, anonymous and separable values.*

(ii) *The Shapley value of v is the unique value that additionally satisfies inclusion.*

(iii) *The proportional value of w uniquely additionally satisfies exclusion.*

Remark 1.1 *Random order consistency is uniquely associated with cooperative value. The reduced game approach only establishes consistency with respect to a particular two-player solution. Complete characterization requires specific axioms for the two-player case (e.g., Hart and Mas-Colell (1988), Theorem B'). Many solution concepts are consistent with respect to the Hart and Mas-Colell reduced game, including equal split, average cost pricing, serial cost sharing (Moulin and Shenker, 1992), fixed proportions, path methods (Friedman, 2004), dictatorial and priority rules (see Leroux (2006) for further discussion). An additional complication of the reduced game approach is still other solutions are consistent with respect to other reduced games.*

1.4 Relative importance likelihood functions

A likelihood L is *exogenous* if it is independent of v, *i.e.*, $\partial L(r)/\partial v(S) = 0$ for all $r \in \mathcal{R}(N)$ and $S \subseteq N$. If L is not exogenous, it is *endogenous*. Random order consistency, anonymity and inclusion identify a unique exogenous likelihood. Random order consistency, separability and exclusion identify a unique endogenous likelihood.

Lemma 1.1 *The unique (up to scaling) endogenous separable likelihood function that satisfies random order consistency is*

$$L^\diamond(r) = \left(\prod_{S \in r} w(S) \right)^{-1}.$$

Further, $L^\diamond$ is anonymous and positive.

Proof: See Appendix A (Sec. 7.1).

Lemma 1.2 *The likelihood function $L^*(r) = c > 0$ with $p(r)=1/n!$ is uniquely identified (up to scaling) by anonymity, inclusion and random order consistency. It also formally satisfies both additively and multiplicative separability.*

Proof: $L^*(r)$ is anonymous and $p(r) = 1/n!$ for any $L^* = c > 0$. It is inclusive since for all r, $p(r) > 0$. It is random order consistent since for any r_0 in $N \setminus \bar{z}$, where z is inert, $\sum_{r \supset r_0} p(r) = n/n! = 1/(n-1)!$. Inclusion is necessary since $L^\diamond$ satisfies anonymity and random order consistency. Random order consistency is necessary because many likelihoods satisfy anonymity and inclusion, *e.g.*, $L^o(r) = \sum_{S \in r} w(S_i^r)$. It is sufficient to show the necessity of anonymity when $n \leq 3$: Let $l_i^\omega(S_i^r) = \omega_{r(i)}/\sum_{j=1}^i \omega_{r(j)}$ and $L^\omega(r) = \prod l_i^\omega(S_i^r)$, where $\omega_i > 0$, $i = 1, 2, 3$, are exogenous *weights*. L^ω is inclusive. It is easy to determine that $L^\omega((1,2,3)) + L^\omega((1,3,2)) + L^\omega((3,1,2)) = L^\omega((1,2))$. (This is the weight system of the weighted Shapley value, *cf., e.g.*, Kalai and Samet (1987).) Thus, L^ω is random order consistent for $m = 2$ and $n = 3$ and anonymity is necessary. Let $l_i^*(S) = c^{1/n}$, $L^*(r) = \prod_{i=1}^n l_i^*(S) = c$ and L^* is multiplicatively separable. Let $l_i^* = c/n$, $L^*(r) = \sum_{i=1}^n l_i^*(S) = c$ and L^* is additively separable. $\quad\square$

Lemma 1.3 *$L^\diamond$ is the unique likelihood function to satisfy separability, random order consistency and exclusion.*

Proof: L^* cannot satisfy exclusion as $p(r) > 0$ for all $r \in \mathcal{R}(N)$. $L^\diamond$ satisfies exclusion. To see, consider an dual attribution problem w_0 where the final marginal contribution of z is very small and a sequence of games $\{w_t\}_{t=1}^\infty$, where $\lim_{t\to\infty} w_t(\bar{z}) = 0$; but for all other sets $S \subseteq N$, $w_t(S) = w_0(S) > 0$. For any $r^\Omega \in \mathcal{R}(N)$ such that $r_1^\Omega = z$, $\lim_{t\to\infty} L_t^\diamond(r^\Omega) = +\infty$, however, clearly $p_t^\diamond(r^\Omega) = L_t^\diamond(r^\Omega)/\sum_{r\in\mathcal{R}(N)} L_t^\diamond(r) \leq 1$. Further, for any r^* with $r_1^* \neq z$, $\lim_{i\to\infty} p_t^\diamond(r^*) = L_t^\diamond(r^*)/\sum_{r\in\mathcal{R}(N)} L_t^\diamond(r) = 0$ since $\lim_{t\to\infty} L^\diamond(r^*)/L^\diamond(r^\Omega) = 0$. A non-anonymous exogenous L could assign an $r \in \mathcal{R}(N)$ zero probability, but exclusion would not be satisfied. $\quad\square$

Remark 1.2 *Note that all separable likelihood functions with $l_i(\bullet) = l_j(\bullet)$ are random order consistent when likelihoods are based on positional marginal contribution.*

1.5 Attribution and value

The Shapley values of v and its dual w are equal and are the expectation when L^* is used in (6). Thus Lemma 1.2 proves

Theorem 1.2 *Let the likelihood of the distribution over orders in the random order model of relative importance satisfy random order consistency, anonymity and inclusion. Then the resulting attribution is equal to the Shapley value of the problem.*

Remark 1.3 *Weighted Shapley (1953) values might be identified by random order consistency and inclusion (as suggested in the proof of Lemma 1.2).*

Define the function $P(S)$ for any $S \subseteq N$ as

$$P(S) = \left(\sum_{r \in \mathcal{R}(S)} L^\diamond(r) \right)^{-1}. \tag{7}$$

$P(N)$ is then the normalizing factor that relates $L^\diamond$ to the implied distribution $p^\diamond$:

$$p^\diamond(r) = P(N)L^\diamond(r). \tag{8}$$

Feldman (2002), Lemma 2.1 shows that formula (7) is one form of the *ratio potential* of a cooperative game and that $P(S)$ is also recursively defined by the formula

$$P(S) = w(S) \left(\sum_{i \in S} P(S \setminus \bar{i})^{-1} \right)^{-1}. \tag{9}$$

Substitution of formula (8) into the expectation (6) gives

$$\varphi_i(w) = P(N) \sum_{r \in \mathcal{R}(N)} L^\diamond(r) m^P_{r(i)}(r). \tag{10}$$

Lemma 1.4 $\qquad P(N \setminus \bar{i}) = \left(\sum_{r \in \mathcal{R}(N)} L^\diamond(r)\, m^P_{r(i)}(r) \right)^{-1}.$

This is proved in Appendix B (Sec. 7.2) and by Feldman (2002), Lemma 2.9.

Corollary 1.1 $\qquad \varphi_i(w) = \dfrac{P(N)}{P(N \setminus \bar{i})}.$

The discrete derivative of the ratio potential with respect to i is one definition of i's proportional value (Feldman (1999) equation (3.6), Ortmann (2000) Definition 2.2). Thus Lemma 1.1, expectation (10) and Lemma 1.4 prove

Theorem 1.3 *Let the likelihood of the distribution over orders in the random order model of relative importance satisfy random order consistency, exclusion and separability. Then the attribution of v is equal to the proportional value of w, its dual.*

2 Properties of the proportional value

2.1 Elementary properties

Random order consistency directly implies the following lemma, where $\varphi(v, S)$ is the proportional value of v when v is limited to the factors in S.

Lemma 2.1 *If z is inert then $\varphi_i(v, S \setminus \bar{z}) = \varphi_i(v, S)$.*

Now let $N = \{1, 2\}$. Then $P(\bar{i}) = w(S_i^{ij}) = w(\bar{i}) = v(\overline{12}) - v(\bar{j})$ and

$$P(\overline{12}) = \frac{w(\overline{12})}{P(\bar{1})^{-1} + P(\bar{2})^{-1}} = \frac{w(\overline{12})}{w(\bar{1})^{-1} + w(\bar{2})^{-1}} = \frac{w(\bar{1})w(\bar{2})w(\overline{12})}{w(\bar{1}) + w(\bar{2})},$$

so that

$$\varphi_i(w) = \frac{P(N)}{P(N \setminus \bar{i})} = \frac{w(\bar{i})}{w(\bar{1}) + w(\bar{2})} \, w(\overline{12}). \tag{11}$$

Myerson (1980) shows the Shapley value is defined by *balanced contributions*:

$$Sh_i(S, v) - Sh_i(S \setminus \bar{j}, v) = Sh_j(S, v) - Sh_j(S \setminus \bar{i}, v).$$

Adding j to S helps i in the same measure that adding i helps j. Ortmann (2000), Theorem 2.6 characterizes the proportional value with an analogous *ratio perserving* condition (called *equal proportional gain* in Feldman (1999 and 2002)):

$$\frac{\varphi_i(S, v)}{\varphi_i(S \setminus \bar{j}, v)} = \frac{\varphi_j(S, v)}{\varphi_j(S \setminus \bar{i}, v)}. \tag{12}$$

Lemma 2.2 *Let v be a cooperative game with n players and require $\hat{c} > 0$.*

(i) *Let $v^*(S) = v(S)$ if $|S| \neq k$ for a $k < n$ and $v^*(S) = v(S) + c$ when $|S| = k$. Then $Sh(v^*) = Sh(v)$.*

(ii) *Let $v^*(S) = v(S)$ if $|S| \neq k$ for a $k < n$ and $v^*(S) = \hat{c}\, v(S)$ when $|S| = k$. Then $\varphi(v^*) = \varphi(v)$.*

(iii) *Let $v^*(S) = v(S)$ if $S \neq N$ and $v^*(N) = v(N) + n\, c$. Then $Sh(v^*) = Sh(v) + c$.*

(iv) *Let $v^*(S) = v(S)$ if $S \neq N$ and $v^*(N) = \hat{c}\, v(N)$, v^* still monotonic. Then $\varphi(v^*) = \hat{c}\, \varphi(v)$.*

Proof: The Shapley results are obvious. Proportional value results follow directly from formulas (7) or (9) and Corollary 1.1. $\qquad\square$

Corollary 2.1 *The proportional value is the unique random order consistent and separable value in the random order model of relative importance that (a) is invariant to a proportional change in the worth of all coalitions of any cardinality $s < n$ or (b) scales with changes in $v(N)$.*

2.2 The probability of being most important

Lemma 2.3 *The probability that factor i is most important in a proportional marginal attribution is $\varphi_i(w)/w(N)$.*

Proof: The probability of i being most important in a PMA implies arriving first in the random order model of relative importance. This also implies being in position n and that worths are relative to the dual game w. Thus

$$
\begin{aligned}
p\left(r_n = i \mid w\right) &= P(N, w) \sum_{\substack{r \in \mathcal{R}(N) \\ r: r_n = i}} \prod_{i=1}^{n} w(S_i^r)^{-1} \\
&= \frac{P(N, w)}{w(N) P(N \setminus \bar{i}, w)} = \frac{\varphi_i(w)}{w(N)}.
\end{aligned}
$$

The first line above follows from formula (8), The substitution on the second line follows from factoring out $w(N)$ and applying Lemma (1.4). $\qquad\square$

2.3 Your enemy's enemy may be your friend

Feldman (1999), Lemma 3.5, shows the proportional value is monotonic. It is shown here the proportional value is not *converse monotonic*. The value of i can increase with an increase in the worth of S even if $i \notin S$.

The *restriction* of a game v by a coalition S^* includes only coalitions $T \supsetneq S^*$ that include S^* as a proper subset. Let $RP(S^*, S)$ be the *restricted potential* for S in v restricted by S^*. Define $RP(S^*, S^*) = 1$ and define RP as in the following lemma.

Lemma 2.4 *For any $T : S^* \subsetneq T \subseteq N$,*

$$
RP(S^*, T) = v(T) \left(\sum_{i \in T \setminus S^*} RP(S^*, T \setminus \bar{i})^{-1} \right)^{-1} = \left(\sum_{\substack{r \in \mathcal{R}(T) \\ S^* = S_s^r}} \prod_{\substack{S \in r \\ S \supseteq S^*}} v(S)^{-1} \right)^{-1}.
$$

The first part of the equivalence is the definition. The proof of the second equivalence is by recursive substitution, as in Feldman (2002), Lemma 2.1. The value of a player $i \notin S^*$ in the game v restricted by S^* is $\varphi_i(S^*, N, v) = RP(S^*, N)/RP(S^*, N \setminus \bar{i})$.

Lemma 2.5 *The derivative of $\varphi_i(v)$ with the worth of any coalition $S^* \subset N$ is*

$$\frac{\partial \varphi_i(v)}{\partial v(S^*)} = \begin{cases} \varphi_i(v) \dfrac{P(N)}{v(S^*)P(S^*)RP(S^*, N)}, & i \in S^*, \\[2ex] \varphi_i(v) \dfrac{1}{v(S^*)P(S^*)} \left(\dfrac{P(N)}{RP(S^*, N)} - \dfrac{P(N \setminus \bar{i})}{RP(S^*, N \setminus \bar{i})} \right), & i \notin S^*. \end{cases}$$

Proof: The result follows from the value of $\partial P(S)/\partial v(S^*)$ for any $S : S^* \subsetneq S \subseteq N$.

$$\begin{aligned} \frac{\partial P(S)}{\partial v(S^*)} &= [P(S)]^2 v(S^*)^{-1} \sum_{r \ni v(S^*)} \prod_{S \in r} v(S)^{-1} \\ &= [P(S)]^2 \Big[v(S^*)P(S^*)RP(S^*, N) \Big]^{-1} \end{aligned}$$

The first result is from straightforward differentiation. The second result follows as the sum of products in the first is equal to $[P(S^*)RP(S^*, S)]^{-1}$ for $S \supsetneq S^*$. $\qquad \square$

It follows that $\partial \varphi_i(v)/\partial v(N) = \varphi_i(v)/v(N)$ (thus $\sum_{i \in N} \partial \varphi_i(v)/\partial v(N) = 1$) and that $\partial \varphi_i(v)/\partial v(S^*) > 0$ for an $i \in S^*$. On the other hand, the sign of $\partial \varphi_i(v)/\partial v(S^*)$ is not clear for an $i \notin S^*$. Multiplying the LHS and RHS results of Lemma 2.5 by $RP(S^*, N)/P(N \setminus \bar{i})$ shows $\sum_{i \in N} \partial \varphi_i(v)/\partial v(S^*) = 0$ and the following.

Corollary 2.2 *For any $i \in N \setminus S^*$,* $\dfrac{\partial \varphi_i(v)}{\partial v(S^*)} \propto \varphi_i(v) - \varphi_i(S^*, N, v)$.

If i's value in v restricted by S^* is less than in v, then φ is not converse monotonic. This seems possible since formula (11) shows that adding a $c > 0$ to all worths can reduce the value of a dominant player. The Shapley and weighted Shapley values are converse monotonic of their linearity.

Table 1 presents an example that demonstrates the proportional value is not converse monotonic. In v, 1 is weakest and 3 is strongest. 2 is strong in combination with 3, but less so in combination with 1. The proportional value (φ), the Shapley value (Sh) and the nucleolus (Nuc) all provide qualitatively similar allocations.

The game v^* is a version of v modified only by increasing 1's individual worth to 10. The proportional value is not converse monotonic because 3's value is larger than in v. Further, player 1's share of value and dominance over the next largest player have both increased. This appears a plausible outcome in economic, political and military competition. The Shapley value shows similar allocations before and after, with 2 still stronger than 1. The nucleolus (Schmeidler, 1969) does not change.

$$
\begin{array}{lll}
v(\bar{1}) = 1, & v(\bar{2}) = 2, & v(\bar{3}) = 25, \\
v(\overline{12}) = 20, & v(\overline{13}) = 30, & v(\overline{23}) = 60, \\
& v(\overline{123}) = 100 &
\end{array}
$$

$$
v^*(S) = v(S), S \neq \bar{1}, \quad v^*(\bar{1}) = 10
$$

$$
\begin{array}{lll}
\varphi_1(v) = 7.58, & \varphi_2(v) = 29.21, & \varphi_3(v) = 63.20 \\
\varphi_1(v^*) = 20.61, & \varphi_2(v^*) = 10.69, & \varphi_3(v^*) = 68.70 \\
\Delta\varphi_1 = 13.03, & \Delta\varphi_2 = -18.52, & \Delta\varphi_3 = 5.50 \\
\end{array}
$$

$$
\begin{array}{lll}
Sh_1(v) = 17.50, & Sh_2(v) = 33.00, & Sh_3(v) = 49.50 \\
Sh_1(v^*) = 20.50, & Sh_2(v^*) = 31.50, & Sh_3(v^*) = 48.00 \\
\Delta Sh_1 = 3.00, & \Delta Sh_2 = -1.50, & \Delta Sh_3 = -1.50 \\
\end{array}
$$

$$
\begin{array}{lll}
Nuc_1(v) = 12.00, & Nuc_2(v) = 36.00, & Nuc_3(v) = 52.00 \\
Nuc_1(v^*) = 12.00, & Nuc_2(v^*) = 36.00, & Nuc_3(v^*) = 52.00 \\
\Delta Nuc_1 = 0.00, & \Delta Nuc_2 = 0.00, & \Delta Nuc_3 = 0.00 \\
\end{array}
$$

Table 1: Your enemy's enemy may be your friend.

Corollary 2.3 *The Shapley value is the unique random order consistent, anonymous, separable and converse monotonic value.*

2.4 Coalition formation

Simple coalition formation results are implied by Sections 2.2 and 2.3. The coalition S forms if its members all arrive before any members of $N \setminus S$. Lemma 2.3 implies the probability that i is most important and j is second most important in the random order model of relative importance is $\varphi_i(N, w)/w(N) \times \varphi_j(N \setminus \bar{i}, w)/w(N \setminus \bar{i})$, or, in terms of potentials, $P(N, w)/P(N \setminus \overline{ij}, w) \times [w(N)w(N \setminus \bar{i})]^{-1}$. When the probability that S forms is summed over all such orders, the sum of worth products is the restricted potential $RP(S, N, w)$. Results change in a standard random order model based on v, where the first in an order is the first to arrive. The player arriving last is now most important. If one player has zero individual worth, this player will arrive first with probability one (and receive zero value, see Feldman (2002), Section 6). Then the probability of S forming in v is the probability that $N \setminus S$ arrive last.

Lemma 2.6 *The probability of a coalition S forming in the random order model of relative importance is*

$$
p(S \,|\, w) = \frac{P(N, w)}{P(N \setminus S, w) \; RP(N \setminus S, N, w)},
$$

where $P(\varnothing, w) \equiv 1$ and $RP(\varnothing, N, w) \equiv P(N, w)$. The probability that S forms in a random order model of the proportional value based on v instead of its dual is

$$p(S \,|\, v) = \frac{P(N, v)}{P(S, v) \, RP(S, N, v)}.$$

Corollary 2.4 *The derivatives of the proportional value for any member $i \in N$ relative any $S \subseteq N$ in the game v or its dual w are as follows.*

$$\frac{\partial \varphi_i(v)}{\partial v(S)} = \begin{cases} \dfrac{\varphi_i(v)}{v(S)} \, p(S \,|\, v), & i \in S, \\[2mm] \dfrac{\varphi_i(v)}{v(S)} \left[p(S, N \,|\, v) - p(S, N \setminus \bar{i} \,|\, v) \right], & i \notin S, \text{ and} \end{cases}$$

$$\frac{\partial \varphi_i(w)}{\partial w(S)} = \begin{cases} \dfrac{\varphi_i(w)}{w(S)} \, p(N \setminus S \,|\, w), & i \in S, \\[2mm] \dfrac{\varphi_i(w)}{w(S)} \left[p(N \setminus (S \cup \bar{i}), N \,|\, w) - p(N \setminus (S \cup \bar{i}), N \setminus \bar{i} \,|\, w) \right], & i \notin S. \end{cases}$$

This corollary results from substitution of the results of Lemma 2.6 into Lemma 2.5. (Note that for any $i \in S$, $\partial Sh_i(v)/\partial v(S) = 1/s \, p_{Sh}(S \,|\, v) = (s-1)!(n-s)!/(s \, n!)$.) This corollary provides the intuition that the proportional value is not converse monotonic for an i with respect to $S \not\ni i$ in v when the importance of S (the probability of forming) is greater in N than in $N \setminus \bar{i}$. The probability that S arrives and then i is determined by Lemma 2.6, giving the following (computationally very inefficient) representation of the proportional value in classical marginal contribution form.

Corollary 2.5

$$\varphi_i(v) = \sum_{S \subset N \setminus \bar{i}} \frac{P(S \cup \bar{i})P(N)}{v(\bar{i})RP(\bar{i}, S \cup \bar{i})P(S) \, RP(S, N)} \left(v(S \cup i) - v(S) \right).$$

2.5 Linear-proportional family of values

Feldman (2005) Section 5.1 shows the linear and proportional values are both included in a one-parameter family of values. This family results from replacing random order consistency with a weaker axiom, which requires that effect on likelihood of a marginal change of the worth of any coalition S_i^r included in an order r be the same.

Axiom 2.1 PROPORTIONAL EFFECT: $\dfrac{\partial \ln L(r)}{\partial \ln w(S_i^r)} = \alpha, \quad i = 1, \ldots, n.$

It is straightforward to show that proportional effect, exclusion and anonymity identify the following likelihood (see Feldman (2005) Sections 3.3 and 5.1)

$$L^\alpha(r) = \left(\prod_{S \in r} w(S) \right)^{-\alpha}, \quad \alpha > 0. \tag{13}$$

The case $\alpha = 0$ is covered by Lemma 1.2. Define the normalizing factor $P^\alpha(N) = \left(\sum_{r \in \mathcal{R}(N)} L^\alpha(r) \right)^{-1}$, and a family of probability distributions p^α, indexed by α results:

$$p^\alpha(r) = P^\alpha(N) L^\alpha(r). \tag{14}$$

The induced expectation when (14) used to generate expectation (6) is is clearly the Shapley value when $\alpha = 0$ and the proportional value when $\alpha = 1$. Note, however, that these are the only random order consistent members of this family and that $P^\alpha(N)$ is not a potential for $\alpha \neq 1$.

3 Econometric attribution

An independent variable completely orthogonal to all others in an econometric model will not affect their model parameters or statistical significance levels. It is inert in the implied attribution problem. Random order consistency requires only that such a variable not affect other attributions. The theory of attribution can thus be applied to tasks such as variance and likelihood decomposition across a wide range of models.

Proposition 3.1 *If adding a factor z to an attribution problem based on likelihood function increases the likelihood of the model but leaves the individual and joint statistical significance of all other variables unchanged, then z is an inert factor.*

Proof: Let v and v^* represent the joint log likelihoods of factors in models based on N and $N^* = N \cup \bar{z}$, respectively, with duals w and w^*. Measure the joint significance of any factor set with the likelihood ratio test, which is based on their joint marginal likelihood contribution. Thus $w^*(S) = w(S)$ for all $S \subseteq N \setminus \bar{z}$, and $v^*(N \cup \bar{z}) - v^*(N \cup \bar{z} \setminus S) = v^*(N) - v^*(N \setminus S)$ for all $S \subseteq N$. Since $w^*(z) > 0$, z is inert in w. $\square$

Exclusion requires that a factor making zero final marginal contribution to model performance, when all others have positive final marginal contribution, receive zero attribution share. Such a factor will have zero statistical significance (*e.g.*, $p = 1$ that $\beta = 0$). Exclusion makes attribution as closely related to statistical significance as possible given the mutual correlation of explanatory factors.

Feldman (2005a) proposes four admissibility criteria for estimators of statistical relative importance: *nonnegativity*, *proper exclusion* (here, simply exclusion), *proper*

inclusion and *full contribution*. Proper inclusion requires factors making positive (final) marginal contribution receive positive attribution. Full contribution requires that the sum of attributions for factors in a set S equal their joint marginal contribution if they are uncorrelated with all factors in $N \setminus S$. PMA is an admissible estimator.

Feldman (2005a) also shows PMA components can be estimated consistently, develops properties and presents examples of LA, PMA and CVD (see fn. 4, below). Feldman (2005a) and Grömping (2007) note the potentially lower level of precision of PMVD components.[2]

In econometric attribution the utility function (1) can be understood to belong to an analyst and to be the model objective or likelihood function. Since OLS maximizes R^2 and R^2 is set monotonic, the marginal contribution to variance or R^2 can be the basis of attribution. Measures such as adjusted R^2, SIC and BIC cannot be used in attribution because they are not set monotonic. A further logical restriction for eligible measures is that it should be possible to use the measure to construct statistical significance tests. Such measures are more likely to be useful for the assessment of relative importance of model factors. F-tests can be constructed from R^2 values.

3.1   Indirect effects

Grömping (2007) proposes that LA should be used when interested in indirect effects. For example, if $x = f(y)$ and $y = g(z)$ then z will receive a zero PMA share in the attribution of $x = f'(y, z)$, but will receive an LA share if x and z are directly correlated. The following is a direct implication of the Fundamental Theorem.

Corollary 3.1 *Linear attribution is the only anonymous random order consistent attribution that recognizes indirect contributions to econometric model performance.*

Indirect effects can also be explicitly attributed with an assumed non-simultaneous causal structure. The result is a nesting of attribution problems. Owen (1977) shows that nested games can be interpreted as a composition of random order models.

Proposition 3.2 *Consider nested OLS models $x = f(y)$ and $y = g(z)$. The variance of x indirectly explained by z is $\sigma_x^2 R_{xy}^2 R_{yz}^2$.*

Proof: First, $\hat{x} = a + \sigma_{xy}/\sigma_y^2 y$. The projection of $\hat{x}$ on z gives $\hat{\mathbf{x}}$, the part of x explained by y that is explained by z: $\hat{\mathbf{x}} = b + \sigma_{\hat{x}z}/\sigma_z^2 z$, where $\sigma_{\hat{x}z} = \sigma_{xy}/\sigma_y^2 \sigma_{yz}$. The variance of $\hat{\mathbf{x}}$ is $[\sigma_{xy}\sigma_{yz}]^2/[\sigma_y^2\sigma_z^2]^2\sigma_z^2$. The variance share of x thus explained by z is then

$$R_{xz}^2 = \frac{[\sigma_{xy}\sigma_{yz}]^2}{\sigma_x^2[\sigma_y^2]^2\sigma_z^2} = \frac{\rho_{xy}^2\sigma_x^2\sigma_y^2\rho_{yz}^2\sigma_y^2\sigma_z^2}{\sigma_x^2[\sigma_y^2]^2\sigma_z^2} = \rho_{xy}^2\rho_{yz}^2 = R_{xy}^2 R_{yz}^2$$

$\square$

---

[2] Grömping (2006) describes the relaimpo package for R statistical freeware, which estimates several relative importance metrics for models and computes bootstrap confidence intervals.

3.2 The structure of econometric attributions

Econometric attribution is based on the model objective function to be maximized:

$$v(S) = \Theta(S) - \Theta(\varnothing), \qquad\qquad \text{for all } S : S \subseteq N, S \neq \varnothing, \qquad (15)$$

where $\Theta(S)$ is the objective function value when the factors in S are in the model and those in $N \setminus S$ are not. $\Theta(\varnothing)$ is the null utility of formula (1), the model objective value when not including any factors in N. For OLS with R^2 as the objective and an intercept, and when the intercept is not a factor, $\Theta(\varnothing) = 0$. With log-likelihoods, typically $v(S) < 0$ without normalization. If $v^*(S) = \text{abs}[\Theta(S)]$, v^* is not monotonic and results are useless. Also, likelihood (*e.g.*, $v^*(S) = \exp[\Theta(S)]$) is not a substitute for log-likelihood because it is not used in significance tests.

Attribution shares, and not magnitudes or differences, are the essential information. It may be difficult to properly determine the null likelihood for some models.[3] In such cases, linear attribution shares cannot usefully be defined because they are a function of $\Theta(\varnothing)$. PMA *shares*, however, are invariant to changes in $\Theta(\varnothing)$, so long as monotonicity is maintained, because in the dual game to v the worths of all $S \subsetneq N$ are invariant to changes in $\Theta(\varnothing)$. Changes in the dual worth $w(N)$ only scale PMA components. These relationships are consequences of Lemma 2.2.

Corollary 3.2 *Let $v^*(S) = c\,v(S) + d, c > 0$ for all S and let w^* be the dual of v^*. Then $\varphi_i(w^*)/w^*(N) = \varphi_i(w)/w(N)$. PMA is* share *scale and translation invariant.*

OLS R^2 attributions sum to the R^2 of the model, but can be normalized to sum to 100%. R^2-like measures that indicate the relative explanatory power of some maximum likelihood models have been constructed (see Maddala (1983), Section 2.11).

3.3 Fair's (1978) infidelity study

Econometric attribution is studied here in the context of Fair's (1978) study of 601 responses of first-time married individuals to a 1969 sex survey conducted by *Psychology Today*. Fair studies the predictors of the frequency of marital infidelity. Green (2003) reanalyzes this publicly available data. Fair's original model is first analyzed using OLS, tobit and Poisson regression. The determinants of marital happiness are then examined with OLS and ordered probit. In addition to model results, LAs, PMAs and covariance decompositions ($CVDs$) are reported.[4]

---

[3]A linear model without an intercept is a simple example.

[4]CVD is a standard approach to variance decomposition in economics and finance. The CVD component of a factor i is $\beta_i 1'_N \Sigma \beta$, where β is the model coefficient vector, Σ is the factor covariance matrix and 1_N is a $n \times 1$ vector of ones. See Feldman (2005a) Sections 4.3, 4.4 and 5 for results concerning CVD and examples. CVD can be described as a form of LA that results when the optimization of formula (1) is applied to only to the complete model.

Factor	beta	st. dev.	t-stat.	PMA	LA	CVD
OLS						
Age	-0.050	0.022	-2.278	4.83%	5.73%	-10.27%
Years married	0.162	0.037	4.387	20.16%	23.59%	38.87%
Religious	-0.476	0.111	-4.279	18.92%	19.63%	18.54%
Occupation	0.106	0.071	1.491	2.14%	1.84%	2.21%
Happiness	-0.712	0.118	-6.021	53.94%	49.21%	50.66%
R^2: 0.1314						
Tobit						
Age	-0.179	0.079	-2.27	5.30%	5.33%	-10.09%
Years married	0.554	0.135	4.12	19.22%	22.64%	39.27%
Religious	-1.686	0.404	-4.18	20.16%	20.78%	21.30%
Occupation	0.326	0.254	1.28	1.89%	1.59%	1.84%
Happiness	-2.285	0.408	-5.60	53.44%	49.65%	47.69%
Log(scale)	2.110	0.067	31.44			
Model ln L: -705.5762,		null model ln L: -744.7375				
Poisson						
Age	-0.032	0.006	-5.51	4.86%	5.68%	-11.28%
Years married	0.116	0.010	11.68	23.12%	25.71%	45.09%
Religious	-0.354	0.031	-11.46	20.88%	21.19%	22.95%
Occupation	0.080	0.019	4.11	2.61%	2.13%	3.08%
Happiness	-0.409	0.027	-14.95	48.54%	45.29%	40.16%
Model ln L: -1180.05,		null model ln L: -1462.75				

Table 2: Reanalysis of Fair (1978). Independent variable: frequency of infidelity.

The response variable studied by Fair is the self-reported frequency of infidelity. The explanatory factors include the respondent's sex, age, education and occupation, the number of years married, whether the respondent has children and the respondent's (self reported) level of religious belief and marital happiness.[5]

Fair excludes sex, education and having children as predictive factors based on their joint statistical insignificance. In attributions based on the complete model, all attribution methods assign these factors small attribution components, consistent with their low individual test statistics. The respective t-statistics are (0.18, 0.41, 0.21). The corresponding PMA components are (0.03%, 0.21% 0.05%). PMA and t-statistics can together better motivate efficient multivariate hypothesis testing than t-statistics alone. LA results, (0.15%, 3.12%, 0.29%), are not as clear for education.

---

[5]Fair (1978) provides descriptive statistics. The data set is available online.

Factor	beta	st. dev.	t-stat.	PMA	LA	CVD
OLS						
Sex	-0.070	0.103	-0.69	1.45%	0.75%	0.27%
Age	-0.006	0.008	-0.78	4.33%	19.92%	11.25%
Years married	-0.035	0.014	-2.52	58.54%	35.13%	48.50%
Children	-0.207	0.120	-1.74	10.74%	19.83%	18.63%
Religious	0.085	0.038	2.22	7.17%	5.44%	2.43%
Education	0.074	0.022	3.41	16.08%	17.71%	19.73%
Occupation	-0.025	0.030	-0.83	1.69%	1.21%	-0.81%
R^2 : 0.0894						
Ordered probit						
Sex	-0.092	0.106	-0.86	2.09%	1.04%	0.92%
Age	-0.005	0.008	-0.68	3.21%	19.72%	9.70%
Years married	-0.037	0.014	-2.60	58.64%	36.18%	49.57%
Children	-0.262	0.125	-2.09	15.60%	23.18%	23.82%
Religious	0.087	0.039	2.20	6.87%	5.16%	2.07%
Education	0.069	0.022	3.10	12.08%	13.70%	14.31%
Occupation	-0.025	0.031	-0.81	1.51%	1.03%	-0.40%
Model $\ln L$: -1180.05,			null model $\ln L$: -1462.75			

Table 3: Data from Fair (1978). Independent variable: Self-rated marital happiness.

Table 2 presents results based on Fair's original model. Factor coefficients are not directly comparable, Greene (2003) discusses the expected relationships. OLS and probit t-statistics are similar. Poisson factor t-statistics are considerably higher. Age and happiness reduce infidelity, but years married increases it.

PMA, LA and statistical significance rankings are identical in this case. CVD consistently finds year married about twice as important as PMA or LA. CVD also finds years married more important than happiness in the Poisson regression. Model choice, in this case, affects coefficients and their precision, but has little effect on LA, PMA or the relative magnitudes of t-statistics. Note that LA and PMA give identical results when factors are uncorrelated. CVD results then also agree for OLS.

The negative CVD components for age in all models are problematic and causes the high CVD importance of years married. When age is removed from the OLS model of infidelity the CVD relative importance of years married falls to 25%, and that of happiness rises to 53%, similar to PMA and LA results.

Table 3 shows OLS and ordered probit models of marital happiness. Estimated coefficients of both models are similar, the except probit estimate of the effect of having children is almost 30% larger, and the PMA component is almost 50% larger. The

increase in the LA component is just above 15%. Explained variance is proportional to the square of the beta in uncorrelated OLS, which would imply a 69% increase.

For the OLS and ordered probit models of marital happiness, all attribution methods indicate years married is most important. The bootstrap probability that PMA years married is larger (more important) than PMA education is greater than 85%. In contrast, t-statistics seem to show education is most precisely estimated. In practice, this result is commonly interpreted as indicating that education is most important.

Age and years married are highly correlated ($\rho = .78$), which degrades statistical test values. Indeed, when age is excluded the years married t-statistic becomes largest (-4.44 for OLS). Age, however, may be a valid predictor for marital happiness. If so, its exclusion from the model could bias the model and attribute its explanatory power to the factors with which it is most correlated. Also, the LA for years married rises to 51% when age is excluded, indicating the relative sensitivity of LA to model definition.

An F-test of the joint significance of sex and age in the OLS model is rejected (p=0.76). But removing age increases the magnitude of the years married beta by 22%. It would be a mistake to exclude age if the goal was to obtained unbiased coefficient estimates unless it could be maintained that age did not belong in the model in the first place. The age PMA of 4.33% signals this result.

4   General microeconomic models

4.1   Inert factors and utility attribution

Is random order consistency an acceptable axiom for utility attribution? Consider this question in the context of portfolio optimization. Uncorrelated assets, *e.g.*, risk-free assets, are not inert factors: Their utility contribution is not independent of other assets. Further, an investment option of zero utility cannot be inert by definition.

The ability to define an inert factor must thus rely on a factor that is not a portfolio choice. Its utility must be independent of the portfolio. These conditions are satisfied if utility is additively separable in the investment and the inert factor.

There is no reason why two completely independent choice problems could not, in principle, be composed into a single optimization problem. This simple fact appears to allow the definition of an inert factor in almost any choice problem. Random order consistency becomes applicable once it is established the existing attribution problem *could* be composed with a suitable independent problem as a larger attribution.

Proposition 4.1 *A factor in an attribution problem is inert if the attribution can be redefined as two independent attributions, one including only this factor.*

19

	Risk and return			Correlation					
Asset	Mean	Std.	Sharpe	LC	SC	Intl	Gvt	Corp	Bills
Large Cap	0.0106	0.041	0.258	1.000	0.729	0.619	0.116	0.217	0.069
Small Cap	0.0110	0.053	0.209	0.729	1.000	0.529	-0.015	0.110	-0.053
Int'l	0.0062	0.048	0.128	0.619	0.529	1.000	0.009	0.082	-0.068
LT Gvt.	0.0082	0.026	0.320	0.116	-0.015	0.009	1.000	0.948	0.079
LT Corp.	0.0079	0.021	0.384	0.217	0.110	0.082	0.948	1.000	0.069
T-Bills	0.0037	0.002	—	0.069	-0.053	-0.068	0.079	0.069	1.000

Table 4: Asset performance data, monthly 1988-2004.

4.2 Portfolio attribution

Let N be a set of n assets with expected return vector μ and covariance matrix Σ. Using quadratic utility and formula (1), $v(S)$ for any set of assets S is

$$v^{mvo}(S) = U(S) - U(\varnothing) = \max(0, \max_{\omega_S \in \Delta^S} \ \omega'_S \mu_S - \lambda \omega'_S \Sigma_S \omega_S), \tag{16}$$

where λ is the index of quadratic risk aversion, μ_S, ω_S and Σ_S are restricted to the members of S, ω_S is the portfolio allocation and Δ^S is the $s-1$ dimensional simplex of nonnegative portfolio weights that sum to unity. Here $U(\varnothing) = 0$. The restriction $v(S) \geq 0$ allows not investing if $v(S) < v(\varnothing)$, as is possible when restricted to some asset combinations. If exponential utility were used, then $U(S) < 0$ for all $S \subseteq N$ and the definition $v(S) = U(S) - U(\varnothing)$ is necessary, as in Section 3.2 for likelihoods.

Table 4 reports the assets used in this example, along with their monthly mean total returns, standard deviations, Sharpe ratios and correlations for the period 1988 to 2004.[6,7] Table 5 presents the optimal portfolios for an individual with quadratic risk aversion level $\lambda = 2$, along with two sets of attributions. One is for all investment options. The other excludes corporate bonds. For each set, PMA and LA results and (discrete) marginal utility contributions are reported. All attributions are normalized to 100%. The final column of Table 5 presents the proportional marginal performance attribution (PMPA) of assets with positive portfolio allocations. The PMPA of an asset is the ratio of its PMA to its portfolio weight.

---

[6]Monthly data from Ibbotson Associates covers the period January 1, 1988 to December 31, 2004. Large cap is the S&P 500, small cap is the Russell 2000, international is the MSCI Europe, Asia and Far East Index, long-term government is the Ibbotson U.S. Long-Term Government Bond Index, long-term corporate is the Ibbotson U.S. Long-Term Corporate Bond Index and t-bills is the Ibbotson U.S. 30 Day T-Bill.

[7]The Sharpe ratio is the excess return of an asset over the risk free rate divided by its standard deviation. The Sharpe ratio is a measure of excess return per unit of risk.

Asset	Port. Wts.	With LT Corp.			W-out LT Corp.			PMPA
		PMA	LA	Marg.	PMA	LA	Marg.	
Large Cap	37%	49%	25%	55%	49%	31%	19%	1.32
Small Cap	17%	7%	17%	19%	7%	21%	6%	0.41
Int'l	0%	0%	3%	0%	0%	4%	0%	—
LT Gvt.	46%	44%	23%	26%	44%	33%	75%	0.96
LT Corp.	0%	0%	23%	0%	—	—	—	—
T-Bills	0%	0%	9%	0%	0%	11%	0%	—

Table 5: Asset allocation and performance attribution when $\lambda = 2$.

The optimal portfolio monthly return and standard deviation are 0.0096 and 0.026 and its Sharpe ratio is 0.364. It is about 54% equity and 46% fixed income. The PMA is in similar proportion (56%/44%). Only assets with positive portfolio weight receive a PMA share, while every asset class receives some LA share. LA gives government and corporate bonds almost equal importance even though corporate bonds receive no portfolio weight. Marginal contributions are heavily weighted to equity because of the substitutability of government and corporate bonds.[8]

PMPA provides a measure of an asset's per dollar contribution to portfolio performance. The PMPA ranking is differs from the Sharpe ratio ranking. Government bonds have the highest Sharpe ratio of all allocated assets, followed by large cap equity. But Large cap equities contribute most to portfolio performance per unit of investment. The PMPA of large cap equity is more than three times the PMPA of small cap equity even though its Sharpe ratio is only 23% higher. The Sharpe ratio measures stand-alone performance, whereas PMPA evaluates performance relative to a specific portfolio. (Note, corporate bonds have the highest Sharpe ratio but receive no asset allocation.) Further, attribution is relative to a specific risk profile.

In the second set of attributions of Table 5, corporate bonds are not an option. The PMA does not change. The LA equity share rises to 52%, but now government bonds are seen as more important than large cap equity. The relative marginal contribution of government bonds is now very high without substitution from corporate bonds, and equity marginal contribution is weak because of equity mutual correlation.

If t-bills and international equity are also removed from the choice set, the LA becomes $(37\%, 25\%, 28\%)$ and equity receives 62% total attribution. PMA results are unchanged. When large cap and small cap shares are aggregated, the PMA is $(52\%, 48\%)$ and the LA is $(53\%, 47\%)$. In all cases, PMA implies that a little over

---

[8]The Covariance decomposition is (83%, 59%, -42%) to the three allocated assets. Only portfolio assets receive non-zero CVD allocations. CVD is based on portfolio weights and Σ (see footnote 4) and is commonly used in portfolio analysis. A marginal increase in government bonds reduces portfolio risk. This is the essential meaning of the negative CVD coefficient. CVD weights marginal risks by portfolio weights. Long term bonds, however, clearly both contribute to risk and utility.

half of portfolio performance is attributable to equity and almost half to bonds. LA implications depend strongly on details of the attribution analysis.

Unutilized assets are irrelevant for PMA because of exclusion. In this example, and generally, PMA is not very sensitive to aggregation; but it is possible to construct contrary examples. Linear attribution is sensitive to both unutilized factors and factor aggregation (as in example 5.1, below). The differences between LA and PMA may be a useful measure of the substitutability of factors. PMA offers a consistent method of attributing joint effects in the creation of investment value, and thereby providing insight into its sources.

5   Joint cost attribution

Cost allocation has traditionally been considered a bargaining problem by game theorists. In cost attribution, the utility function of formula (1) is an aggregate cost (disutility) function. The arguments of the cost function are typically aggregate demands for products or services over cost centers. The applicability of random order consistency to cost attribution raises the same types of issues as considered for utility attribution in Section 4.1. The potential existence of an inert cost center is no different the potential existence of an inert factor in a utility or econometric attribution.

5.1   Linear or proportional marginal cost attribution?

Neither inclusion nor exclusion appear obviously appropriate for cost attribution. While exclusion might seem appropriate for the attribution of profits, many might find it problematic that somebody is getting something for nothing. Exclusion in cost allocation requires that a party that imposes no final marginal cost should have no attributed cost if all other parties impose final marginal costs.

Inclusion may appear an innocuous axiom in this context, however linear cost allocation methods have not been widely adopted and were not well-received by accounting researchers. Moriarity (1975), Banker (1981) and Gangolly (1981) explicitly support or propose proportional alternatives. Banker writes that linear attribution is "not consistent with our axioms, and hence with traditional cost allocation methods." (1981:127). Thomas (1977) provides examples that demonstrate the vulnerability of linear attribution to strategic definition of cost centers.[9]

Example 5.1 *The potential for strategic manipulation in cost attribution is illustrated in a simple sunk cost attribution problem. Assume n cost centers, sunk cost C and fixed marginal cost c. Cost center i has usage x_i. The cost of service to any*

---

[9]Note also the conference proceedings Moriarity (1981). See Feldman (1999), Section 9.2, for more on accountant's perspectives.

coalition of centers S is $v(S) = C + c\sum_{i \in S} x_i$. The marginal cost of any coalition of centers $S \neq N$, is represented in the dual game by $w(S) = c\sum_{i \in S} x_i$. Thus, the game $w(S)$ is additive for all $S \subsetneq N$: $w(S) = \sum_{i \in S} w(\bar{i})$.

Lemma 2.2 implies $PMA_i = \varphi(w)_i = C/x_i$ because w can seen as a rescaling of $w^*(N)$ in a w^* where $w^*(N) = c\sum_{i \in N} x_i$. Similarly, $LA_i = cx_i + C/n$. When C is very small, LA and PMA are approximately the same since then $v(N) \approx c\sum_{i \in N} x_i$. However, allocations diverge as C increases. When $c\sum_{i \in N} x_i \ll C$, LA for all centers approaches C/n, regardless of the disparities in usage between cost centers, providing incentive for group managers to aggregate reporting units.

Corollary 5.1 *The proportional marginal share of costs attributed to any cost center is invariant with the magnitude of the fixed costs when fixed costs are scale invariant.*

6   Discussion and conclusion

This paper introduces a new expectation-based random order approach to cooperative value. Linear and proportional value are characterized in a common setting without assuming linearity or proportionality. Random order consistency, the key element of this approach, assumes only that we can see the "expectational mechanics" of the random order process, and that these mechanics do not change if invisible inert players or factors are added. These mechanics have specific interpretation regarding the relative importance of factors in the theory of attribution developed here.

Random order consistency is key to both theoretical and applied results. Contrast this with the limited power of reduced game consistency discussed in Remark 1.1. Random order consistency is strongly and uniquely associated with cooperative value.

Consider the scope of this theory if based on reduced game consistency. The use of reduced game consistency must imply some correspondence between the reduced game and the original attribution. But few, if any, properties of a game or cost function are preserved by reduction. The reliance on the dual game is also problematic. Consider an attribution problem with a vector of demands x and cost function $c(y) = y^\alpha$, $\alpha \neq 1$, so $w(S) = c(x_N) - c(x_{N \setminus S}) = [\sum_{i \in N} x_i]^\alpha - [\sum_{i \in S} x_i]^\alpha$. Let w^* be the game resulting from the reduction of player k, so that $N^* = N \setminus \bar{k}$ and $w^*(S) = w(S \cup \bar{k}) - \phi_k(w)$, where $\phi_i(w)$ is $Sh_i(w)$ or $\varphi_i(w)$. In general, there will be no vector of demands x^* such that $w^*(S) = c(x_{N^*}^*) - c(x_{N^* \setminus S}^*)$ for all S. Nor, in general, will there be a cost function c^* such that $w^*(S) = c^*(x_{N^*}^*) - c^*(x_{N^* \setminus S}^*)$ for all S.[10]

Random order consistency places minimal demands on the structure of an attribution problem. Section 4.1 proposes it appropriate for any attribution problem that can be composed with an independent attribution.

---

[10]Compare this perspective on the reduced game with Dubey's (1982) discussion regarding the the benefits of a direct characterization of a cost allocation problem based on its specific features.

The specific applications studied here demonstrate practical potential applications. PMA appears to provide theoretically sound measures of the econometric relative importance of model independent variables. These measures also appear to allow comparison of relative importance across models and to be useful diagnostics for model construction and evaluation. Feldman (2005a and 2006) applies a similar approach with financial factor models to provide variance decomposition of an asset manager's return stream to identify the relative importance of factor drivers.

The portfolio attribution example shows the considerable knowledge of data and preferences necessary for utility attribution, but, also, the potential benefit. Portfolio PMA and PMPA are new investment analysis tools and examples of practical feasible applications. They are distinctive in that they evaluate an asset's total, as opposed to marginal, performance both relative to a portfolio and an investor's risk preferences.

In this theory of attribution, linear and proportional value emerge as parallel expressions of an underlying value principle. This relative importance approach to cooperative value is developed specifically for the single-utility function attribution environment, but applies directly to TU bargaining as well. Feldman (2002) develops similar results for the more general nontransferable utility (NTU) games using ratio potentials. In this "dual" approach to value either the linear or proportional "mode" might be more appropriate for a particular application.

Game theoretic interest in proportional solutions, including the proportional value, has been hampered by the belief that they are translation dependent (see, *e.g.*, Hart and Mas-Colell (1988: 595*n*)). Real allocations seem to change with affine translation of a player's utility function. This outcome should arouse scrutiny as it implies the incoherence of a basic distributive principle recognized at least since Aristotle.[11]

The apparent translation dependence of proportional solutions results from a missing element in the standard NTU cooperative game. The *basis point*, a vector of utilities, one for each player, allows determination of the marginal utility to any player of any bargaining outcome. Feldman (2005b) demonstrates the existence of the basis point and shows that proportional solutions are translation invariant when a player's basis utility and utility function are subjected to the same affine transformation.[12]

This theory of attribution is based on a single decision maker and, hence, would be unaffected even if proportional bargaining solutions were translation dependent. Examples presented here clearly demonstrate the relevance of the basis point in attribution problems. Section 3.2 shows that econometric attributions are not uniquely defined unless the null model objective function value can be defined and used as the basis point. (Interestingly, PMA shares are invariant to the choice of the basis point.) Similarly, Section 4.2 shows that utility attribution must specifically define null model utility for utility functions where null model utility is not zero.

---

[11]Young (1994: 64) writes "[p]roportionality is deeply rooted in law and custom as a norm of distributed justice." Moulin (1999) starts by quoting Aristotle: "Equals should be treated equally, and unequals, unequally in proportion to relevant similarities and differences."

[12]Recognition of the basis point implies *intrapersonal* but *not* interpersonal comparison of utility.

Finally, researching the accounting literature, I was often surprised by the intensity of the arguments against linear cost allocation, and sometimes by the palpable frustration. Thomas (1977: 43), in his study of cost allocation and transfer pricing, chooses the following epigraph (from Lousia M. Alcott's *Jack and Jill*) for his chapter on cost allocation using the Shapley value: "...Molly retired to wet her pillow with a few remorseful tears, and to fall asleep, wondering if real missionaries ever killed their pupils in the process of conversion." Neutrality between linear and proportional methods is theoretically sound and evolutionarily wise.

7 Proofs

7.1 Appendix A: Proof of Lemma 1.1

Proof: **I.** First consider multiplicative likelihoods. Endogenous random order consistency requires that functions of the worths of factor sets including an inert factor factor out of the sum of likelihoods generated by a suborder. The terms in $\sum_{r \supset r_o} \prod_{i \in N} l_i(w(S_i^r))$ do not factor directly. An alternate approach is required.

Consider a single inert factor $z \in N$. Select any r_0 in $\mathcal{R}(N \setminus \bar{z})$. By inserting z at all possible positions in r_0, all orders in N that include r_0 are generated. Define $r_0^i \in \mathcal{R}(N)$ as the order resulting when z is inserted at the ith position: $r_0^i = (r_{0_1}, \ldots, r_{0_{i-1}}, z, r_{0_i}, \ldots, r_{0_{n-1}})$. Then let $R^* = \{r_0^i\}_{i=1}^n = (r_0^1, r_0^2, \ldots, r_0^n)$.

Define $\bar{S}$ as all nonempty factor sets included in r_0, so that $\bar{S} = \{S : S \subset r_0\} = \{S_i^{r_0}\}_{i=1}^{n-1}$. Let $\bar{S}^z$ contain the singleton $\bar{z}$ and the union of each factor set in $\bar{S}$ with z: $\bar{S}^z = \{S : S = T \cup z, T \in \bar{S} \text{ or } T = \varnothing\}$. Index the members of $\bar{S}$ and $\bar{S}^z$ by their cardinality: $\bar{S}_3$ has cardinality three. Define $S^* = \bar{S} \cup \bar{S}^z$.

Set of sets S^* helps to form a common denominator for multiplicative likelihoods. For any $r_0^i \in R^*$, let R_i^* be the factor sets that are included in r_0^i: $R_i^* = \{S \mid S \in r_0^i\}$. $R_i^* \subset S^*$. Define T_i^* as the sets in S^* not included in r_0^i: $T_i^* = S^* \setminus R_i^*$. Let $K = \prod_{S \in S^*} l_i(S)^{-1}$, where i is the cardinality of S. Then, for all $i \in N$,

$$L(r_0^i) = \prod_{S \in r_0^i} l_i(S) = \frac{1}{K} \prod_{S \in T_i^*} l_i(S)^{-1} \equiv \frac{1}{K} \hat{L}(r_0^i), \tag{A1}$$

where $\hat{L}$ is a partial likelihood with $\hat{L}_i = l_i^{-1}$. For every $j < n$ there are two sets in S^* with j factors. Only one of these factor sets is in any T_i^*. These are $S_j^{r_o} = \bar{S}_j$ and $\bar{S}_j^z = S_{j-1}^{r_o} \cup z$. If one is in R_i^* then it is not in T_i^*, and vice-versa. For $j < i$, $\bar{S}_j^z \in T_i^*$ because $\bar{S}_j \in R^*$, and for $j \geq i$, $\bar{S}_j \in T_i^*$ because $\bar{S}_j^z \in R^*$. Thus, $T_i^* = \{\bar{S}^z\}_{j=1}^{i-1} \cup \{\bar{S}\}_{j=i}^{n-1}$. Note $\bar{S}_n^z = N \in R_i^*$ and $N \notin T_i^*$.

Let T_j^{*i} be the factor set in T_j^* of cardinality i. Then

$$\hat{L}(r_0^1) + \hat{L}(r_0^2) \;=\; \prod_{i=1}^{n-1} l_i(T_1^{*i})^{-1} + \prod_{i=1}^{n-1} l_i(T_2^{*i})^{-1}$$

$$=\; \left(l_1(\bar{S}_1^z)^{-1} + l_1(\bar{S}_1)^{-1} \right) \prod_{i=2}^{n-1} l_i(\bar{S}_i)^{-1}.$$

When $w(\bar{S}_1^z) = w(\bar{z}) > 0$ factorization requires $l_1(S) = l_2(S) = w(S)^{-1}$. Then $l(S)^{-1} = w(S)$ and, since z is inert, $w(\bar{S}_2^z) = w(\bar{z}) + w(\bar{S}_1)$. Note that any l with $l(0) = \infty$ trivially allows factorization if $w(\bar{z}) = 0$. (Thus the requirement that inert factors have positive individual worth.) If and only if $l_1(S) = l_2(S) = w(S)^{-1}$ can these terms be combined and absorbed into the product

$$\hat{L}(r_0^1) + \hat{L}(r_0^2) = w(\bar{S}_2^z) \prod_{i=2}^{n-1} l_i(\bar{S}_i)^{-1}.$$

Induction is used to sum likelihoods. Assume that $l_i(S) = w(S)^{-1}$ for all $i < j$ and that the following pattern holds for sums of the first $j-1$ orders in R^*:

$$\sum_{i=1}^{j-1} \hat{L}(r_0^i) = \prod_{i=2}^{j-1} w(\bar{S}_i^z) \prod_{i=j-1}^{n-1} l_i(\bar{S}_i)^{-1}.$$

Under the above assumptions, $\hat{L}(r_0^j)$ may be written

$$\hat{L}(r_0^j) = \prod_{i=1}^{j-1} w(\bar{S}_i^z) \prod_{i=j}^{n-1} l_i(\bar{S}_i)^{-1}.$$

Adding and factoring the last two expressions gives the first formula below. The second follows if and only if $l_j(S) = w(S)^{-1}$. The induction is a "zipper."

$$\sum_{i=1}^{j} \hat{L}(r_0^i) \;=\; \left(w(\bar{S}_1^z) + w(\bar{S}_{j-1}) \right) \prod_{i=2}^{j-1} w(\bar{S}_i^z) \prod_{i=j}^{n-1} l_i(\bar{S}_i)^{-1}$$

$$=\; \prod_{i=2}^{j} w(\bar{S}_i^z) \prod_{i=j}^{n-1} l_i(\bar{S}_i)^{-1} \qquad\qquad (A2)$$

Evaluating (A2) at $j = n$ and recalling (A1) and that $\bar{S} = \{S : S \in r_o\}$ gives

$$\frac{1}{K} \sum_{i=1}^{n} \hat{L}(r_0^i) = \frac{1}{K} \prod_{i=2}^{n-1} w(\bar{S}_i^z) = w(\bar{z})^{-1} \prod_{S \in r_0} w(S)^{-1} = w(\bar{z})^{-1} L^{\diamond}(r_0).$$

The sum of likelihoods generated by an $r_0^* \in \mathcal{R}(N)$ divided by those generated by r_0 thus equals $L^\diamond(r_0^*)/L^\diamond(r_0)$, proving $p^\diamond(r_0|w_0) = \sum_{r_0^i \supset r_0} p^\diamond(r_0^i|w)$, where w_0 is the dual game generated by $N \setminus \bar{z}$ and $p^\diamond$ is the distribution induced by $L^\diamond$. Thus, $L^\diamond$ is random order consistent between games of cardinality $n-1$ and n for any $n > 2$. Consistency between games of cardinality $m > 2$ and n is established by repeat reduction of individual inert factors. $L^\diamond$ is random order consistent.

Any $L' = c\,L^\diamond$, where $c > 0$, is also random order consistent, but still generates $p^\diamond$. An L' created by dropping any l_i or combination of l_i is not random order consistent because the zipper will fail. The only way that the terms that depend on all $\bar{S}^z$ can be factored out of the sum $\sum_{r \supset r_0} L(r)$ is by combining terms using the relationship $l_i(\bar{S}_i) + l_1(\bar{S}_1^z) = l_{i+1}(\bar{S}_{i+1}^z)$, possible because z is inert. This restricts l_i to $l_i(S_i^r) = w(S_i^r)$ or $l_i(S_i^r) = w(S_i^r)^{-1}$. Using the terms in $L^\diamond$ directly (as opposed to $\hat{L}^\diamond$) allows combination when $l_i(S_i^r) = w(S_i^r)$, but combination is only possible between r_0^1 and r_0^2. There are no other endogenous, nontrivially different $\{l_i'\}_{i=1}^n$ that simultaneously allow combination of terms and addition of likelihoods. For example, with $l_i' = w(S_i^r)^2$, $l_i'(\bar{S}_i)^{\frac{1}{2}} + l_1'(\bar{S}_1^z)^{\frac{1}{2}} = l_{i+1}'(\bar{S}_{i+1}^z)^{\frac{1}{2}}$, but combining on this basis implies $L'(r_0)^{\frac{1}{2}} = \sum_{r \supset r_0} L'(r)^{\frac{1}{2}}$. Thus, $L^\diamond$ is unique up to scaling.

$L^\diamond$ is endogenous and multiplicatively separable. $L^\diamond$ is anonymous. Non-anonymous variants cannot be random order consistent because distinguishing orders will cause the zipper to fail. Finally, $L^\diamond$ is clearly positive when v is strongly monotonic.

II. Now consider additively separable endogenous likelihoods. Additive separability allows the collection and regrouping of sublikelihood functions, yielding

$$\sum_{r \supset r_0} L(r) = L(r_0) + \sum_{i=1}^{n-1}(n-i-1)\,l_i(\bar{S}_i) + \sum_{i=1}^{n} i\,l_i(\bar{S}_i^z).$$

Consider two orders r_0 and r_0^* in $\mathcal{R}(N \setminus \bar{z})$ and the ratio $\sum_{r \supset r_0^*} L(r)/\sum_{r \supset r_0} L(r)$. This ratio must equal $L(r_0^*)/L(r_0)$ if L is random order consistent. This implies

$$\frac{L(r_0^*)}{L(r_0)} = \frac{\sum_{i=1}^{n-1}(n-i-1)\,l_i(\bar{S}_i(r_0^*)) + \sum_{i=1}^{n} i\,l_i(\bar{S}_i^z(r_0^*))}{\sum_{i=1}^{n-1}(n-i-1)\,l_i(\bar{S}_i(r_0)) + \sum_{i=1}^{n} i\,l_i(\bar{S}_i^z(r_0))},$$

where the dependence on orders is indicated to avoid ambiguity. No further factorization of terms involving z is possible. Thus, this equality can hold for arbitrary v only when $l_i(\bar{S}_i^z(r_0)) = l_i(\bar{S}_i^z(r_0^*))$ for all i, i.e., when L is exogenous. $\square$

7.2 Appendix B: Proof of Lemma 1.4

Proof: Select any $r \in \mathcal{R}(N \setminus \bar{i})$ and let $r^j \in \mathcal{R}(N)$, $j = 1, \ldots, n$ be the order created by inserting i at position j, so that $r_j^j = i$. If $L^\diamond(r) = \sum_{j=1}^{n} L^\diamond(r^j) m_{r^j(i)}^P(r^j)$, the result

follows. The key to this reduction is more apparent when $m^P_{r^j(i)}(r^j)$ is represented as $w(S^j_j) - w(S^j_{j-1})$. The sum can then be expanded into $2n - 1$ separate terms. Consider first the order n term, where $w(S^n_n) = w(N)$ is the numerator and also in the denominator. This term simplifies to $L^\diamond(r)$, the LHS value. Now pair terms for orders r^j and r^{j-1}, for every $j = 1, \ldots, n$, and the following equality results:

$$\frac{w(S^{j-1}_{j-1})}{w(S^{j-1}_1) \cdots w(S^{j-1}_{j-1}) \cdots w(S^{j-1}_n)} + \frac{-w(S^j_{j-1})}{w(S^j_1) \cdots w(S^j_{j-1}) \cdots w(S^j_n)} = 0.$$

First, for all $k \neq j - 1$, $w(S^{j-1}_k) = w(S^j_k)$ since i has either not yet entered or has already entered both S^j_k and S^{j-1}_k. Then, for position $j - 1$, where $S^{j-1}_{j-1} \neq S^j_{j-1}$, here $w(S^{j-1}_{j-1})$ and $w(S^j_{j-1})$ are each on the top and bottom and cancel. $\qquad\square$

8 References

AUMANN, ROBERT J. AND LLOYD SHAPLEY (1974): *Values of non-atomic games.* Princeton: Princeton University Press.

BANKER, R. D. (1981): "Equity considerations in traditional full cost allocation practices: An axiomatic perspective," in Moriarity (1981), 110-130.

BILLERA, L. J. AND D. C. HEATH (1982): "Allocation of shared cost: A set of axioms yielding a unique procedure," *Mathematics of Operations Research*, **7**:32-39.

CALVO, EMILIO AND THEO H. S. DRIESSEN (2003): *The multiplicative value in cooperative TU-games*, Working paper, Department of Economic Analysis, University of Valencia.

CHEVAN, ALBERT AND MICHAEL SUTHERLAND (1991): "Hierarchical partitioning," *The American Statistician*, **45**:90-96.

DUBEY, PRADEEP (1982): "The Shapley value as aircraft landing fees-revisited," *Management Science*, **28**:869-874.

FAIR, RAY (1978): *A Theory of extramarital affairs*, *Journal of Political Economy*, **86**:45-61.

FELDMAN, BARRY (1999): *The proportional value of a cooperative game*, `netec.mcc.ac.uk/WoPEc/data/Papers/ecmwc20001140.html`.

—— (2002): *A dual theory of cooperative value*, `papers.ssrn.com/abstract=317284`.

—— (2005a): *Relative importance and value*, working paper, `www.prismanalytics.com/docs/RelativeImportance.pdf`.

—— (2005b): *Lost in translation: Basis utility and proportionality in games*, `papers.ssrn.com/abstract=764004`.

—— (2006): "Using PMVD to understand hedge fund performance drivers," in *Portfolio Analysis: Advanced Topics in Performance Measurement, Risk and Attribution*, Timothy Ryan (ed.), London: Risk Books.

FIRTH, DAVID (1998): *Relative importance of explanatory variables*, prepared for *Statistical*

Issues in the Social Sciences, Stockholm, October 1998, Oxford: Nuffield College, `www.nuff.ox.ac.uk/sociology/alcd/relimp.pdf`.

FRIEDMAN, ERIC J. (2004) "Paths and consistency in additive cost sharing," *International Journal of Game Theory*, **32**:501-518.

GANGOLLY, JAGDISH S. (1981): "On joint cost allocation: Independent cost proportional scheme (ICPS) and its properties," *Journal of Accounting Research*, **19**:299-312.

GREENE, WILLIAM (2003): *Econometric Analysis*, 5th ed., Prentice Hall.

GRÖMPING, ULRIKE (2006): "Relative importance for linear regression in R: The package relaimpo," *Journal of Statistical Software*, **17(1)**, `jstatsoft.org/v17/i01/v17i01.pdf`.

—— (2007): "Estimators of relative importance in linear regression based on variance decomposition," *American Statistician*, forthcoming, `www.tfh-berlin.de/~groemp/downloads/amstat_preprint.pdf`

GUL, FARUK (1989): "Bargaining foundations of the Shapley value," *Econometrica*, **57**:81-95.

HART, SERGIU (2006): "Shapley value," in *The New Palgrave: A Dictionary of Economics*, 2nd Edition, Steven N. Durlauf and Lawrence E. Blume, eds., Palgrave MacMillan, forthcoming. (`www.ma.huji.ac.il/~hart/abs/val-palg2.html`)

—— AND ANDREU MAS-COLELL (1989): "Potential, value and consistency," *Econometrica*, **57**:589-614.

—— AND —— (1996): "Bargaining and value," *Econometrica*, **64**:357-380.

HOFFMAN, P. J. (1960):, "The paramorphic representation of clinical judgment," *Psychological Bulletin*, **57**:116-131.

KALAI, EHUD AND DOV SAMET (1987): "On weighted Shapley values," *International Journal of Game Theory*, **16**:205-222.

KHMELNITSKYA, ANNA B. (1999): "Marginalist and efficient values for TU games," *Mathematical Social Sciences*, **38**:45-54.

—— AND THEO S. H. DRIESSEN (2001): "Semiproportional values for TU games," *Mathematical Methods of Operations Research*, **53**:495-511.

KRUSKAL, WILLIAM H. (1987): "Relative importance by averaging over orderings," *American Statistician*, **41**:6-10.

LEROUX, JUSTIN (2006): *A discussion of the consistency axiom in cost allocation problems*, HEC Montréal, Institut D'économie Appliquée working paper IEA-06-13.

LINDEMAN, RICHARD. H., PETER F. MERENDA, AND RUTH Z. GOLD (1980): *Introduction to Bivariate and Multivariate Analysis*, Scott Foresman, Glenview, IL.

MADDALA, G. S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge: Cambridge University Press.

MCNALLY, RALPH (2000): "Regression and model-building in conservation biology, biogeography and ecology: The distinction between and reconciliation of predictive and explanatory models," *Biodiversity and Conservation*, **9**:655-671.

MORIARITY, SHANE (1975): "Another approach to allocating joint costs." *The Accounting Review*, pp. 791-95.

—— , ED. (1981): *Joint Cost Allocations: Proceedings of the University of Oklahoma*

Conference on Cost Allocations, April 23-24, 1981. Norman: Center for Economic and Management Research.

MOULIN, HERVÉ (1999): "Axiomatic cost and surplus-sharing," Chapter 6 in the *Handbook of Social Choice and Welfare*, Kenneth J. Arrow, Amartya K. Sen, and Kotaro Suzumura, eds. Amsterdam: Elsevier.

—— AND SCOTT SHENKER (1992): "Serial cost sharing," *Econometrica*, **65**:1009-37.

MYERSON, ROGER (1980): "Conference structures and fair allocation rules," *International Journal of Game Theory*, **9**:169-82.

ORTMANN, K. M. (2000): "The proportional value for positive cooperative games," *Mathematical Methods of Operations Research*, **51**:235-48.

OWEN, GUILLERMO (1977): "Values of games with a priori unions," in *Essays in Mathematical Economics and Game Theory*, R. Hein and O. Moeschlin, eds., pp. 76-88, Berlin: Springer-Verlag.

ROTH, ALVIN, ed. (1988) *The Shapley Value: Essays in Honor of Lloyd Shapley*, Cambridge: Cambridge University Press.

SCHMEIDLER, DAVID (1969): "The nucleolus of a characteristic function game," *SIAM Journal of Applied Mathematics*, **17**:1163-1170.

SHAPLEY, LLOYD S. (1953): "A value for n-person games," in *Contributions to the Theory of Games*, Vol. II, eds. H.W. Kuhn and A.W. Tucker, Princeton University Press, 307-317.

—— (1964): "Values of large games VII: A general exchange economy with money," *RAND RM* 4248-PR.

SHUBIK, MARTIN (1982): "Incentives, decentralized control, the assignment of joint costs and internal pricing," *Management Science*, **8**:325-343.

SOOFI, EHSAN S., JOSEPH J. RETZER, AND MASOUD YASAI-ARDEKANI (2000): "A framework for measuring importance of variables with application to management research and decision models," *Decision Sciences* **31**:1-31.

THOMAS, ARTHUR L. (1977): *A Behavioral Analysis of Joint-Cost Allocation and Transfer Pricing*, Chicago: Arthur Andersen & Co.

VON NEUMANN, JOHN (1928/1959): "On the theory of playing games," English translation in *Contributions to the Theory of Games IV*, ed. by R. D. Luce and A. W. Tucker. Princeton, Princeton University Press, pp. 13-42.

WEBER, ROBERT J. (1998): "Probabilistic values for games," in Roth (1988), pp. 101-119.

YOUNG, H. PEYTON (ed.) (1985): *Cost Allocation: Methods, Principles, Applications*, Amsterdam: North-Holland.

—— (1994): *Equity: In Theory and Practice.* Princeton: Princeton University Press.