Jahrb. f. Nationalök. u. Stat. (G. Fischer Verlag, Stuttgart 1995) Bd. (Vol.) 214/1

Principal-Agent-Theorie ohne Substitutionsaxiom*)

Principal-Agent-Theory Without the Independence Axiom

Von Hans Rau-Bredow*, München

1. *Einführung*

Fast alle Modelle der formalen Principal-Agent-Theorie gehen von der Theorie des Erwartungsnutzens aus. Diese Erwartungsnutzentheorie wird jedoch in einer bis auf Allais (1953) zurückgehenden und in den letzten Jahren verstärkt geführten Diskussion kritisiert und durch alternative bzw. allgemeinere Entscheidungsregeln zu ersetzen versucht[1]). Zahlreiche Untersuchungen wie z. B. von Kahneman/Tversky (1979) haben gezeigt, daß das für diesen Ansatz grundlegende Substitutionsaxiom in der Realität häufig nicht beachtet wird. Die Erwartungsnutzentheorie ist daher nicht für die Principal-Agent-Theorie geeignet, auch wenn man diesen Ansatz zumindest als normative Vorgabe für Risikoentscheidungen akzeptiert. Denn bei der Berechnung einer optimalen Entlohnungsfunktion kommt es ja nicht darauf an, wie sich der Agent verhalten soll, sondern wie er sich tatsächlich verhält.

In der Principal-Agent-Theorie wird daher ein Ansatz benötigt, der nicht im Widerspruch zum tatsächlich beobachtbaren Risikoverhalten steht. Die wohl wichtigste Arbeit zur Weiterentwicklung der Erwartungsnutzentheorie stammt von Machina (1982). Bei diesem Ansatz wird das umstrittene Substitutionsaxiom nicht mehr vorausgesetzt, sondern lediglich eine sehr viel weniger einschränkende Differenzierbarkeitsbedingung verwendet. Mit Hilfe dieser Differenzierbarkeitsbedingung kann dann die Existenz einer „lokalen", auch von der Wahrscheinlichkeitsverteilung abhängigen Nutzenfunktion bewiesen werden, durch deren Eigenschaften sich ganz analog zum Fall einer herkömmlichen von Neumann/Morgenstern Nutzenfunktion die individuelle Risikoeinstellung bestimmen läßt. Die grundsätzlichen Ergebnisse der Erwartungsnutzentheorie lassen sich damit auf einen wesentlich allgemeineren Bereich übertragen.

Bisher wurde aber noch nicht untersucht, wie sich eine pareto – optimale Risikoallokation und die Lösung von Principal-Agent-Problemen mit Hilfe dieses allgemeinen Ansatzes bestimmen läßt. Um die entsprechenden Maximierungsaufgaben lösen zu können, muß jedoch bezüglich der individuellen Risikopräferenzen nicht mehr

*) Ich danke *Gerhard Illing* für hilfreiche Hinweise zu einer früheren Fassung dieses Beitrages.

[1]) Vgl. z. B. *Machina* (1987), *Sugden* (1986), *Weber/Camerer* (1987).

als Differenzierbarkeit vorausgesetzt werden[2]. In diesem Beitrag soll deshalb das grundlegende Principal-Agent-Modell von Holmström (1979) mit Hilfe der Theorie der differenzierbaren Risikopräferenzen von Machina (1982) verallgemeinert werden. Obwohl dann der formale Aufwand größer wird, ist es für die ökonomische Theorie bedeutsam, inwieweit die bisher erzielten Ergebnisse von den umstrittenen Voraussetzungen der Erwartungsnutzentheorie abhängig sind.

Der Beitrag ist wie folgt aufgebaut: Zunächst wird ein knapper Abriß der Theorie von Machina (1982) gegeben (Abschnitt 2). Anschließend wird als erstes Ergebnis gezeigt, daß eine pareto-optimale Risikoallokation auch im allgemeinen Fall dadurch gekennzeichnet ist, daß die nun anhand der lokalen Nutzenfunktionen zu berechnenden Grenzraten der Substitution übereinstimmen (Abschnitt 3). Im second-best Fall ist dagegen eine Abweichung von der pareto-optimalen Risikoallokation erforderlich, um einen risikoaversen Agenten zu einem höheren Arbeitseinsatz zu motivieren. Eine entsprechend modifizierte Lösung des Risikoallokationsproblems wird berechnet, wobei sich zeigt, daß der Vergleich von second-best und first-best Lösung nun schwieriger ist als bei dem von Holmström (1979) untersuchten Spezialfall der Erwartungsnutzentheorie (Abschnitt 4).

2. *Die Theorie der differenzierbaren Risikopräferenzen*

Gegeben sei ein auf der Menge der möglichen Verteilungsfunktionen F(x) einer zufälligen Größe X definiertes Präferenzfunktional U(F). Mit einer Nutzenfunktion u(x) ist das Präferenzfunktional üblicherweise als der Erwartungswert des Nutzens definiert:

$$U(F) = E_F\, u(x) = \int u(x)\, dF(x) \tag{1}$$

Dieser spezielle Ansatz und seine axiomatischen Voraussetzungen werden jedoch kritisiert[3]. Machina (1982) setzt deshalb bezüglich des Präferenzfunktionals lediglich eine allgemeine Differenzierbarkeitsbedingung (im Sinne von Fréchet) voraus[4]. Die durch das Substitutionsaxiom bedingte Linearität des Präferenzfunktionals ist damit nicht länger gültig. Es ist aber dann eine lokale Näherung durch einen linearen Ansatz möglich: Es kann die Existenz einer lokalen, auch von der Verteilungsfunktion F(x) abhängigen Nutzenfunktion u(x,F) bewiesen werden, so daß für beliebige F*(x) und F(x) die Differenz U(F*) − U(F) durch die Differenz der jeweiligen Erwartungsnutzen approximiert werden kann:

$$U(F^*) - U(F) = \int u(x,F)\, (dF^*(x) - dF(x)) + o(\| F^* - F \|) \tag{2}$$

Dabei steht o( ) für Restterme höherer Ordnung, d. h. es gilt $o(x)/x \to 0$ für $x \to 0$; der „Abstand" $\| F^* - F \|$ zwischen zwei Verteilungsfunktionen wird durch $\| F^* - F \| = \int | F^*(x) - F(x) |\, dx$ definiert. Anhand der lokalen Nutzenfunktion kann der verallgemeinerte Pratt-Arrow Index

$$r(x,F) = - \frac{u_{11}(x,F)}{u_1(x,F)} \tag{3}$$

gebildet werden, der die individuelle Risikoeinstellung mißt[5]. Ein Individuum mit einem größeren Pratt-Arrow Index verlangt auch eine größere Risikoprämie.

Um die empirischen Ergebnisse zum Risikoverhalten und insbesondere die charakteristischen Verletzungen des Substitutionsaxioms im Rahmen dieser allgemeinen Theorie erklären zu können, werden von Machina (1982) S. 300 in einem weiteren Schritt noch zwei Hypothesen formuliert, die den Pratt-Arrow-Index jeweils als Funktion einer der beiden Variablen betrachten:

Hypothese I: r(x,F) ist eine nichtfallende Funktion von x.

Hypothese II: Wenn F*(x) die Verteilung F(x) stochastisch dominiert, dann gilt $r(x,F^*) > r(x,F)$.

Von besonderer Bedeutung ist dabei die Hypothese II, danach steigt die Risikoaversion beim Übergang zu einer „besseren" Verteilung. Wie gezeigt werden kann, gibt diese Hypothese II eine plausible Erklärung für die z. B. von Allais (1953) oder Kahneman/Tversky (1979) angegebenen charakteristischen Verletzungen des Substitutionsaxioms[6]. In diesem Beitrag wird die Hypothese II außerdem bei der Untersuchung der second-best Lösung von Bedeutung sein.

3. *Optimale Risikoallokation*

Ein durch eine Zufallsvariable X mit Verteilung F(x) gegebenes unsicheres Einkommen soll nun pareto-optimal zwischen zwei Individuen aufgeteilt werden. Die beiden Ergebnisanteile werden mit y(X) und z(X) = X − y(X) bezeichnet, die zugeordneten Verteilungsfunktionen heißen $F_y(x)$ und $F_z(x)$. U(F) bzw. V(F) sind die Präferenzfunktionale der beteiligten Individuen, u(x,F) und v(x,F) die entsprechenden lokalen Nutzenfunktionen. Das Problem kann als ein Maximierungsproblem eines Prinzipals formuliert werden, mit der Nebenbedingung, daß der Agent den Reservationsnutzen v_{min} erhält:

$$\underset{y(\ )}{\text{Max:}}\ U(F_z) \tag{4}$$

$$\text{N.B.} \qquad V(F_y) \geqq v_{min}$$

[2]) Principal-Agent-Probleme lassen sich mathematisch exakt mit Hilfe der Variationsrechnung lösen. Die Aussagen der Variationsrechnung lassen sich mit Hilfe der sogenannten Fréchet-Ableitung herleiten (vgl. z.B. *Klingbeil* 1977, S. 22 ff.), und dasselbe Konzept der Fréchet-Ableitung wird von *Machina* (1982) auf Risikopräferenzen angewendet.

[3]) Vgl. *Machina* (1982) S. 280 ff. sowie die in FN 1 angegebene Literatur.

[4]) Zu den genauen mathematischen Details der folgenden Darstellung vgl. *Machina* (1982) S. 293 f.

[5]) Vgl. *Machina* (1982) S. 299. u_1 bzw. u_{11} bezeichnen die erste bzw. zweite partielle Ableitung nach dem ersten Argument.

[6]) Vgl. *Machina* (1982) S. 302 f. Besonders deutlich kann dies durch sogenannte Dreiecksdiagramme veranschaulicht werden, vgl. dazu auch die eingängige Darstellung bei *Sugden* (1986) S. 5 f.

Die Lagrangefunktion mit Multiplikator λ lautet:

$$L(y) = U(F_z) + \lambda\,(V(F_y) - v_{min}) \tag{5}$$

Die Fréchet-Ableitung $DL_y(\ )$ dieser Lagrangefunktion $L(y)$ ist ein für Funktionen $y^*(\ )$ definiertes lineares Funktional, das die folgende Näherung erfüllt:

$$L(y + y^*) - L(y) = DL_y(y^*) + o(\| y^* \|) \tag{6}$$

Im Maximum hat die Fréchet-Ableitung eine Nullstelle[7]), d.h. bei einer optimalen Beteiligungsfunktion $y(\ )$ gilt $DL_y(y^*) = 0$ für alle y^*. Die Fréchet-Ableitung kann mit Hilfe einer linearen Näherung, durch welche die Auswirkung einer marginalen Veränderung der Beteiligungsfunktion auf das Präferenzfunktional beschrieben wird, berechnet werden. Das folgende, im Anhang bewiesene Lemma formuliert eine solche lineare Näherung:

Lemma 1:

$$V(F_{y+y^*}) - V(F_y) = \int v_1(y(x),F_y)\, y^*(x)\, dF(x) + o(\| y^* \|_\infty)$$

Hier wurde die Maximums-Norm $\| y^* \|_\infty = \max |y(x)|$ verwendet, $v_1(\ ,\ )$ bezeichnet die partielle Ableitung nach dem ersten Argument. Angewendet auf die Gleichung (5) ergibt dies unter Berücksichtigung von $z(x) = x - y(x)$ folgende Fréchet-Ableitung:

$$DL_y(y^*) = \int [- u_1(z(x),F_z) + \lambda\, v_1(y(x),F_y)]\, y^*(x)\, dF(x) \tag{7}$$

Aus $DL_y(y^*) = 0$ für alle y^* folgt direkt:

$$\frac{u_1(z(x),F_z)}{v_1(y(x),F_y)} = \lambda \tag{8}$$

Formuliert man dieses Ergebnis für zwei verschiedene Einkommen x und $\tilde{x}$, dann ergibt sich:

$$\frac{u_1(z(\tilde{x}),F_z)}{u_1(z(x),F_z)} = \frac{v_1(y(\tilde{x}),F_y)}{v_1(y(x),F_y)} \tag{9}$$

Damit wird das bekannte Ergebnis für lokale Nutzenfunktionen verallgemeinert: Eine pareto-optimale Risikoallokation ist dadurch gekennzeichnet, daß die nun anhand der lokalen Nutzenfunktionen zu berechnenden Grenzraten der Substitution zwischen den für die verschiedenen Ergebnisse festgelegten Anteilen bei Prinzipal und Agent übereinstimmen.

Eine andere elegante Interpretation ergibt sich, wenn man die Gleichung (8) implizit nach x differenziert und das Ergebnis mit Hilfe der Pratt-Arrow-Indizes $r_P = r_p(z(x),F_z)$ bzw. $r_A = r_A(y(x),F_y)$ von Prinzipal bzw. Agent formuliert[8]):

[7]) Vgl. z. B. *Klingbeil* (1977) S. 34.

[8]) Für die entsprechende Formulierung im Fall der Erwartungsnutzentheorie vgl. *Rees* (1985) S. 9.

$$y'(x) = \frac{r_P}{r_P + r_A} \tag{10}$$

Mit diesem Resultat kann die Aufteilung des Einkommens direkt anhand der individuellen Risikoeinstellungen berechnet werden. So werden einem risikoaversen Agenten ($r_A > 0$) Einkommensschwankungen nur teilweise zugerechnet ($y'(x) < 1$). Bei einem risikoneutralen Prinzipal ($r_P = 0$) bezieht der Agent sogar lediglich ein ergebnisunabhängiges Fixum ($y'(x) = 0$).

4. *Optimale Verteilungsregeln bei Moral-Hazard*

Es wird nun davon ausgegangen, daß die Verteilungsfunktion $F(x,a)$ des Ergebnisses auch vom Arbeitseinsatz a des Agenten abhängig ist. Ein höherer Arbeitseinsatz führt zu einer stochastisch dominanten Verteilung, aber auch zu größeren privaten Kosten $c(a)$ des Agenten. Im second-best Fall kann der Arbeitseinsatz des Agenten nicht kontrolliert werden, sondern ergibt sich aus seinem individuellen Rationalitätskalkül. Zur Erzielung einer besseren Motivation kann es dann vorteilhaft sein, einen risikoaversen Agenten stärker am Ergebnis zu beteiligen.

Der Nutzen des Agenten sei durch $V(F_y(\ ,a)) - c(a)$ gegeben (additive Separierbarkeit). Das Maximierungsprogramm ist dann um eine Anreizbedingung zu ergänzen, weil der Agent seinen Arbeitseinsatz so wählen wird, daß Grenzerträge $\delta/\delta a\ V(F_y(\ ,a))$ und Grenzkosten $c'(a)$ übereinstimmen[9]):

$$\max_{y(\ )}\ U(F_z(\ ,a)) \tag{11}$$

$$\text{N.B.}\quad V(F_y(\ ,a)) - c(a) \geqq v_{min}\ \text{(Teilnahmebedingung)}$$

$$\frac{\delta}{\delta a} V(F_y(\ ,a)) - c'(a) = 0\ \text{(Anreizbedingung)}$$

Die zugehörige Lagrangefunktion mit den Multiplikatoren λ und μ lautet:

$$L(y) = U(F_z(\ ,a)) + \lambda\,[V(F_y(\ ,a)) - c(a) - v_{min}] + \mu \left[\frac{\delta}{\delta a} V(F_y(\ ,a)) - c'(a)\right] \tag{12}$$

Wie im vorherigen Abschnitt wird die Fréchet-Ableitung mit Hilfe von Lemma 1 berechnet. Dabei wird die Vertauschbarkeit von Fréchet-Ableitung und partieller

[9]) *Rogerson* (1985) konnte im Rahmen der Erwartungsnutzentheorie zeigen, daß dieser sogenannte first-order-approach ohne die Berücksichtigung höherer Ableitungen zu einer eindeutigen Lösung führt, wenn ein monotoner Likelihood-Quotient und eine konkave Verteilungsfunktion vorliegen. Etwas schwächere Bedingungen leitet *Jewitt* (1988) bei einem risikoneutralen Prinzipal ab.

Ableitung nach a vorausgesetzt. Mit $f(x,a)$ wird die Dichte der Verteilung $F(x,a)$ bezeichnet, $f_a(x,a)$ ist die partielle Ableitung der Dichte nach a. Man erhält:

$$
\begin{aligned}
DL_y(y^*) = {} & \int - u_1(z(x),F_z(\ ,a))\, y^*(x)\, dF(x,a) \\
& + \lambda \int v_1(y(x),F_y(\ ,a))\, y^*(x)\, dF(x,a) \\
& + \mu \frac{\delta}{\delta a} \int v_1(y(x),F_y(\ ,a))\, y^*(x)\, dF(x,a) \\
= {} & \int [- u_1(z(x),F_z(\ ,a))\, f(x,a) + \lambda\, v_1(y(x),F_y(\ ,a))\, f(x,a) \\
& + \mu \frac{\delta}{\delta a} v_1(y(x),F_y(\ ,a))\, f(x,a) \\
& + \mu\, v_1(y(x),F_y(\ ,a))\, f_a(x,a)]\, y^*(x)\, dx
\end{aligned}
\tag{13}
$$

Die notwendige Bedingung $DL_y(y^*) = 0$ für alle y^* ergibt jetzt folgende Charakterisierung der optimalen Entlohnungsfunktion:

$$
\frac{u_1(z(x),F_z(\ ,a))}{v_1(y(x),F_y(\ ,a))} = \lambda + \mu \left[\frac{f_a(x,a)}{f(x,a)} + \frac{\frac{\delta}{\delta a} v_1(y(x),F_y(\ ,a))}{v_1(y(x),F_y(\ ,a))} \right]
\tag{14}
$$

Damit ist die gesuchte Verallgemeinerung des Principal-Agent-Ansatzes gegeben. Dabei ist das bekannte Resultat von Holmström (1979) S. 77 als Spezialfall enthalten, denn im Fall der Erwartungsnutzentheorie ist die Nutzenfunktion von der Verteilung und somit auch vom Arbeitseinsatz a unabhängig, d. h. es gilt $\delta/\delta a\; v_1(y(x),F_y(\ ,a)) = \delta/\delta a\; v_1(y(x)) = 0$.

Für eine abschließende Interpretation ist das erzielte second-best Ergebnis noch mit dem first-best Fall zu vergleichen. Wenn die Gleichung (14) ebenfalls implizit nach x differenziert wird, dann erhält man nach einer etwas längeren, aber elementaren Rechnung in abkürzender Schreibweise:

$$
y'(x) = \frac{r_P + \mu \frac{v_1}{u_1} \frac{\delta}{\delta x} \frac{f_a}{f}}{r_P + r_A + \mu \frac{v_1}{u_1} \frac{\delta}{\delta a} r_A}
\tag{15}
$$

Das entsprechende Ergebnis für den first-best Fall ist durch Gleichung (10) gegeben. Bei einem Vergleich der beiden Ergebnisse kommt es zunächst darauf an, daß μ positiv ist. Dazu soll eine heuristische Überlegung genügen[10]: Aus der Formulierung der Lagrangefunktion in Gleichung (12) ist ersichtlich, daß der Multiplikator μ angibt, wie sich eine Veränderung des Grenzertrages $\delta/\delta a\; V(F_y(\ ,a))$ auf den Nutzen des Prinzipals auswirkt. Dieser Grenzertrag des Agenten wird jedoch tendenziell geringer ausfallen,

[10]) Einen Beweis gibt *Holmström* (1979) S. 90 für den Fall der Erwartungsnutzentheorie, der sich aber nicht direkt auf den Fall der differenzierbaren Risikopräferenzen übertragen läßt.

wenn der Agent weniger stark an den Ergebnisschwankungen beteiligt wird. Im Optimum kann daher nicht $\mu < 0$ gelten, weil der Prinzipal sonst durch eine geringere Ergebnisbeteiligung eine weitere Nutzensteigerung erzielen könnte, ohne daß sich gleichzeitig die vom Agenten verlangte Risikoprämie erhöht.

In Gleichung (15) sind auch $v_1(\ ,\ )$ und $u_1(\ ,\ )$ positiv[11]), $\delta/\delta x\; f_a(x,a)/f(x,a)$ ist bei einem monotonen Likelihood-Quotienten ebenfalls größer als Null[12]). Weil im speziellen Fall der Erwartungsnutzentheorie $\delta/\delta a\; r_A(y(x),F_y(\ ,a)) = \delta/\delta a\; r_A(y(x)) = 0$ gilt, folgt dann aus einem Vergleich der Gleichungen (10) und (15), daß der Agent im second-best Fall aus Motivationsgründen stärker an den Ergebnisschwankungen beteiligt wird als im first-best Fall. Eine solche einfache Aussage ist aber nun im allgemeinen Fall der differenzierbaren Risikopräferenzen nicht mehr möglich.

Wenn man zumindest voraussetzt, daß die Entlohnung des Agenten eine monoton steigende Funktion des Ergebnisses ist ($y'(x) > 0$), dann führt ein höherer Arbeitseinsatz zu einer stochastisch dominanten Verteilung des Anteils des Agenten am Ergebnis und dies wiederum aufgrund der Hypothese II zu größerer Risikoaversion, d. h. die Gültigkeit der Hypothese II impliziert die Aussage $\delta/\delta a\; r_A > 0$. Bei einem Vergleich der Gleichungen (10) und (15) kann dann nicht ausgeschlossen werden, daß $\delta/\delta a\; r_A$ zumindest für einige x-Werte so groß ausfällt, daß der Agent an diesen Stellen im second-best Fall sogar geringer als im first-best Fall am Ergebnis beteiligt ist. Die Hypothese II wirkt also tendenziell einer aus Motivationsgründen angezeigten stärkeren Ergebnisbeteiligung des Agenten entgegen, was intuitiv wie folgt begründet werden kann: Eine stärkere Ergebnisbeteiligung induziert einen höheren Arbeitseinsatz des Agenten und damit eine stochastisch dominante Ergebnisverteilung, die Hypothese II besagt aber, daß der Agent bei einer stochastisch dominanten Verteilung eine größere Risikoprämie verlangt.

Abschließend ist noch zu bemerken, daß in diesem Beitrag lediglich das Basismodell der Principal-Agent-Theorie verallgemeinert wurde. Mehrperiodische Ansätze oder die Implementierung von Informations- und Kontrollsystemen wurden z. B. nicht behandelt. Möglicherweise können sich jedoch aus der Vorgehensweise dieses Beitrages auch Anregungen für entsprechend weiterführende Forschungen ergeben.

Anhang: Beweis von Lemma 1

Aus der Existenz einer lokalen Nutzenfunktion, aus der Definition von F_{y+y^*} und F und aus einer Taylor-Entwicklung der lokalen Nutzenfunktion nach dem ersten Argument folgt der Reihe nach:

$$
\begin{aligned}
V(F_{y+y^*}) - V(F_y) &= \int u(x,F_y)\,(dF_{y+y^*}(x) - dF_y(x)) + o(\| F_{y+y^*} - F_y \|) \\
&= \int u(y(x) + y^*(x), F_y) - u(y(x), F_y)\, dF(x) + o(\| F_{y+y^*} - F_y \|) \\
&= \int u_1(y(x), F_y)\, y^*(x)\, dF(x) + o(\| y^* \|_\infty) + o(\| F_{y+y^*} - F_y \|)
\end{aligned}
$$

[11]) Aus dem Dominanzprinzip folgt, daß die lokalen Nutzenfunktionen eine nichtfallende Funktion von x sind, vgl. *Machina* (1982) S. 296.

[12]) Ein monotoner Likelihood-Quotient wurde auch in FN 9 verwendet. Nach *Milgrom* (1981) kann der Prinzipal unter dieser Bedingung von einem besseren Ergebnis auf einen höheren Arbeitseinsatz des Agenten schließen.

Es bleibt zu zeigen, daß sich $o(\| F_{y+y^*} - F_y \|)$ durch $o(\| y^* \|_\infty)$ abschätzen läßt. Dies folgt aus:

$$\begin{aligned} \| F_{y+y^*} - F_y \| &= \int | F_{y+y^*}(x) - F_y(x) | \, dx \\ &\leqq \int F_y(x + \| y^* \|_\infty) - F_y(x - \| y^* \|_\infty) \, dx \\ &= 2 \, \| y^* \|_\infty \end{aligned}$$

q.e.d.

Literatur

Allais, M. (1953), Le comportement de l'homme rationnel devant le risque: Critique des postulates et axiomes de l'école Americain. In: Econometrica, 21, 503–543.

Holmström, B. (1979), Moral Hazard and Observability. In: Bell Journal of Economics, 10, 74–91.

Jewitt, I. (1988), Justifying the First-Order-Approach to Principal-Agent-Problems. In: Econometrica, 56, 1177–1190.

Kahneman, D., Tversky, A. (1979), Prospect Theory: An Analysis of Decisions under Risk. In: Econometrica, 47, 263–291.

Klingbeil, E. (1977), Variationsrechnung. Mannheim usw.

Machina, M. J. (1982), "Expected Utility" Analysis without the Independence Axiom. In: Econometrica, 50, 277–323.

Machina, M. J. (1987), Choice under Uncertainty: Problems Solved and Unsolved. In: Journal of Economic Perspectives, 1, 121–154.

Milgrom, P. (1981), Good News and Bad News: Representations Theorems and Applications. In: Bell Journal of Economics, 12, 380–391.

Rees, R. (1985), The Theory of Principal and Agent. In: Bulletin of Economic Research, 37, 3–26.

Rogerson, W. R. (1985), The First-Order-Approach to Principal-Agent-Problems. In: Econometrica, 53, 1357–1367.

Sugden, R. (1986), New Developments in the Theory of Choice under Uncertainty. In: Bulletin of Economic Research, 38, 1–24.

Weber, M., Camerer, C. (1987), Recent Developments in Modelling Preferences under Risk. In: OR Spektrum, 9, 129–151.

Zusammenfassung

Dieser Beitrag präsentiert eine Verallgemeinerung der Principal-Agent-Theorie, indem die Erwartungsnutzentheorie durch den allgemeineren Ansatz von Machina (1982) ersetzt wird. Dabei wird statt des umstrittenen Substitutionsaxioms lediglich eine viel weniger einschränkende Differenzierbarkeitsbedingung vorausgesetzt, das Risikoverhalten kann dann durch eine sogenannte „lokale" Nutzenfunktion charakterisiert werden. Es wird gezeigt, daß dann bei einer pareto-optimalen Risikoallokation die nun anhand der „lokalen" Nutzenfunktion zu berechnenden Grenzraten der Substitution übereinstimmen müssen. Im second-best Fall kann dagegen aus Motivationsgründen eine Abweichung von der pareto-optimalen Risikoallokation vorteilhaft sein. Auch hierzu wird ein entsprechend modifiziertes Ergebnis abgeleitet, welches das bekannte Resultat von Holmström (1979) als Spezialfall enthält.

Summary

Almost all research in the field of principal-agent-theory is done in the expected utility framework. A more general approach to decision making under risk has been provided by Machina (1982), in which the crucial independence axiom is replaced by a much weaker differentiability assumption. Risk aversion can than be described by the characteristics of a "local" utility function. This paper provides a corresponding generalization of the approach of Holmström (1979) to principal-agent-problems. First it is shown that an optimal risk allocation implies that the marginal rates of substitution, now expressed in terms of the "local" utility function, are identical for both principal and agent. In the presence of Moral Hazard, a deviation from an optimal risk allocation is necessary in order to set better incentives. A solution to this problem is derived, which contains the result of Holmström as a special case.

Dr. Hans Rau-Bredow, Knorrstraße 41, 80807 München.