Title
A Fresh Look at the Rotten Kid Theorem

Permalink
https://escholarship.org/uc/item/8z525458

Journal
Journal of Political Economy, 97(5)

Author
Bergstrom, Ted

Publication Date
1989-10-01

Peer reviewed
A Fresh Look at the Rotten Kid Theorem—and Other Household Mysteries

Theodore C. Bergstrom

University of Michigan

Gary Becker's "Rotten Kid theorem" asserts that if all family members receive gifts of money income from a benevolent household member, then even if the household head does not precommit to an incentive plan for family members, it will be in the interest of selfish family members to maximize total family income. I show by examples that the Rotten Kid theorem is not true without assuming transferable utility. I find a simple condition on utility functions that is necessary and sufficient for there to be the kind of transferable utility needed for a Rotten Kid theorem. While restrictive, these conditions still allow one to apply the strong conclusions of the Rotten Kid theorem in an interesting class of examples.

I. The Rotten Kid Theorem

The famous "Coase theorem" (1960) has a younger sibling—also from the streets of Chicago—called the "Rotten Kid theorem." This theorem seems to have been first stated by Gary Becker in "A Theory of Social Interactions" (1974) and continues to play a lively role in discussions of the theory of the family. Becker asserts that if a family has a head who "cares sufficiently about all other members to transfer general resources to them, then redistribution of income among members [of the household] would not affect the consumption of any member, as long as the head continues to contribute to all" (p. 1076). "The major, and somewhat unexpected, conclusion is that if a head exists, other members also are motivated to maximize family income and consumption, even if their welfare depends on their own consumption alone" (p. 1080). In A Treatise on the Family (1981), Becker restates the Rotten

© 1989 by The University of Chicago. All rights reserved. 0022-3808/89/9705-0001$01.50

1198
Kid theorem: "Each beneficiary, no matter how selfish, maximizes the family income of his benefactor and thereby internalizes all effects of his actions on other beneficiaries" (p. 183).

If it is generally correct, the Rotten Kid theorem must be one of the most remarkable and important results in the theory of incentives. For it tells us that a sufficiently benevolent household head would automatically internalize all the external effects that family members have on each other. Benevolent parents of intelligent, though selfish, children can breathe easier. In the family there will be no free riders or principal-agent problems. Elaborate incentive schemes and detection devices are unnecessary. All that is needed is to explain the Rotten Kid theorem to each family member and they will all (except possibly for a few irrational lapses) behave in the common interest. Not only would this be remarkable good news for parents, it would suggest a promising way of avoiding the incentive problems that be-devil firms and other social organizations. Shouldn't it be possible to find group incentive structures similar to those of families with benevolent heads?

This paper argues that the Rotten Kid theorem applies less generally than has sometimes been believed. But the news is not all bad. The theorem remains true for a restrictive, but still interesting, class of preferences and technologies. As with all theorems that have strong and interesting conclusions, it is worthwhile to find the limits of generality in which the Rotten Kid theorem applies and to illustrate these limits by examples.

Becker states his formal model and offers a proof in his 1974 paper: One family member, the household head, is benevolent toward all other family members and is rich enough so that he chooses voluntarily to give some money to each family member. There is a single consumption good, and \( X_i \) denotes the amount of consumption by family member \( i \). All family members except the head of the household are selfish and interested only in their own consumptions. The head of the household is altruistic, and his utility depends positively on the consumption of each household member. Therefore, one can write the head's utility as

\[
U(X_1, \ldots, X_n).
\]  

Let \( I_i \) be the income of family member \( i \) before any intrafamily transfers occur. Total expenditures on consumption for family members must satisfy

\[
\sum_i X_i = \sum_i I_i.
\]
If the head of the household is making transfers of income to all other household members, then after the transfers are made, the distribution of consumption in the family will be the one that maximizes the household head’s utility subject to the constraint that total family consumption equals total family income. That is, the allocation of consumption in the family will be the vector \((X_1, \ldots, X_n)\) that maximizes the utility function (1) subject to the budget constraint (2). This is like a standard problem in consumer theory in which the “goods,” \(X_i\), have prices of unity and income is \(\Sigma I_i\). Given the reasonable assumption that the \(X_i\)’s are all “normal goods” for the head of the household, it follows that each \(X_i\) is a monotonic increasing function of total family income. Therefore, any of the selfish children who has an opportunity to increase total family income, even if it is at the cost of reducing his or her own pretransfer income, will choose to do so. After all, the only way to increase one’s own consumption after gifts from the head of the household are accounted for is to increase family income. This proves the Rotten Kid theorem.

II. Some “Failures” of the Rotten Kid Theorem

The Case of the Lazy Rotten Kids

It is worth noticing, but not very surprising, that the Rotten Kid theorem fails to apply when there is asymmetric information. For example, let the income of each family member \(i\) be a function \(I_i(Y_i)\), where \(Y_i\) is an index of how hard \(i\) works. Let each child’s utility depend on his consumption of goods and leisure, where leisure of \(i\) is measured as \(1 - Y_i\). Suppose that the parent can observe incomes but not individual effort levels and that he chooses an allocation of consumption to maximize the utility function (1) subject to the constraint (2). Then a selfish child will have insufficient incentive to work since he will not receive the marginal product of an extra bit of effort but only this marginal product times the household head’s marginal propensity to spend on him.

But suppose that the head of the household is able to observe both the incomes and the levels of work effort by his children, and suppose that the head’s utility depends on the utilities of his children rather than just on their consumptions. Then will the Rotten Kid theorem hold? If the head of the household is able only to give gifts of money and does not direct the distribution of leisure, the answer is, “in general, no.”

The following example shows why. The head of the household has two children (named 1 and 2). The utility function of child \(i\) is a function

\[U_i = X_i(1 - Y_i),\]  
(3)
where $Y_i$ is $i$'s work effort. To simplify calculations, assume that the head does not work or consume himself, but he has an income $I_0$ and his utility function is
\[
U_0 = U_1^{1/2} + U_2^{1/2} = X_1^{1/2}(1 - Y_1)^{1/2} + X_2^{1/2}(1 - Y_2)^{1/2}.
\] (4)

Each child earns an income $I_i = wY_i$, and the head of the household is able to observe the work effort of each child. The head will make gifts to his children in such a way that their consumptions maximize (4) subject to the constraints
\[
X_1 + X_2 = I_0 + wY_1 + wY_2
\] (5)

and
\[
X_i > I_i \quad \text{for } i = 1, 2.
\] (6)

Following Becker, assume that the household head's income, $I_0$, is large enough so that he will always choose to make positive gifts to each child. Then the constraints (6) are not binding and can be ignored. Simple computations show that for the utility functions (3) and (4), the maximum of (4) subject to (5) occurs when $X_1/X_2 = (1 - Y_1)^2/(1 - Y_2)^2$. Using this fact and the budget equation (5), one can explicitly calculate each child's consumption as a function of his own work effort and of the other child's work effort. But all that one really needs to notice here is that each child's share of total family income is a decreasing function of his work effort. Therefore, the incentive problem is even worse in this example than in the case in which the household head does not observe individual effort.

Of course the head of the household might devise a better incentive scheme for his family. He could, for instance, pay "wages" for work effort by family members. This example simply shows that the Rotten Kid theorem does not allow us to claim that household incentive problems are solved automatically by the presence of a benevolent household head.

The Case of the Controversial Night-Light

Becker (1974, 1981) offers another provocative "application" of the Rotten Kid theorem. This example (slightly embellished) is the case of a husband who wishes to read at night although the light bothers his wife. The husband loves his wife and gives her consumption goods. He is aware that the night-light annoys her, and his desire for her to be happy makes him read less at night than he otherwise would. But he still uses the night-light more than she would like. The wife is entirely selfish. One day while the husband is away at work, an electri-
cian calls on the wife and offers to disconnect the night-light in such a way that the husband would think it was an accident and would not be able to use it again. The wife ponders whether she should accept the electrician’s offer. According to Becker, although she is selfish and although she dislikes the night-light, she should refuse. He reasons that if she has the night-light disconnected, the husband will be worse off. Even though he will not blame her for the loss of the night-light, the effect is like a loss in family income. If her utility is a “normal good” for him, the loss of family income will lead the husband to reduce his gifts of consumption goods to her so that her utility after the night-light is disconnected is lower than it had been before. Becker explains his reasoning as follows: “Perhaps a surprising implication of this analysis is that both are made better off when an altruist or his selfish beneficiary decides to eat with his fingers or read in bed. Since the utility of the altruist would be raised, he would increase his contribution to her by more than her initial harm from his actions or reduce his contribution by more than his initial harm from her actions” (1981, pp. 179–80).

Is Becker right? Let us look at an explicit model, using the apparatus of public goods theory. In this model, $X_h$ is goods consumption of the husband, $X_w$ is goods consumption of the wife, $I$ is family income, and $Y$ is the number of hours that the husband spends reading in bed. Following Becker, suppose that the husband is altruistic and the wife is selfish. Let the husband’s utility function take the form $U_h(u_h(X_h, Y), u_w(X_w, Y))$, where the wife’s utility function is $u_w(X_w, Y)$. In accordance with Becker’s story, $u_w$ is an increasing function of $X_w$ and a decreasing function of $Y$. The function $u_h(X_h, Y)$ is an aggregator function representing the husband’s “private preferences” over alternative combinations of night-light and goods consumption for himself, with the wife’s utility held constant. The function $U_h(\cdot, \cdot)$ is an increasing function of its two arguments, $u_h$ and $u_w$. The husband chooses the distribution of private consumption in the family subject to the family budget constraint.

Now consider an example in which the utility functions are

$$U_h(u_h, u_w) = u_h u_w^a,$$

where $0 < a < 1$ and where

$$u_h(X_h, Y) = X_h(Y + 1)$$

and

$$u_w(X_w, Y) = X_w e^{-Y}.$$

This example has the qualitative features of Becker’s discussion and displays strictly convex preferences for both people. One can substi-
tute (8) and (9) into (7) to write the husband's utility in terms of the decision variables, \( X_h, X_w, \) and \( Y \). This yields

\[
U_h = X_h X_w^a (Y + 1) e^{-aY}.
\]

(10)

If the husband determines the division of consumption between himself and his wife, he would seek to maximize (7) subject to the constraint

\[
X_h + X_w = I.
\]

(11)

Simple calculations show that the husband would choose to divide income so that \( X_w = aI/(1 + a) \). This means that his choice of how to divide the family income is completely independent of the amount of night-light, \( Y \). If the husband can choose \( Y \), he will choose \( Y = (1/a) - 1 > 0 \). On the other hand, since the wife's allowance of private goods is independent of \( Y \) and since her utility is a decreasing function of \( Y \), she would choose \( Y = 0 \) if she had her choice in the matter.

In this example, the outcome when the wife chooses the amount of reading is not Pareto optimal while the outcome when the altruistic husband chooses everything must be Pareto optimal. Nevertheless, the wife is better off when she chooses \( Y \) than when he does. Of course, the husband could bribe the wife to allow him to read, but there is not an automatic incentive for the wife to choose a Pareto-optimal amount of public good without an explicit bargain being struck.

*The Case of the Prodigal Son*

Lindbeck and Weibull (1988) discuss a problem that is similar to that of the lazy rotten kids and that arises in the allocation of consumption over time. The discussion has more ancient roots in the biblical parable of the prodigal son. Suppose that a child has a certain allotment of wealth in the first period, which he can either spend or save. In the second period of life he knows that he will receive a gift from a benevolent parent. If the parent can make no precommitment to punish profligate first-period behavior, the kid typically has an incentive to overspend in the first period.\(^1\)

I illustrate with a simplification of an example found in Lindbeck and Weibull. Consider a selfish kid (named \( k \)) with a benevolent parent. There are two time periods. The utility function of the kid is \( u_k(c^1_k, c^2_k) = \ln c^1_k + \ln c^2_k \), where \( c^t_k \) is the kid's consumption in period \( t \).

\(^1\) Lindbeck and Weibull are particularly interested in the case of prodigal parents, who spend too much in their youth in the expectation that their children will support them. The authors make an interesting case for forced savings programs, such as social security, based on this effect.
The utility function of the parent is

\[ U_p(c_p^1, c_p^2, c_k^1, c_k^2) = \ln c_p^1 + \ln c_p^2 + \alpha u_k \]
\[ = \ln c_1^p + \ln c_2^p + \alpha \ln c_k^1 + \alpha \ln c_k^2. \]

The child realizes that, whatever is done in the first period, the allocation chosen in the second period will be the one that maximizes the parent’s utility subject to the budget constraint that applies at that time. Since the parent’s utility function is just the log of a Cobb-Douglas function, familiar calculations will tell us that the utility-maximizing parent will choose an allocation that divides the family’s second-period wealth so that the fraction \(1/(1 + \alpha)\) goes to the parent and the fraction \(\alpha/(1 + \alpha)\) goes to the kid. The kid knows, therefore, that if he saves an extra dollar, he will increase second-period wealth by \(1 + r\) dollars, where \(r\) is the interest rate. But when we take into account the change in his parent’s giving that is induced, we find that the extra consumption he gains from this saving is only \([\alpha/(1 + \alpha)](1 + r)\). He will therefore consume too much in the first period and too little in the second period from the standpoint of Pareto efficiency.

Where Did the Rotten Kid Theorem Go Wrong?

I have shown examples in which the Rotten Kid theorem does not apply, at least in a straightforward way. Why does this happen? In each example, trouble arose when a second commodity appeared. In the case of the lazy rotten kids, introducing the private commodity, work effort (or, equivalently, leisure), caused the trouble. In the case of the prodigal son, it was the introduction of two dated commodities. In the case of the controversial night-light, it was the introduction of a public good, the night-light.

A good way to see what went wrong is to look at the “utility possibility frontier.” In the case of the night-light, consider the locus of feasible combinations of \(u_h\) and \(u_w\) that can be achieved by redistributing income, conditional on a given choice of \(Y\). From equations (8), (9), and (11), we find that this frontier is described by the equation

\[ \frac{u_h}{Y + 1} + u_w e^Y = I. \]  

(12)

In the case \(Y = 0\), one sees by substitution into (12) that the equation for the utility possibility frontier is just \(u_h + u_w = I\). In figure 1, this frontier is the line \(AB\). In the case \(Y = 1\), the utility possibility frontier becomes \((u_h/2) + u_w e = I\). This frontier is depicted by the line \(CD\). Notice that these two utility possibility frontiers are not “nested.” A change in \(Y\) changes the slope as well as the position of the utility possibility frontier. If the wife chooses \(Y\) before the husband chooses
the family distribution of income, then he must choose a point on the utility possibility frontier conditioned on her choice of \( Y \). Formally, the husband’s choice problem is to choose \( u_h \) and \( u_w \) to maximize (7) subject to (11). Since (7) is of the Cobb-Douglas form, he will choose

\[
u_h = \frac{Y + 1}{1 + a}, \quad u_w = \frac{ae^{-Y}}{1 + a}.
\]  

(13)

The wife can gain by choosing an amount of night-light services, \( Y \), different from what her husband would choose because by manipulating \( Y \) she can “twist” the utility possibility frontier in a way that is favorable to her. As is seen from (11), the smaller \( Y \) is in this example, the “cheaper” it is to supply \( u_w \) and the more “expensive” it is to supply \( u_h \). Consequently, as (13) shows, the smaller \( Y \) is, the more \( u_w \) the husband will choose. The selfish wife will want \( Y \) to be as small as possible.

Much the same thing happens in the case of the lazy rotten kid and the case of the prodigal son. In each case, if a rotten kid can commit himself to an action before the household head can commit himself to an incentive scheme, the kid will find it profitable to choose actions that are inefficient but that distort the utility possibility frontier in a way favorable to him.

### III. The Game Rotten Kids Play

The Rotten Kid theorem can be described succinctly as a description of equilibrium in a two-stage game. The players are a set of “family members,” including a benevolent “head.” In the first stage of
the game, each family member chooses an "action." The actions of the family members will, in general, influence each others' utilities directly and will also affect the total wealth available for consumption by family members. In the second stage, the head decides how to distribute income among the family members. The Rotten Kid theorem claims that the subgame perfect equilibrium for this game is the head's most preferred outcome. To say this in another way, the theorem is true if a benevolent head can achieve his most preferred outcome, even though he is not able to "precommit" to incentive plans that require punishments or rewards that he would not choose to undertake ex post.

In order to sharpen the analysis and to relate this discussion to the large body of recent work on incentive theory, it is useful to tell the story of the rotten kid in the vocabulary that has become standard in the theory of social decision mechanisms (Groves 1982; Laffont and Maskin 1982). There are $n$ agents (the kids). In the first period, each agent, $i$, chooses an action $a_i$ from a closed bounded set $A_i$ of possible actions. In the second period, each agent $i$ will receive an allotment of money, $t_i$. Each agent has a continuous utility function $U_i(a, t_i)$, where $a = (a_1, \ldots, a_n)$ is a list of the actions taken by all agents and $t_i$ is the money $i$ receives. The total amount of money available to be distributed in the second period is a function $I(a)$. There is a planner (the household head in the rotten kid story) who observes the actions $a$ and then assigns an income $t_i(a)$ to agent $i$ in such a way that $\Sigma_{j=1}^{n} t_j(a) = I(a)$.

For any $a$, the utility possibility set conditional on $a$ is the set of distributions of utility, $UP(a)$, that can be achieved if $a$ is the list of actions chosen by the agents and if the amount of money available to be distributed is $I(a)$. That is, $UP(a)$ is the set

$$\left\{(u_1, \ldots, u_n) | (u_1, \ldots, u_n) \leq (u_1(a, t_1), \ldots, u_r(a, t_n)), (t_1, \ldots, t_n) \geq 0 \text{ and } \sum_{j=1}^{n} t_j = I(a)\right\}.$$

In the first stage of the game, each agent $i$ chooses $a_i$. After observing the vector of choices, $a$, the planner chooses the distribution of money income. The planner has preferences over alternative distributions of utility among the agents and chooses the income distribution he likes best from the set $UP(a)$.

**Definitions.** We shall refer to the game described above as the "Game Rotten Kids Play." We shall say that rotten kids are well behaved if the subgame perfect equilibrium for the Game Rotten Kids Play is the same as the Pareto optimum, which is the household head's most preferred outcome.
If rotten kids are well behaved, the incentive problem for the group becomes a problem of the kind that Marschak and Radner (1972) call "team theory." All organizational problems are merely problems of coordination and not problems of conflicting objectives.\footnote{These problems of coordination could be solved, e.g., by having the planner instruct all the agents on which \textit{a} to choose. Well-behaved rotten kids would all find it in their interest to obey the planner's instructions.}

IV. Can the Rotten Kid Theorem Be Rehabilitated?

Transferable Utility and the Rotten Kid Theorem

The term "transferable utility" appears to have originated in game theory, but the idea that it represents is familiar to all economists. There is transferable utility in an \( n \)-person economy if whenever a utility distribution \((u_1, \ldots, u_n)\) is possible, any distributions of utility in which the sum of the utilities is the same as \(\sum_i u_i\) (subject to some lower bound on each of the \(u_i\)'s) are also possible. More formally, there is transferable utility if there exists some utility representation of individual preferences such that the utility possibility frontier is a set of the form

\[
\left\{ (u_1, \ldots, u_n) \succeq (\bar{u}_1, \ldots, \bar{u}_n) \left| \sum_i u_i = I \right. \right\}
\]

for some \(I\) and some \((\bar{u}_1, \ldots, \bar{u}_n)\). Geometrically, this means that the utility possibility frontier is a simplex. When it is a simplex, it is natural to speak of transferable utility because, starting from a Pareto optimum, any "transfer of utilities" that preserves the sum of utilities (and does not lower anyone's utility below the minimal amount \(\bar{u}_i\)) represents another feasible and Pareto-optimal distribution.

Each choice of the vector of actions by the kids determines a "conditional utility possibility frontier," \(UP(a)\). Different points on this frontier are reached by different distributions of money income. For some utility representation of preferences, if the conditional utility possibility frontiers are all simplexes, then Becker's proof for the single-commodity case applies directly. The actions of the kids simply move the boundary of the utility possibility simplex inward or outward in parallel fashion. The role of the benevolent head of the household is to choose a point on the utility possibility frontier. The actions of the kids determine the position of this frontier. As long as the head of the household is benevolent and treats the utility of each of the kids as a normal good, the kids will all agree that an outward shift of the utility
possibility frontier is a "good thing" and an inward shift is a "bad thing."°

We shall say that there is conditional transferable utility if there
exists some utility representation for each consumer such that the
conditional utility possibility frontier, UP(a), is a simplex for every a.
From our discussion, we are able to claim the following result.

PROPOSITION 1. A Rotten Kid theorem.—Rotten kids are well behaved
when there is conditional transferable utility.

Utility Functions That Imply Conditional
Transferable Utility

A rather large and interesting class of preferences will yield transfer-
able utility. A simple condition on the form of utility functions that is
both necessary and sufficient follows.

PROPOSITION 2. In the Game Rotten Kids Play, there is conditional
transferable utility if and only if the preferences of every consumer
can be represented in the functional form \( u_i(m_i, a) = A(a)m_i + B_i(a) \).

Proof. We need to show that the existence of utility representations
of the form \( u_i(m_i, a) = A(a)m_i + B_i(a) \) is necessary and sufficient for
the set \( UP(a) \) to be a simplex for every \( a \). For sufficiency, notice that
for any \( a \), \( UP(a) \) is just the set of utilities distributions attainable by
distributing \( I(a) \) among the agents. This set is

\[
\{(u_1, \ldots, u_h) \geq (B_1(a), \ldots, B_h(a)) | \Sigma u_i \leq A(a)I(a) + \Sigma B_i(a)\},
\]

which is a simplex.

Necessity follows from a standard result in the theory of functional
equations, which is known as Pexider's functional equation (Aczél
1966). Let the functions \( f_i(x_i, y) \) have the property that whenever \( \Sigma x_i = \Sigma x'_i \) then \( \Sigma f_i(x_i, y) = \Sigma f_i(x'_i, y) \). Then if these functions are all
continuous, they must be of the form \( f_i(x_i, y) = A(y)x_i + B_i(y) \). If
there is transferable utility, then the utility functions \( u(m_i, a) \) must
have this property and therefore must be of the functional form
claimed in proposition 1. Q.E.D.

The Rotten Kid Theorem with Household Public Goods:
The Night-Light Revisited

A good way to model family externalities is to use the standard public
goods model. Let there be a vector \( Y \) of public goods that is jointly

° A benevolent household head is not the only household choice mechanism that has
this property. The Nash bargaining solution in a family in which everyone with any
bargaining power is nonmalevolent and someone with bargaining power is actually
benevolent would also do this.
determined by the actions of family members and let the total amount of money to be divided among agents be a function $I(Y)$. If the utility functions are of the form $u_i(m_i, Y) = A(Y)m_i + B_i(Y)$, then according to proposition 2, there will be conditional transferable utility. To help us understand the restrictions imposed by this form, consider some special cases within this class. If $A(Y) = 1$, then utility is of the quasi-linear form, $u_i(m_i, Y) = m_i + f_i(Y)$. In this case, each consumer’s marginal rate of substitution between private goods and public goods depends only on his consumption of public goods and not on his consumption of private goods. Another special case is one in which $A(Y)$ depends on $Y$, but $B_i(Y)$ is a constant for each $i$. In this case, preferences are identical for all consumers and Cobb-Douglas in $m_i$ and the aggregator $A(Y)$. Allowing both $A(Y)$ and $B_i(Y)$ to vary in $Y$ permits a rich variety of cases. For example, income elasticities may be as large or small as one wishes, preferences need not be homothetic or identical, and there can be components of the vector $Y$ that are “goods” for some consumers and “bads” for others. Nevertheless, the assumption that preferences can be represented in this form is restrictive, as one sees from the counterexample for the case of the controversial night-light.

In the case of the night-light, the wife’s action, $a$, is to choose an amount, $Y$, of a household public good, the night-light. After she has made her choice, the husband chooses the distribution of income in the family. Suppose that the husband and wife have private utility functions of the form $u_w(m_w, Y) = A(Y)m_w + B_w(Y)$ and $u_h(m_h, Y) = A(Y)m_h + B_h(Y)$ and suppose that the husband’s benevolence is represented by a utility function of the form $U_h(u_h(m_h, Y), u_w(m_w, Y))$. Let $u_h$ be an increasing function of $Y$ and $u_w$ a decreasing function of $Y$. Then for any choice of $Y$, the set of combinations $(u_h, u_w)$ that are attainable by alternative distributions of income is bounded by a straight line segment with a slope of $−1$. Choosing a different $Y$ will simply shift this line segment in or out in parallel fashion. If the husband treats the wife’s utility as a “normal good,” then it will always be in her interest to choose the $Y$ that shifts this line segment furthest out. Thus she would choose the same Pareto-optimal amount of

---

4 This result was first reported in Bergstrom and Cornes (1981, 1983). They showed that this condition is sufficient and (subject to technical qualifications) also necessary for there to be transferable utility in a public goods model. They also show that this is the most general class of cases for which the optimal amount of public goods is independent of the distribution of private goods. Bergstrom and Varian (1985) find necessary and sufficient conditions for transferable utility in a pure exchange economy with private goods.

5 If $A(Y) > 0$ for all $Y$, then a function of this form will be quasi-concave if and only if $\alpha(Y) = 1/A(Y)$ is a convex function and $\beta_i(Y) = B_i(Y)/A(Y)$ is concave in $Y$. 
night-light that her husband would choose if he were calling all the shots.⁶

*The Rotten Kid Theorem and the Hicks Composite Commodity Theorem: Motivating Lazy Rotten Kids and Restraining Prodigal Kin*

In the example of lazy rotten kids, the Rotten Kid theorem breaks down when a second private good, leisure, is introduced. But even with many private goods, there is an easy application of the Rotten Kid theorem. No matter how many private goods there are, if all family members are price takers for each of these goods, then we can, by the Hicks composite commodity theorem, treat “income” as the single private good. If there is only one good, then, trivially, there is transferable utility. In a fixed-price environment, money income can serve as the utility function for each child. Redistributions of income preserve the sum of utilities, which is what is required for transferable utility.

This fact allows one resolution of the case of the lazy rotten kids. If each kid can earn some fixed hourly wage in the marketplace, then each kid can be made a price taker for two marketable goods, leisure and consumption. The single-commodity model applies if the head of the household makes monetary gifts that determine each child’s “full income” (his income if he works all the time) and then allows each child to “buy” his preferred bundle of leisure and goods with this income. In the same spirit, the case of the prodigal son can be reduced to a one-commodity model if the kid has access to competitive markets for borrowing and lending and if the head of the household commits to making a single lump-sum gift. In each case, the kid would be allowed to make his own allocative choices with the budget he is given. The market prices that he faces enforce the same efficient outcome that the parent would choose if the parent acted as dictator.

But the theorem, interpreted in this way, requires that the household head be able to precommit to a strategy in which the kid is rewarded for work effort. As our earlier examples showed, if the kids were able to commit to a level of effort or to a first-period consumption before the benevolent household head decides how to distribute income, the outcome would not in general be Pareto efficient. In the absence of parental precommitment, the kids will be able to manipulate their indulgent parent by working too little or spending too much in the first period.

⁶ In the Appendix, I work out an example for specific functional forms.
The Rotten Kid Theorem with Several Private Goods and No Parental Precommitment

Where there is more than one private good, there is also a special class of utility functions for which the Rotten Kid theorem applies without precommitment by the household head. In the story of the lazy rotten kids and in the story of the prodigal son, there are two private goods and a budget constraint. The kids are able to choose their consumptions of one of the goods before the parent decides what the kids’ incomes will be. In the notation for the Game Rotten Kids Play, each kid chooses an action \( a_i \) that is his consumption of one of the goods: leisure in the case of the lazy rotten kids, first-period consumption in the case of the prodigal. Let \( a \) be the vector \((a_1, \ldots, a_n)\) and \( m_i \) be the money income given to \( i \). According to proposition 2, there will be transferable utility if and only if the utility functions of all kids are of the form \( u_i(m_i, a_i) = A(a)m_i + B_i(a) \).

Notice that \( A(a) \) has no subscript. This piece of \( u_i \) must be the same function for all consumers. But when the \( a_i \)'s are pure private goods, it must be that \( u_i \) is independent of \( a_j \) for all \( j \neq i \). The only way in which this is possible is if \( A(a) \) is a constant. Without loss of generality, we can let this constant be one. Therefore, where the \( a_i \)'s are private goods, there is transferable utility if and only if the utility functions take the quasi-linear form \( u_i(m_i, a_i) = m_i + B_i(a) \).

For example, suppose that each kid has a utility function, \( u(m_i, Y_i) = m_i + B_i(Y_i) \), where \( m_i \) is \( i \)'s money income and \( Y_i \) is \( i \)'s work effort, and suppose that family income is a function \( I(Y_1, \ldots, Y_n) \) of the work efforts of the kids. The household head, after observing the work effort of each kid, chooses to distribute money income in the family in the way that maximizes his utility \( U(u_1, \ldots, u_n) \) over the set of feasible utility distributions. But with utility functions of this form, the utility possibility frontier is the set \( \{(u_1, \ldots, u_n) | \Sigma u_i = I(Y_1, \ldots, Y_n) + \Sigma B_i(Y_i) \} \). Therefore, if the household head treats the utility of each kid as a normal good, then each kid will benefit from any change in activities that increases \( I(Y_1, \ldots, Y_n) + \Sigma B_i(Y_i) \). So all the kids will choose their work effort so as to maximize this sum. In particular, this means that each kid will work up to the point at which the marginal contribution of an extra unit of effort to total family income is exactly equal to his marginal rate of substitution between income and leisure. That is, \( \left[ \frac{\partial I(Y_1, \ldots, Y_n)}{\partial Y_i} \right] + B_i(Y_i) = 0 \).\(^7\)

Applying this special utility form to the case of the prodigal son, we have kids whose utility functions are linear in second-period consumption, that is, \( u_i(c_i^1, c_i^2) = c_i^2 + B_i(c_i^1) \). Just as in the case of the lazy

\(^7\) An example for a specific functional form of utility is worked out in the Appendix.
rotten kids, the payment scheme implied by the parent's benevolent utility function will be one that induces efficient behavior. As our earlier example shows, this does not happen in general. With other kinds of utility functions, efficiency can be achieved only if the parent is able to precommit to a wage scheme that is not subgame perfect.

It is nice to know that there are nontrivial examples in which the Rotten Kid theorem applies to slothful and prodigal children. But the class of utility functions for which it applies is very restrictive. In particular, the assumption of quasi-linear utility implies that the lazy rotten kid's income elasticity of demand for leisure is zero. Similarly, for the Rotten Kid theorem to work, the prodigal son's income elasticity of demand for first-period consumption has to be zero.

V. Is Conditional Transferable Utility Necessary for the Rotten Kid Theorem?

We have proved a Rotten Kid theorem that assumes conditional transferable utility. We have also shown that a necessary and sufficient condition for conditional transferable utility is that utility functions are of the form $A(a)m_i + B_i(a)$. The question remains whether conditional transferable utility is necessary for a Rotten Kid theorem to hold. With strong special assumptions about technology or about the tastes of the head of the household, it is possible to find cases in which rotten kids are well behaved, even without transferable utility. Some special assumptions of this kind may turn out to be reasonable and interesting. If so, there are "rotten kid theorems" that apply beyond the domain we explore. Here we investigate the assumptions on preferences needed so that rotten kids will be well behaved, regardless of the technology, and with no stronger assumptions on the household head's preferences than that he treats each kid's utility as a normal good. It turns out that, subject to technical conditions and an assumption that "money is important enough," this requires conditional transferable utility.

The condition that "money is important enough" means that for some reference vector of activities, $a_0$, and for any other vector $a$, there is an amount of money that will exactly compensate any consumer for the difference between the vector $a_0$ and the vector $a$. The technical conditions are that preferences are representable by a utility function that has a positive derivative with respect to income. These conditions are stated more formally in the following assumption.

Assumption M. Preferences of each kid $i$ are representable by a

---

8 The assumption of differentiability could be dispensed with in favor of strict monotonicity and continuity, but differentiability makes exposition much easier.
function \( u_i(m_i, \mathbf{a}) \) that has a positive derivative with respect to \( m_i \) for all \( m_i > 0 \) and all feasible vectors \( \mathbf{a} \). There is some vector of actions \( \mathbf{a}_0 \) such that for every kid \( i \), all incomes \( m_i \), and all feasible \( \mathbf{a} \), there exists income \( m'_i \) such that \( u_i(m'_i, \mathbf{a}_0) = u_i(m_i, \mathbf{a}) \).

Assumption M, alone, is by no means sufficient to imply transferable utility. Indeed our examples of lazy rotten kids, night-lights, and prodigal sons all have utility functions that satisfy condition M but do not have transferable utility. The main result of this section is the following proposition.

**Proposition 3.** If assumption M holds and if the rotten kids are well behaved for all technologies and for any benevolent household head who treats all the kids as normal goods, then there must be conditional transferable utility and utility functions must be of the form \( A(\mathbf{a})m_i + B_i(\mathbf{a}) \).

We prove the result for the case of two kids, where the geometry allows an easy, intuitive proof. Extension to higher dimensions is not difficult, but the exposition is tedious. Let \( UP(M, \mathbf{a}) \) be the utility possibility frontier that can be achieved by distributions of the total income \( M \) when the vector of activities is \( \mathbf{a} \). Suppose that the Rotten Kid theorem holds for any technology and for any benevolent household head for whom each kid's utility is a normal good for the head. Then it must be that there are no \( (M, \mathbf{a}) \) and \( (M', \mathbf{a}') \) such that the utility possibility frontier \( UP(M, \mathbf{a}) \) crosses the frontier \( UP(M', \mathbf{a}') \). For if the two curves crossed, we can always construct a technology and benevolent preferences for the parent so that without parental precommitment, at least one of the kids will act against the interests of the parent.

Figure 2 justifies this last claim. Let \( UU \) mark the set \( UP(M, \mathbf{a}_0) \) and \( U'U' \) mark the set \( UP(M', \mathbf{a}') \). We can always draw indifference curves for a benevolent parent in such a way that the parent's best choice from \( UP(M', \mathbf{a}') \) is at the point \( X \), where the curves cross, and such that the parent prefers some utility distribution \( Y \) from \( UP(M, \mathbf{a}) \) to the distribution \( X \). But one of the kids, call him kid 2, will prefer the distribution \( X \) to the distribution \( Y \). Since we have made no special assumptions about technology, we can always construct an example in which kid 2 can choose an action that changes the vector of actions from \( \mathbf{a} \) to \( \mathbf{a}' \) and family income from \( M \) to \( M' \). To do so would benefit the kid but would make the household head and the other kid worse off. Therefore, if rotten kids must be well behaved, it follows that the curves \( U(M, \mathbf{a}) \) and \( U(M', \mathbf{a}') \) do not cross each other.

We show that if assumption M holds and if the utility possibility frontiers never cross, then there exist utility functions for the kids of the form \( A(\mathbf{a})m_i + B_i(\mathbf{a}) \). Given assumption M, we can define a function \( u_i^*(m_i, \mathbf{a}) \) such that for all \( \mathbf{a} \) and \( m_i \), \( u_i(u_i^*(m_i, \mathbf{a}), \mathbf{a}^0) = u_i(m_i, \mathbf{a}) \). In
words, $u_i^*(m_i, a)$ is the amount of expenditures that $i$ would need to be as well off when the activity vector is $a^0$ as he is when the activity vector is $a$ and his expenditure is $m_i$. Since preferences are monotonic in $m_i$, it must be that the function $u_i^*(m_i, a)$ represents $i$'s preferences. Also, it is evident from the construction that for all $m_i$, $u_i^*(m_i, a^0) = m_i$. Therefore, for any $M$, the utility possibility frontier, $UP(M, a)$ is just a straight line segment with a slope of $-1$.

Consider any vector of actions $a$ and income $M$. From assumption M, it follows that any point on $UP(M, a)$ is also a point on $UP(M', a_0)$ for some $M'$. But the only way that the utility possibility frontiers $UP(M, a)$ and $UP(M', a_0)$ can meet without crossing is that they have the same slope at the point where they meet. Since movements along either frontier are accomplished only by redistributing income, the slope of $UP(M, a)$ at the point $(u(m_1, a), u(m_2, a))$ must equal the ratio $[\partial u_1^*(m_1, a)/\partial m_1]/[\partial u_2^*(m_2, a)/\partial m_2]$. But this slope must also equal the slope of $UP(M', a_0)$, which is $-1$. Therefore, it must be that for all positive $m_1$ and $m_2$

$$\frac{\partial u_1^*(m_1, a)}{\partial m_1} = \frac{\partial u_2^*(m_2, a)}{\partial m_2}. \quad (14)$$

Separately integrating the two sides of (14), one sees that the only way that this can happen is if the $u_i^*$'s are of the form $u_i^* = A(a)m_i + B_i(a)$. From proposition 2, it follows that there is conditional transferable utility. This proves proposition 3 for the case of two kids.
VI. A Rotten Kid Theorem for Nice Kids

We have so far dealt with a family that has only one unselfish member, the head. Now consider the case in which several family members are concerned about each others' welfare. Becker calls this the case of "reciprocal altruism" and offers some useful suggestions of how the Rotten Kid theorem would extend to this case.

There are at least two reasonable ways of modeling consumers who are benevolent toward each other. The first way is to treat individual preferences over allocations as the fundamental concept. Let a household allocation be represented as \((X_1, \ldots, X_n, Y)\), where \(X_i\) is the vector of private goods consumed by family member \(i\) and \(Y\) is a vector of household public goods and of externality-generating activities by family members. Suppose that the preferences of any individual over allocations are weakly separable between individuals and respond benevolently to the level of private preferences of each \(i\). Then \(i\)'s preferences can be represented by a utility function of the form

\[
U_i(u_1(X_1, Y), \ldots, u_n(X_n, Y)),
\]

where \(u_i(X_i, Y)\) represents \(i\)'s preferences over alternative commodity bundles in which the \(u_j\) for \(j \neq i\) are held constant. We shall call the preferences represented by \(u_i\) "private preferences"; we shall call the preferences represented by \(U_i\) "\(i\)'s total preferences." If all family members are nonmalevolent, then \(U_i\) is a nondecreasing function of each of the \(u_j\)’s. If there is a household head who decides the allocation of resources in the family and if the private utility of each family member is a normal good for the head, then our previous analysis applies without modification. Anything that shifts the private utility possibility frontier outward in a parallel fashion will lead the head to choose an outcome with higher private utilities for everyone. But since everyone is assumed to be nonmalevolent, it must be that if private utility, \(u_i\), increases for everyone, then so will total utility \(U_i\). Therefore, the Rotten Kid theorem will apply if there is conditional transferable utility.

There is a possible snag. What right have we to assume, in a family with complicated reciprocal entanglements of benevolence, that family members' private utilities will be "normal goods" for the head of the household? It is helpful to look at a second way of modeling benevolence. Suppose that each consumer’s total utility, \(U_i(X, Y)\), is determined by his private utility, \(u_i(X_i, Y)\), and by the total utilities, \(U_j(X, Y)\), of other family members. Then the utility function of family member \(i\) is written as

\[
U_i = U_i(U_1(X, Y), \ldots, u_i(X_i, Y), \ldots, U_n(X, Y)).
\]
Let $\mathbf{A}$ be the $n \times n$ matrix with zeros on the diagonal and with $A_{ij} = \partial U_i / \partial U_j$ for $i \neq j$. According to the implicit function theorem (given sufficient smoothness), if $\mathbf{I} - \mathbf{A}$ is nonsingular, the system of equations (16) locally determines a system of equations of the form (15). With assumptions a little stronger than the assumption that $\mathbf{I} - \mathbf{A}$ is nonsingular over the relevant domain (see Nikaido 1968), the system of equations (16) has a global inverse of the form (15). When this is the case, we know that representability in either form implies representability in the other form, and we even have a recipe for moving from one form of representation to the other.

Although preferences can be represented in the form (15), it may be more difficult to decide what is reasonable to assume about these utility functions than to decide what is reasonable to assume about utility functions in the form (16). For example, one may be willing to assume that the total utilities $U_j$ are normal goods for consumer $i$. Then an increase in household wealth would make consumer $i$ want to give all the other $i$'s increases in total utility. But this does not necessarily imply that private utilities, $u_i$, are also normal goods.

There is a nice class of cases in which the assumption that total utilities $U_i$ are normal goods implies that private utilities are normal goods as well and in which transferable utility in private utilities implies transferable utility in total utilities. The case is this. Suppose that each $i$ has preferences representable in the additively separable form

$$U_i = u_i + \sum_{j \neq i} A_{ij} U_j,$$

where $A_{ij} \geq 0$ for all $i$ and $j$. Suppose further that the row sums of the matrix $\mathbf{I} - \mathbf{A}$ are all positive. Nonnegativity of the $A_{ij}$'s simply means nonmalevolence. Positivity of the row sums means that individuals are “partially selfish” in the sense that they weigh their own utilities more heavily than the sum of the weights they place on others. Under these assumptions, the theory of dominant diagonal matrices applies and tells us that the matrix $\mathbf{B} = (\mathbf{I} - \mathbf{A})^{-1}$ exists and is nonnegative in every entry. It follows from (17) that

$$U_i = \sum_j B_{ij} u_j$$

for all $i$.

Equation (18) is just enough to allow us to apply the Rotten Kid theorem. Since the $U_i$'s are related by a linear transformation of the $u_i$'s, the utility possibility frontier expressed in terms of the $U_i$'s will be a simplex if and only if the utility possibility frontier expressed in terms of the $u_i$'s is a simplex. Furthermore, an outward shift in the
latter frontier will imply an outward shift in the former. It follows that if there is transferable utility in \( u_i \)'s, then there will be transferable utility in the \( U_i \)'s. Therefore, if the head of the household regards the \( U_i \)'s as normal goods, the Rotten Kid theorem applies.

VII. Final Remarks

The kinship between the Rotten Kid theorem and the Coase theorem is intriguing. As Cooter (1982) points out, Coase refrained from stating a general “theorem” underlying his examples. Quite wisely, it seems, Coase chose to present the profession with a “Coase insight” and left formal details to be worked out by others. Statements of the Coase theorem appear in the works of others. For example, Zerbe (1976) proposes the following statement of the Coase theorem: “In a world of perfect competition, perfect information, and zero transaction costs, the allocation of resources will be efficient and invariant with respect to legal laws of liability” (p. 29). This statement of the Coase theorem has two distinct claims. One is that “in the absence of transactions costs and with perfect information,” private bargaining will lead to an efficient outcome regardless of the legal assignment of property rights. The second is that the “allocation of resources” will be invariant to the assignment of property rights. The general validity of both claims has been much disputed in the literature.\(^9\) There is little dispute, however, that Coase's examples have profoundly improved the profession’s understanding of the possibilities for private “bargaining in the shadow of the law.”

The Coasian claim of “invariance of resource allocation with respect to the assignment of property rights” has been shown not to be generally true. It is easy to show examples in which the Pareto-efficient level of an externality-generating activity depends on how utility is distributed.\(^10\) But the point on the utility possibility frontier that is selected depends on the initial allocation of property rights. As it turns out, the conditions that allow the Rotten Kid theorem to be salvaged are precisely the conditions needed for Coasian invariance to obtain within the family. The result that the Rotten Kid theorem implies Coasian invariance is almost immediate from the statements of the two results.

We see that the two Chicago siblings prosper in similar environments. There is, I think, another and perhaps more important family resemblance. Each “theorem” is at least as useful as a heuristic genera-


\(^10\) One instance is the example in this paper in the case of the night-light.
tor of insight as it is as a formal proposition. Anyone who only reads about the Coase theorem without reading Coase’s insightful discussion of examples has missed the best of the story. Much the same can be said of the Rotten Kid theorem. The richness of insight that Becker has to offer about the economics of the family is only partially captured in formal analysis of rotten kid theorems. Many insights into the economics of the family and the economics of incentives that have escaped formal treatment in this paper are to be found in Becker’s discussion and examples in the *Treatise on the Family* and in his 1974 article.

Appendix

Two Specific Examples

A Happy Outcome in the Case of the Night-Light

I promised to work out an explicit example of the case of the night-light, where \( u_h \) and \( u_w \) are both of the functional form \( A(Y)m_i + B_i(Y) \). Let \( u_h(m_h, Y) = Ym_h + Y \) and \( u_w(m_w, Y) = Ym_w - Y^2 \) and suppose that \( U_h(u_h, u_w) = u_hu_w^a \). Let the wife choose \( Y \) and suppose that the husband has a fixed amount of money income, \( I \), to allocate between himself and his wife. He decides how to allocate this money after he observes the \( Y \) that his wife has chosen. Since it must be that \( m_h + m_w = I \), we see that given any choice of \( Y \), the set of vectors \((u_h, u_w)\) that can be achieved by alternative distributions of money income is just \([(u_h, u_w)|u_h + u_w = IY + Y - Y^2]\). Given the Cobb-Douglas form of his utility function in the variables \( u_h \) and \( u_w \), the husband will choose to distribute money income in such a way that \( u_w = [a/(1 + a)](IY + Y - Y^2) \) and \( u_h = [1/(1 + a)](IY + Y - Y^2) \). Since her ultimate utility level is proportional to \( IY + Y - Y^2 \), it is in the interest of the wife to choose \( Y \) so as to maximize this expression. This is a Pareto-optimal choice of \( Y \) and is also the quantity the husband would choose if he were allowed to choose \( Y \) as well as the distribution of income.

Some Well-Behaved, Lazy Rotten Kids

Here I show in detail, for specific quasi-linear utility functions, how the Rotten Kid theorem works out for lazy rotten kids. Suppose that there are two kids, each of whom has a utility function \( u(m_i, Y_i) = m_i + (1/Y_i) \). The household head’s utility function is \( u_1u_2 \) and family income is \( I_0 + w(Y_1 + Y_2) \). If the kids choose work efforts \( Y_1 \) and \( Y_2 \), then the utility possibility frontier is defined by \( u_1 + u_2 = I_0 + w(Y_1 + Y_2) + (1/Y_1) + (1/Y_2) \). Maximizing his utility subject to this constraint, the household head will choose an income distribution so that \( u_1 = .5[I_0 + w(Y_1 + Y_2) + (1/Y_1) + (1/Y_2)] \). Since the utility function of kid 1 is \( u(X_1, Y_1) = X_1 + (1/Y_1) \), the head accomplishes this by giving 1 an income \( X_1 \), where \( X_1 + (1/Y_1) = .5[I_0 + w(Y_1 + Y_2) + (1/Y_1) +

\[11\] Solving this out for \( m_w \), we find that the husband will choose \( m_w = [1/(1 + a)][a(I + 1) + Y] \), where \( Y \) is the wife’s choice of \( Y \).
(1/Y_2)]. This implies that \( X_1 = 0.5[wY_1 - (1/Y_1)] + 0.5I_0 + 0.5wY_2 \). Realizing that this is the way in which his income is related to work efforts, kid 1 knows that if the levels of effort are \( Y_1 \) and \( Y_2 \), then his utility will be \( X_1 + (1/Y_1) = 0.5[wY_1 + (1/Y_1)] + 0.5I_0 + 0.5wY_2 \). The best he can do for himself is to choose \( Y_1 \) to maximize \( wY_1 + (1/Y_1) \). Similar reasoning shows that the best that kid 2 can do for himself is to choose \( Y_2 \) so as maximize \( wY_2 + (1/Y_2) \). Therefore, each kid is motivated to work to the point at which his wage rate equals his marginal rate of substitution between consumption and leisure.

References


