Bequests as Signals: An Explanation for the Equal Division Puzzle

B. Douglas Bernheim

Stanford University

Sergei Severinov

University of Wisconsin—Madison and Duke University

In the United States, more than two-thirds of decedents with multichild families divide their estates exactly equally among their children. In contrast, gifts given before death are usually unequal. These findings challenge the validity of existing theories regarding the determination of intergenerational transfers. In this paper, we develop a theory that accounts for this puzzle based on the notion that the division of bequests provides a signal about a parent's altruistic preferences. The theory can also explain the norm of unigeniture, which prevails in other societies.

I. Introduction

In the United States, more than two-thirds of testate decedents with multichild families divide their estates exactly equally among their chil-

Seminar participants at the National Bureau of Economic Research, Stanford University, the University of Wisconsin—Madison, University of California at Los Angeles, Washington University in St. Louis, Pennsylvania State University, and the Econometric Society meetings provided helpful comments. We thank Katherine Carman for able research assistance and the National Science Foundation for financial support through grant SBR95-11321.

[*Journal of Political Economy,* 2003, vol. 111, no. 4]

dren.[1] This finding challenges the validity of existing theories regarding the determination of intergenerational transfers. If bequests reflect altruism (as in Barro [1974]) or intrafamily exchange (as in Bernheim, Shleifer, and Summers [1985]), the optimal division of an estate should vary with the characteristics of the children as well as with the attitudes and preferences of the parent, and equal division should be a "knife-edge" case.[2] If bequests are accidental (because the length of life is uncertain and annuity markets are imperfect), then equal division might reflect indifference concerning the division of assets.[3] However, indifference cannot account either for the prevalence of equal division among those who go to the trouble of writing wills or for the widespread and apparently deliberate inequality of gifts (see, e.g., Dunn and Phillips 1997; McGarry 1998). Likewise, if parents simply feel that fairness requires equal division of bequests, the same principle should apply to gifts. These observations give rise to the "equal division puzzle."

The absence of a coherent theoretical explanation for equal division represents a serious gap in the literature.[4] Bequests feature prominently in theoretical and empirical discussions of capital accumulation, fiscal policy, income distribution, and other issues (see, e.g., Barro 1974; Bernheim and Bagwell 1988). Moreover, the altruistic model of bequests is often invoked to justify the practice of studying models with infinite-lived agents ("dynastic" families). As long as one of the most notable empirical regularities concerning bequests remains unexplained, econ-

[1] According to Wilhelm (1996), 68.6 percent of all decedents with multichild families divide their estates exactly equally among their children, and 76.6 percent divide their estates so that each child receives within 2 percent of the average inheritance across all children. Two studies by Menchik (1980, 1988) place the frequency of exact equal division at, respectively, 62.5 percent and 84.3 percent. Tomes (1981) obtains a significantly lower figure (21.1 percent), though he also finds that children received within $500 of the average inheritance in 50.4 percent of all cases. Menchik (1988) argues that the lower frequency of equal division in Tomes's sample reflects data problems.

[2] Bernheim et al. argue that the strategic exchange motive has fewer difficulties with the prevalence of equal division than other theories of bequests, but they do not provide a theoretical framework that yields equal division as a robust prediction (see also Unur 1998).

[3] The accidental bequest hypothesis is inconsistent with the observation that many individuals appear to resist annuitization. For additional evidence, see, e.g., Bernheim et al. (1985), Hurd (1987), Bernheim (1991), and Gale and Scholz (1994).

[4] Any arbitrary rule for dividing bequests, including equal division, is ex post optimal for an altruistic parent provided that the parent anticipates the application of the rule and fully compensates through gifts prior to death. Thus, in some settings, the optimal division of bequests is indeterminate. Lundholm and Ohlsson (2000) propose a model that resolves this indeterminacy in favor of equal division. They proceed (as we do) from the premise that gifts are observable whereas bequests are not. However, in contrast to our analysis, their model assumes—and therefore does not explain—the existence of an equal division norm. Moreover, their theory is inconsistent with the available evidence indicating that gifts are only partially compensatory. In practice, the frequencies and magnitudes of gifts do not seem sufficient to offset the effects of changing resources, preferences, and other conditions on the optimal division of bequests.

omists must carefully qualify all conclusions that are linked to assumptions about transfer motives.

In this paper, we propose a theory of intergenerational transfers that accounts for the equal division puzzle, including the unequal division of gifts. We consider a model in which an altruistic, utility-maximizing parent divides his or her estate between potentially heterogeneous children. To this relatively standard framework we add a new element: each child's perception of parental affection directly affects his or her subjective well-being. This assumption is grounded in psychological evidence (see, e.g., Coopersmith 1967; Bednar and Peterson 1995). In particular, children care about the extent to which they are loved or valued by a parent, relative to brothers and sisters (see, e.g., Tesser 1980; Brody, Stoneman, and McCoy 1994; Bank and Kahn 1997).

Our theory requires one additional plausible assumption: children cannot directly observe the parent's preferences and instead infer these preferences from the parent's actions, including bequests. The altruistic parent must then consider the possibility that an unequal bequest may cause the children to infer that they are loved either more or less than their siblings. In this setting, bequests serve as signals of parental affection. Under conditions identified in the text, no separating equilibrium exists, but there is an attractive equilibrium in which a positive fraction of the population adheres to a norm of equal division. For appropriately chosen parameter values, this fraction can be arbitrarily large.

The intuition for our central result is straightforward. If parents prefer to appear less partial than they actually are, then those who love their children unequally have incentives to imitate the behavior of those with relatively little bias. To differentiate himself or herself from any particular potential imitator, a relatively impartial parent must give the child who is less loved by the imitator a larger share than the child would receive if the parent's preferences were observable. Since each child is favored by some potential imitator, it therefore is impossible to divide the estate of a relatively impartial parent in a way that discourages all imitation. Consequently, equilibrium tends to produce a pool at the center of the parental type space.

Our analysis also accounts for the unequal division of gifts prior to death. The key difference between gifts and bequests relates to observability: the division of bequests is perfectly observable by all concerned parties, whereas the division of gifts need not be. As long as "secret" gifts are feasible, neither child is in a position to verify that the parent's resources have been divided equally, and an equal division norm cannot survive.

There are also conditions under which parents prefer to appear more partial than they actually are. In such cases, parents shade their choices toward favored children to differentiate themselves from those who love their children more equally. This gives rise to pools at the boundaries of the feasible choice set, with a single child receiving the parent's entire estate. Thus, under appropriate parametric assumptions, our model can also account for the pattern known as "unigeniture." This is of interest since unigeniture is a common norm in many societies (see, e.g., Chu 1991; Guinnane 1992).

II. The Model

A. *The Environment*

We consider interactions among three parties: a parent (P) and two children ($i = 1, 2$). The parent is endowed with wealth $w_P > 0$, which he or she divides between the children by making nonnegative bequests, $b_i \geq 0$. We focus on the division of bequests and abstract from the possibility that the parent might consume some portion of his or her resources (imagine that death is imminent and that rapid consumption is not attractive).[5] Thus the parent chooses bequests to satisfy the constraint $b_1 + b_2 = w_P$. Each child i is endowed with wealth $w_i > 0$ and consumes $c_i = w_i + b_i$. For simplicity, we assume that all parties can observe each others' endowments. We examine the role of this last assumption in Section III*C*1.

It is convenient to think of the parent as dividing the family's total resources, $W \equiv w_P + w_1 + w_2$. Specifically, the parent picks $x \in [\underline{x}, 1 - \bar{x}]$, where $\underline{x} \equiv w_1/W$ and $\bar{x} \equiv w_2/W$; child 1 consumes xW, and child 2 consumes $(1 - x)W$. With this change of variables, it is important to keep in mind that the phrase "equal division" refers to the parent's endowment rather than to the family's resources. That is, equal division occurs when $b_1 = b_2$ or, equivalently, when

$$x = x^E \equiv \frac{1 + (w_1/W) - (w_2/W)}{2}.$$

B. *Preferences*

We use U_P and U_i to denote the utilities of the parent and children, respectively. We assume that the parent is an altruist. Since the parent does not consume anything directly, his or her utility depends only on

[5] Our theory does not attempt to explain why so many individuals reach the end of life with positive bequeathable assets (indeed, it suggests that parents should prefer to make transfers as gifts rather than as bequests). This phenomenon is potentially attributable to factors outside the model, such as uncertainty concerning the length of life combined with imperfections in annuity markets. Individuals may also derive feelings of security, control, or satisfaction from asset ownership.

the outcomes for the children: $U_P = tU_1 + (1-t)U_2$, where $t \in [0, 1]$. Parents differ according to the relative weight t that they attach to the first child's utility. We assume that t is known to the parent but not to the children. Though children may have had many opportunities to learn about their parents' preferences, significant uncertainty remains. Children's prior beliefs about t are given by some atomless cumulative distribution function F, and the support of F is the interval $[0, 1]$. We use f to denote the density function associated with F, and we assume that f is symmetric around $\frac{1}{2}$ (i.e., $f(t) = f(1-t)$).[6]

We assume that each child cares about his or her own consumption, c_i, as well as about t. That is, each child's sense of well-being is affected by the extent to which he or she feels "loved" relative to a sibling. Though children cannot observe t directly, they may attempt to infer it from aspects of the parent's behavior, including the choice of x. When the children believe that $t = \hat{t}$, their utilities are given by $U_1 = u(c_1) + \beta v(\hat{t})$ and $U_2 = u(c_2) + \beta v(1-\hat{t})$, where u is defined over $[0, +\infty)$, and v is defined over $[0, 1]$. The parent's utility is therefore

$$\begin{aligned} U_P &= [tu(xW) + (1-t)u((1-x)W)] + \beta[tv(\hat{t}) + (1-t)v(1-\hat{t})] \\ &\equiv U(x, t) + \beta V(\hat{t}, t). \qquad (1) \end{aligned}$$

Assumption 1. The functions u and v are strictly increasing, strictly concave, and twice continuously differentiable on (respectively) $(0, W]$ and $[0, 1]$, $\lim_{c\to 0} cu'(c) = +\infty$, $v'(0)$ is finite, and $v'(1) = 0$.

Most of assumption 1 is reasonably standard. Weaker conditions would suffice for most of our results.[7] Since it is possible to live without parental affection but not without consumption, it is reasonable to assume that the derivative of v is finite at $\hat{t} = 0$, whereas the derivative of u is infinite at $c = 0$. It is also natural to assume that a child is satiated when he or she has all of the parent's affection. From assumption 1, it follows that $U(x, t)$ is twice continuously differentiable on $[\underline{x}, 1-\bar{x}] \times [0, 1]$, $V(\hat{t}, t)$ is twice continuously differentiable on $[0, 1]^2$, $U_{11}(\cdot) < 0$, $V_{11}(\cdot) < 0$, $U_{12}(\cdot) > 0$, and $V_{12}(\cdot) > 0$.

Thus far, we have confined our discussion of preferences to cases in which children are certain about the parent's type ("degenerate" beliefs). To analyze pooling equilibria, we also need to describe payoffs when the children have nondegenerate beliefs (i.e., they are not certain about the parent's type). We imagine that there is a mapping $\hat{B}$ from beliefs about types (probability distributions) into types, with the fol-

[6] The symmetry assumption is not essential, but it allows us to simplify some of the proofs. Conceptually, cases in which t is distributed asymmetrically are similar to cases in which the children have unequal endowments, which we treat explicitly.

[7] For example, since $c_i \geq w_i > 0$, the properties of u near zero are inconsequential.

lowing interpretation: if the children's beliefs about the parent's type are summarized by the probability distribution ϕ, their reaction is the same as though they knew with certainty that the parent's type was $\hat{B}(\phi)$. The parent's utility is then given by equation (1), where $\hat{B}(\phi)$ replaces the term $\hat{t}$.

In much of the subsequent analysis, we focus on cases for which children learn only that the parent's type lies between some lower bound r and some upper bound s. The corresponding posterior probability distributions have the form $\phi(t) = [F(t) - F(r)]/[F(s) - F(r)]$ for $t \in [r, s)$, with $\phi(t) = 1$ for $t \geq s$ and $\phi(t) = 0$ for $t < r$, where $0 \leq r < s \leq 1$. On this restricted domain, one can write $\hat{B}(\phi)$ as $B(r, s)$.

Naturally, it is difficult to proceed analytically unless we impose some restrictions on the mapping $\hat{B}$. One possibility is to assume that each child's utility depends on the subjective expectation of v (in which case $\hat{B}$ is a certainty equivalent). Since, however, uncertainty about the parent's preferences does not entail a lottery over outcomes (if t is never fully revealed, then the children's uncertainty is never completely resolved, and the utilities $v(t)$ and $v(1-t)$ are never actually realized), standard justifications for using subjective expectations may not apply. We therefore proceed by imposing a small number of minimal and relatively unobjectionable restrictions.

Assumption 2. (i) If, for some t', $\phi(t') = 1$ and $\phi(t) = 0$ for $t < t'$, then $\hat{B}(\phi) = t'$. (ii) For ϕ' and ϕ'' such that $\phi'(t) \leq \phi''(t)$ for all t with strict inequality for some t, we have $\hat{B}(\phi') > \hat{B}(\phi'')$. (iii) The function $B(r, s)$ is twice continuously differentiable, and $B(r, s) = 1 - B(1-s, 1-r)$.

Part i is essentially a tautology. Part ii requires that, if the children's beliefs shift toward higher types (in the sense of first-order stochastic dominance), they react as though the parent is a higher type. Part iii includes a technical differentiability condition along with the requirement that B is symmetric around $\frac{1}{2}$. From these assumptions, one can derive two additional properties: first, $B(r, s)$ is increasing in r and s; second, for $r < s$, we have $B(r, s) \in (r, s)$. Note in particular that the expectations operator satisfies assumption 2.

C. Parental Bliss Points

Ignoring for the moment the possibility that children may infer $\hat{t}$ from x, we optimize $U(x, t)$ over $x \in [0, 1]$ to find the parent's "action bliss point," $X(t)$. Because $u(\cdot)$ is strictly concave, $X(t)$ is the solution of the following first-order condition:[8]

$$tu'(X(t)W) = (1 - t)u'([1 - X(t)]W). \quad (2)$$

From assumption 1, it follows that $X(t)$ is well defined, single-valued, strictly increasing, and continuous with $X(0) = 0$, $X(1) = 1$, and $X(\frac{1}{2}) = \frac{1}{2}$. Moreover, $U(x, t)$ is single-peaked in x, with a maximum at $x = X(t)$.

EXAMPLE 1. The utility function $u(c) = c^\gamma/\gamma$, with $\gamma < 0$ (so that $\lim_{c\to 0} cu'(c) = +\infty$, as required in assumption 1). Then $X(t) = (1 - t)^\eta[t^\eta + (1 - t)^\eta]^{-1}$, where $\eta = 1/(\gamma - 1)$.

Similarly, we maximize $V(\hat{t}, t)$ over $\hat{t} \in [0, 1]$ to find the parent's "perception bliss point," $p(t)$. Since $v(\cdot)$ is strictly concave, $p(t)$ is the solution to the following first-order condition:[9]

$$tv'(p(t)) = (1 - t)v'(1 - p(t)). \quad (3)$$

From assumption 1, it follows that $p(t)$ is well defined, single-valued, strictly increasing, and continuous with $p(0) = 0$, $p(1) = 1$, and $p(\frac{1}{2}) = \frac{1}{2}$. Moreover, $V(\hat{t}, t)$ is single-peaked in $\hat{t}$, with a maximum at $\hat{t} = p(t)$.

EXAMPLE 2. The function $v(\hat{t}) = -(1 - \hat{t})^2$. Then $p(t) = t$.

In figure 1, we exhibit indifference contours for two types of parents, t and t'. Notice that these contours are ellipses and generally cross twice or not at all. Consequently, preferences do *not* satisfy the familiar Spence-Mirrlees single-crossing property.

D. *The Direction of Imitation*

The perception bliss point function, $p(t)$, plays a critical role in our analysis. Its relation to t is particularly important. We focus on two special cases.

CONDITION 1. *Imitation toward the center.*—The function $p(t) > t$ for $t \in (0, \frac{1}{2})$, $p(t) < t$ for $t \in (\frac{1}{2}, 1)$, and $p'(\frac{1}{2}) < 1$.

CONDITION 2. *Imitation toward the extremes.*—The function $p(t) < t$ for $t \in (0, \frac{1}{2})$, $p(t) > t$ for $t \in (\frac{1}{2}, 1)$, and $p'(\frac{1}{2}) > 1$.

Figure 2 depicts perception bliss point functions satisfying conditions 1 and 2. Under condition 1, all types wish to be perceived as loving their children more equally than they actually do, so imitation tends to occur toward the center of the type space. Under condition 2, all types wish to be perceived as loving their children less equally than they actually do, so imitation tends to occur toward the extremes of the type

[8] Since $\lim_{c\to 0} u'(c) = +\infty$, we know that the solution to the first-order condition is interior for all $t \in (0, 1)$. Plainly, $X(0) = 0$ and $X(1) = 1$.

[9] Since $v'(0)$ is finite and $v'(1) = 0$, we know that the solution to the first-order condition is interior for all $t \in (0, 1)$. Plainly, $p(0) = 0$ and $p(1) = 1$.

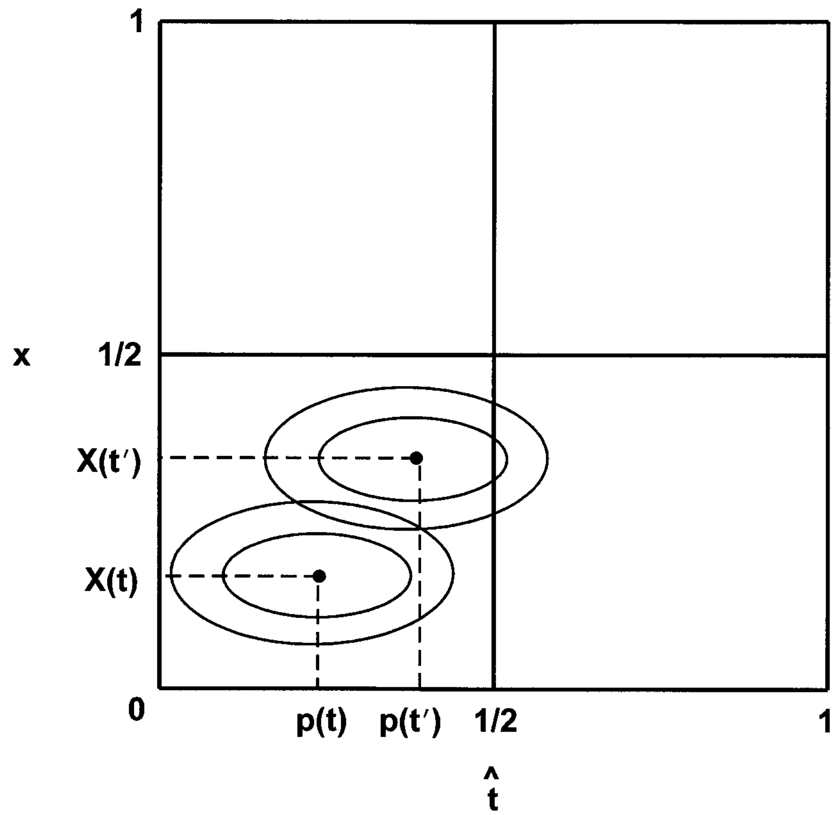

FIG. 1.—Illustration of indifference contours

space. Naturally, it is possible to have mixed configurations, but we do not examine them in the current paper.

EXAMPLE 3. Suppose that $v(\hat{t}) = h(-(1 - \hat{t})^2)$ and that v otherwise satisfies assumption 1. Then condition 1 is satisfied if h is concave, and condition 2 is satisfied if h is convex. In example 2, we considered the boundary case in which h is linear, so $p(t) = t$.

The curvature of v (and hence the shape of p) plays an important role in determining the qualitative properties of equilibria. Indeed, systematic variation in the shape of utility functions across different cultures emerges as a potential explanation for the observed differences in norms (equigeniture vs. unigeniture). One would expect the function v to exhibit greater curvature in societies in which the returns to factors associated with parental affection (e.g., assistance with securing a job) decrease at a more rapid rate. This characteristic may be related to the availability of good substitutes for parental support and attention. In

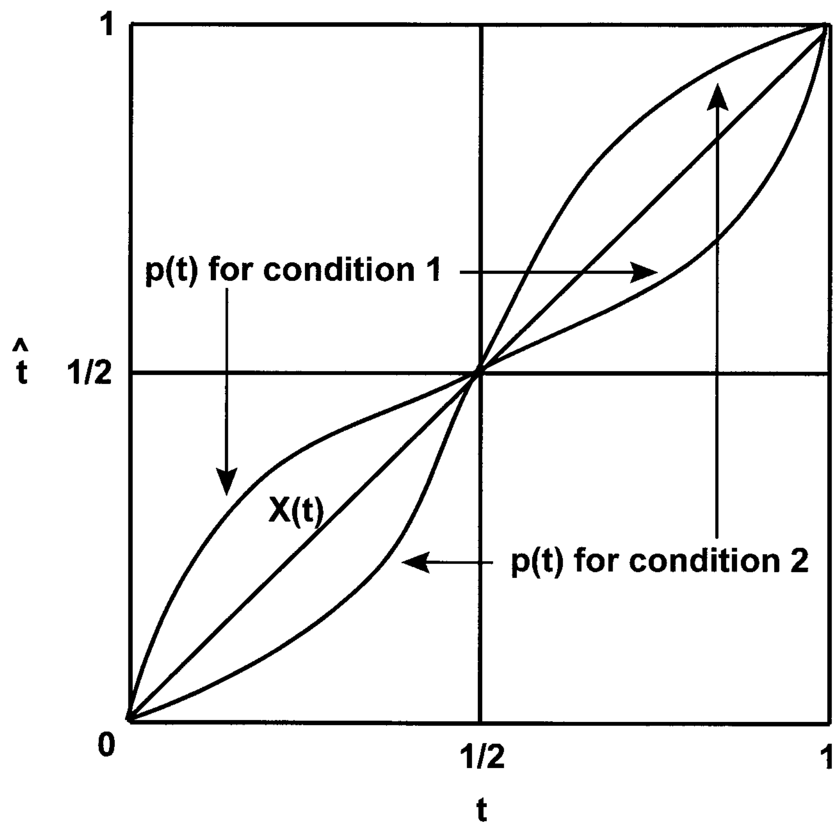

FIG. 2.—Illustration of conditions 1 and 2

open, highly mobile societies, individuals who have achieved a basic level of functionality can successfully strike out on their own, whereas in closed societies, such individuals may remain closely tied to their communities and therefore more dependent on parental goodwill. These observations suggest that condition 1 may be more likely to prevail in developed economies, whereas condition 2 may be more likely to prevail in less developed economies.

E. *Decisions, Inferences, and Equilibrium*

The structure of the game is simple. After observing t, the parent selects x and may also send a message, $m \in [0, 1]$. This message represents "pure" communication about the parent's type, in the sense that the value of m does not directly enter the utility function of any party (in the pertinent literature, this is usually referred to as "cheap talk"). We elaborate on the role of pure communication in the next section. Children observe x and m and draw inferences about the parent's preference parameter, t. The preceding expressions for U_1, U_2, and U_P describe the resulting payoffs.[10]

In this setting, the parent's choices of x and m can signal the parent's type, t. Formally, the model is a "signaling game" in the sense of Banks and Sobel (1987) or Cho and Kreps (1987): the parent is a "sender," the children are "receivers," (x, m) is the sender's "message," and $\hat{t}$ is both the receivers' inference and the receivers' "response." While it is somewhat unconventional to identify the receivers' inference with the receivers' response, this is easily reconciled with standard formulations of signaling.[11]

A signaling equilibrium involves a pair of choice functions, $\mu(t)$ and $\gamma(t)$, mapping the parent's type t to, respectively, decisions concerning the division of bequests, x, and a pure message, m, as well as an inference function $\phi(\hat{t}, x, m)$ mapping all feasible choices into probability distributions over perceived type, $\hat{t}$.[12] The choice function must prescribe optimal decisions for all types t given the inference function. The inference function must be consistent with the choice function, in the sense that it is derived from the choice function by applying Bayes' law for all choices occurring with positive likelihood in equilibrium.

III. Imitation toward the Center

In this section, we explain the prevalence of equal division by identifying conditions that give rise to equilibria in which a substantial fraction of the population chooses $x = x^E$. We account for the emergence of this norm by demonstrating that the corresponding pooling equilibria have attractive properties, and we attribute the stability of this norm to the robustness of the equilibria.

[10] Implicitly, we assume that the parent correctly anticipates and cares about the inferences that children will make after the parent's death and that the children attempt to make the best inferences possible. We do not explore the interesting possibility that children might have incentives to engage in self-deception, intentionally forming incorrect inferences.

[11] Instead of assuming that the parent cares directly about a child's inference, assume that the parent cares about the child's reaction to his or her inference. One can then renormalize the set of possible reactions to conform with the set of possible inferences. In other words, one can use $\hat{t}$ to denote a child's reaction to the inference that the value of the parent's altruism parameter is $\hat{t}$.

[12] Naturally, $\mu(t)$ and $\phi(\hat{t}, x, m)$ also depend on the endowments $(\underline{x}, \bar{x})$. We omit the dependence on these parameters for notational brevity.

A. Equilibria without Social Norms

Even the simplest signaling models can give rise to vast sets of equilibria (see, e.g., Spence 1974). Typically, many of these equilibria entail some pooling. If one associates pools with social norms, the potential existence of norms in the present environment may not seem particularly surprising at first. However, in many familiar signaling problems, equilibria with pooling have unattractive properties, and full separation is a more plausible outcome (see, e.g., Cho and Kreps 1987). Consequently, the emergence of a social norm—let alone a norm of equal division—is not a foregone conclusion. To make the case that signaling provides a plausible explanation for equal division, we must provide an explicit justification for studying equilibria with norms (pooling) rather than equilibria without norms (full separation). In this subsection, we provide a particularly compelling justification: for some ranges of the model's parameters, pooling is unavoidable because full separation is infeasible.

Our first result identifies some important properties of equilibrium separating action functions (when they exist).

THEOREM 1. Suppose that condition 1 is satisfied. Then, in any separating equilibrium with endowments given by some pair $(\underline{x}, \bar{x})$, the function $\mu(t)$ is strictly monotonically increasing and continuous, $\mu(t) > X(t)$ for $t \in (0, \frac{1}{2})$, $\mu(t) < X(t)$ for $t \in (\frac{1}{2}, 1)$, and $\mu(\frac{1}{2}) = \frac{1}{2}$.

Thus, in any separating equilibrium, a child's equilibrium share of the family's resources increases monotonically with the relative weight that the parent attaches to him or her.[13] Moreover, parents who favor one child or the other "lean" toward egalitarianism, whereas parents who place equal weight on both children divide resources equally.

Our next task is to characterize the circumstances under which it is possible to construct a separating equilibrium satisfying the properties listed in theorem 1. Note that any separating equilibrium for our model must also constitute an equilibrium when the type space is restricted to either $[0, \frac{1}{2})$ or $(\frac{1}{2}, 1]$. With either restriction, our model presents a relatively standard signaling problem in which either every type prefers to be mistaken for a higher type ($t \in [0, \frac{1}{2})$) or every type prefers to be mistaken for a lower type ($t \in (\frac{1}{2}, 1]$). Consequently, on an intuitive level, we can proceed in two steps. First, we solve for a separating equilibrium action function, $\underline{\mu}(t)$, when the set of types is restricted to $[0, \frac{1}{2})$, as well as for another separating equilibrium action function,

[13] In standard signaling models wherein preferences satisfy the single-crossing property, it is easy to prove that actions must be monotonic in type. In our model, the single-crossing property is not satisfied, so the standard argument does not apply. The proof of monotonicity is surprisingly subtle and depends heavily on the assumption that the type space is a continuum. With a finite number of types, it is sometimes possible to construct separating equilibria that violate monotonicity. Even with a continuum of types, it is sometimes possible to construct pooling equilibria that violate monotonicity.

$\bar{\mu}(t)$, when the set of types is restricted to $(\frac{1}{2}, 1]$. Second, we check to see whether it is possible to construct an equilibrium for the complete model by "pasting" these functions together.

Assume for the moment that the set of types is $[0, \frac{1}{2})$. The indifference contours of each type $t \in [0, \frac{1}{2})$ must be tangent to $\underline{\mu}(t)$ at the equilibrium outcome assigned to that type. Thus $\underline{\mu}(t)$ corresponds to the solution of the differential equation

$$\underline{\mu}'(t) = -\frac{\beta V_1(t, t)}{U_1(\underline{\mu}(t), t)} \tag{4}$$

on the interval $t \in [0, \frac{1}{2})$. The existence and uniqueness of $\underline{\mu}(t)$ for a given initial condition (i.e., a value for $\underline{\mu}(0)$) follow from modifications of standard arguments. If the set of types is $(\frac{1}{2}, 1]$, the characterization of $\bar{\mu}(t)$ is analogous.

Since parents lean toward egalitarianism (theorem 1), we know that $\lim_{t\uparrow 1/2}\underline{\mu}(t) \geq \frac{1}{2}$ and that $\lim_{t\downarrow 1/2}\bar{\mu}(t) \leq \frac{1}{2}$. Since the action function for a separating equilibrium is necessarily monotonic (also theorem 1), it is impossible to construct a separating equilibrium for the complete model by "pasting" $\underline{\mu}(t)$ and $\bar{\mu}(t)$ together if either of the preceding inequalities is strict. Consequently, full separation of types can occur only if

$$\lim_{t\uparrow 1/2}\underline{\mu}(t) = \lim_{t\downarrow 1/2}\bar{\mu}(t) = \tfrac{1}{2}. \tag{5}$$

Equation (5) may or may not hold in any given instance. If it fails to hold with the most extreme feasible initial conditions ($\underline{\mu}(0) = \underline{x}$ and $\bar{\mu}(0) = 1 - \underline{x}$), then it fails to hold for all other feasible initial conditions. As one increases the utility attached to perceptions (β) or children's endowments ($\underline{x}$ and $\bar{x}$) or both, any given type $t \in (0, \frac{1}{2})$ must choose a larger x, and any given type $t \in (\frac{1}{2}, 1)$ must choose a lower x, to discourage imitation. As a result, full separation (the absence of a norm) is impossible when children have sufficient resources or attach sufficient importance to parental affection. The following theorem states this formally.

Theorem 2. Suppose that condition 1 is satisfied. For any pair of endowments $(\underline{x}, \bar{x})$, there exists $\beta^*(\underline{x}, \bar{x})$ such that a fully separating equilibrium exists if and only if $\beta \leq \beta^*(\underline{x}, \bar{x})$. Moreover, $\beta^*(\underline{x}, \bar{x})$ is decreasing in $\max\{\underline{x}, \bar{x}\}$ and strictly positive iff $\max\{\underline{x}, \bar{x}\} < \frac{1}{2}$.

It follows that any equilibrium must involve some pooling when $\beta > \beta^*(\underline{x}, \bar{x})$. In the next subsection, we investigate the structure of pooling equilibria.

B. Equilibria with Social Norms

1. Central Pooling Equilibria

As is usually the case with signaling models, our model gives rise to a wide variety of pooling equilibria. We focus attention on equilibria characterized by "central pooling." In Section *B* of the Appendix, we justify this focus by discussing and applying formal criteria for selecting among equilibria.

In a central pooling equilibrium, full separation occurs for individuals with sufficiently extreme preferences, but all those with intermediate preferences select the same action (i.e., they conform to a social norm). Despite conforming with respect to actions, intermediate types may nevertheless differentiate themselves to a limited extent through credible verbal statements (m). Thus, for example, when a norm of equigeniture prevails, parents who divide their bequests equally may nevertheless make informative statements about the extent to which they favor one child or the other.

Formally, a central pooling action function is characterized by three variables: x_p (the action chosen by all intermediate types), $t_l < \frac{1}{2}$ (the lowest intermediate type), and $t_h > \frac{1}{2}$ (the highest intermediate type). We refer to x_p as the *social norm* of the central pooling equilibrium and to $[t_l, t_h]$ as the *action pool*. The action function is constructed as follows:

$$\mu(t) = \begin{cases} x_p & \text{for } t \in [t_l, t_h] \\ \underline{\mu}(t) & \text{for } t < t_l \\ \bar{\mu}(t) & \text{for } t > t_h. \end{cases}$$

Feasibility requires $x_p \in [\underline{x}, \bar{x}]$. We restrict attention to cases in which $\underline{\mu}(t_l) \leq x_p \leq \bar{\mu}(t_h)$ to assure that the action function is monotonic. Note that a separating equilibrium is a special (degenerate) case of a central pooling equilibrium wherein $t_l = \underline{\mu}(t_l) = x_p = \bar{\mu}(t_h) = t_h = \frac{1}{2}$.

Types within the action pool may also differentiate themselves through cheap-talk messages. For any m with $\gamma(t) = m$ for some $t \in [t_l, t_h]$, we refer to $\{t \in [t_l, t_h] | \gamma(t) = m\}$ as a *segment* of the action pool (it is the set of types that conform to the norm x_p while conveying the message m). Under assumption 1, these segments must consist of N consecutive intervals, $[t_1, t_2]$, $[t_2, t_3]$, …, $[t_N, t_{N+1}]$, where $t_l = t_1 < t_2 < \cdots < t_{N+1} = t_h$.

Equilibrium obtains when a collection of indifference conditions are satisfied. First, type $t_l = t_1$ must be indifferent between separating and joining the lowest segment of the central pool:

$$U(\underline{\mu}(t_1), t_1) + V(t_1, t_1) = U(x_p, t_1) + V(B(t_1, t_2), t_1). \tag{6}$$

Second, each type on a boundary between consecutive segments of the central pool must be indifferent between the inferences associated with those segments:[14]

$$V(B(t_{n-1},\ t_n),\ t_n) = V(B(t_n,\ t_{n+1}),\ t_n) \quad \text{for } n = 2,\ \ldots,\ N. \tag{7}$$

Third, type $t_h = t_{N+1}$ must be indifferent between separating and joining the highest segment of the central pool:

$$U(x_p,\ t_{N+1}) + V(B(t_N,\ t_{N+1}),\ t_{N+1}) =$$
$$U(\bar{\mu}(t_{N+1}),\ t_{N+1}) + V(t_{N+1},\ t_{N+1}). \tag{8}$$

In some instances, it is of course possible to have equilibria with $N = 1$, in which case cheap talk is unnecessary. However, as a general matter, cheap talk is important because one cannot guarantee the existence of a central pooling equilibrium without it. In particular, a problem arises when the parent's perception bliss point, $p(t)$, is sufficiently close to his or her own type, t. In that case, it may be impossible to find values of $t_1 < \frac{1}{2}$ (with $\underline{\mu}(t_1) < x_p$) and $t_2 > \frac{1}{2}$ that satisfy condition (6).[15]

Our next result establishes that, with cheap talk, a (nondegenerate) central pooling equilibrium necessarily exists whenever $\beta > \beta^*(\underline{x},\ \bar{x})$. Notably, this is precisely when separating equilibria fail to exist.

Theorem 3. Suppose that condition 1 is satisfied and that $\beta > \beta^*(\underline{x},\ \bar{x})$. Then there exists a central pooling equilibrium with some norm x_p in which parent types choosing x_p separate themselves into a finite number (possibly just one) of subgroups through pure communication (m) and each of these subgroups is a connected interval.

2. Central Pooling Equilibria with a Norm of Equal Division

So far, nothing guarantees that it is possible to sustain equal division as a social norm ($x_p = x^E$). When the children have identical endowments, the model is symmetric, so one would naturally expect to obtain a central pool with equal division. This, however, would not account for the robustness of the norm. Exact equality of endowments is a "measure zero" event, and yet in practice many parents are heavily predisposed to divide bequests equally even when children's resources are unequal.

[14] This portion of our analysis is reminiscent of Crawford and Sobel (1982). It differs from Crawford and Sobel's in two respects. First, in our model, it is not the case that $p(t) > t$ for all t within the group choosing x_p. Second, the equilibrium in pure communication must be consistent with the equilibrium conditions for the overall game. Section A of the Appendix includes further comments on related signaling models.

[15] Since $X(t) \leq \underline{\mu}(t_1) \leq x_p$, we know that $U(\underline{\mu}(t_1),\ t_1) \geq U(x_p,\ t_1)$. Thus, if $V(t,\ t) > V(B(t,\ \frac{1}{2}),\ t)$ for all $t < \frac{1}{2}$, then the left-hand side of (7) exceeds the right-hand side for all values of $t_1 < \frac{1}{2}$ and $t_2 > \frac{1}{2}$. Intuitively, segmentation of the central pool through cheap talk may be necessary because, otherwise, the central pool may be too large (in the sense that the inference associated with the central pool, $B(t_l,\ t_h)$, is less attractive to the boundary types t_l and t_h than full disclosure).

Our model provides a plausible explanation for the equal division puzzle because a norm of equigeniture can prevail even with unequal endowments, provided that the degree of inequality is not too great. To understand why, note that, if we think of x_p as a parameter rather than as an endogenously determined variable, expressions (6)–(8) form a system of $N+1$ equations in $N+1$ unknowns $(t_1, \dots, t_{N+1})$. Suppose for the moment that the children's endowments are equal ($\underline{x} = \bar{x} = x^0$), and consider the equilibrium with a norm of equal division ($x_p = \frac{1}{2}$). By the implicit function theorem, as long as the system of equations is locally invertible, there also exists a solution for every $(x_p, \underline{x}, \bar{x})$ in some neighborhood of $(\frac{1}{2}, x^0, x^0)$. This neighborhood includes points of the form $(x^E(\underline{x}, \bar{x}), \underline{x}, \bar{x})$ for all $(\underline{x}, \bar{x})$ with $\underline{x} \neq \bar{x}$ sufficiently close to (x^0, x^0). Each of these solutions corresponds to a central pooling equilibrium with a norm of equal division for an environment with unequal endowments. One can also show that, generically (i.e., "almost always"), the system is locally invertible. Thus equigeniture is a robust social norm.

Formally, let Λ denote the set of functions v satisfying assumption 1 for which condition 1 holds. Using a standard mathematical notion of genericity, we obtain the following result.[16]

Theorem 4. Suppose that condition 1 is satisfied. Consider any $x^0 \in (0, 1)$. For almost all $v \in \Lambda$, a central pooling equilibrium with equal division exists for all endowments within some neighborhood of $(\underline{x}, \bar{x}) = (x^0, x^0)$.

The nature of the preceding argument makes it plain that x_p is locally indeterminate. Why select x^E? When many outcomes are consistent with equilibria, "meeting in the middle" is often the most natural rule for coordinating activity (Schelling 1960). Moreover, any equilibrium not involving some form of equal division would require the parties to share a common understanding of the "ordering" of the children (i.e., which child has received x and which has received $1 - x$). If, for example, the norm is a 60–40 split, there may be confusion as to whether a 40–60 split constitutes a deviation from the norm. Families could base the ordering on some objective criterion such as age, but there are many competing criteria (gender, income, and so forth). With a norm of unequal division, the ordering always favors one child over the other, so the children might even take the ordering itself as a signal of t.

Equal division of the family's resources (as opposed to equal division of bequests) is also appealing as a focal norm. However, unlike equal division of bequests, it is not always feasible (one child may have more

[16] We say that a property holds for "almost all" v if it holds on an open-dense subset of Λ, where Λ is endowed with the topology of uniform C^1 convergence (i.e., two functions v and $\tilde{v}$ are close if their values and first derivatives are close everywhere on [0, 1]).

than half of the family's resources), and it requires parents to constantly adjust their wills through time when children accumulate resources at different rates. We also doubt that it would be robust if one added some degree of asymmetric information about w_1 and w_2; unlike equal division of bequests, equal division of total family resources is not easily verified.

3. Computations

Our next objective is to determine whether the model can generate outcomes in which equal division of estates is the predominant mode of behavior, even among families for which children's endowments are substantially unequal. We investigate this issue computationally using the following parameterization: $u(c) = c^\gamma/\gamma$ with $\gamma < 1$ (as in example 1) and $v(\hat{t}) = -(1-\hat{t})^\lambda$ with $\lambda > 2$ (a special case of example 3 wherein condition 1 is satisfied).[17] Under these assumptions, we can rewrite equation (1), which defines the parent's utility, as

$$\frac{U_P}{W^\gamma} = \left[t\frac{x^\gamma}{\gamma} + (1-t)\frac{(1-x)^\gamma}{\gamma}\right] - \left(\frac{\beta}{W^\gamma}\right)[t(1-\hat{t})^\lambda + (1-t)\hat{t}^\lambda].$$

We have divided through by the constant W^γ to highlight the fact that behavior depends on W and β only through the ratio β/W^γ. For all the calculations presented here, we assume that $\gamma = 0.5$ and $\lambda = 3$.

Figure 3 shows that an equilibrium norm of equal division is consistent with greater inequality between children's endowments when children care more about parental affection. This is intuitive. A necessary condition for the existence of a central pooling equilibrium with a norm of equal division is that x^E lies between $\bar{\mu}(\frac{1}{2})$ and $\underline{\mu}(\frac{1}{2})$ (otherwise actions would not be monotonic in type). As we increase β, it becomes more difficult to discourage imitation by more extreme types, so decisions become more distorted from $p(t)$, until one reaches a point at which this necessary condition is satisfied.

For the purpose of the figure, we assume that children collectively own two-thirds of the family's resources ($x_k \equiv \underline{x} + \bar{x} = 0.66$), and we consider all possible values of $\beta/W^\gamma \in [0, 25]$ and $\underline{x} \in [0, 0.66]$ (equal division of children's endowments corresponds to $\underline{x} = 0.33$). The lightly shaded area identifies parameter values for which the aforementioned necessary condition is satisfied. The dark area identifies parameter val-

[17] To compute equilibria, we numerically approximate the solutions to the differential equations that define the separating functions $\underline{\mu}$ and $\bar{\mu}$. For any candidate value for t_l, we compute the implied segments of the central pool using the indifference conditions (6) and (7). This generates a candidate for t_h. If the indifference condition (8) is satisfied, the configuration is an equilibrium. To find all equilibria, we search exhaustively over all possible values of t_l in $[0, \frac{1}{2})$.

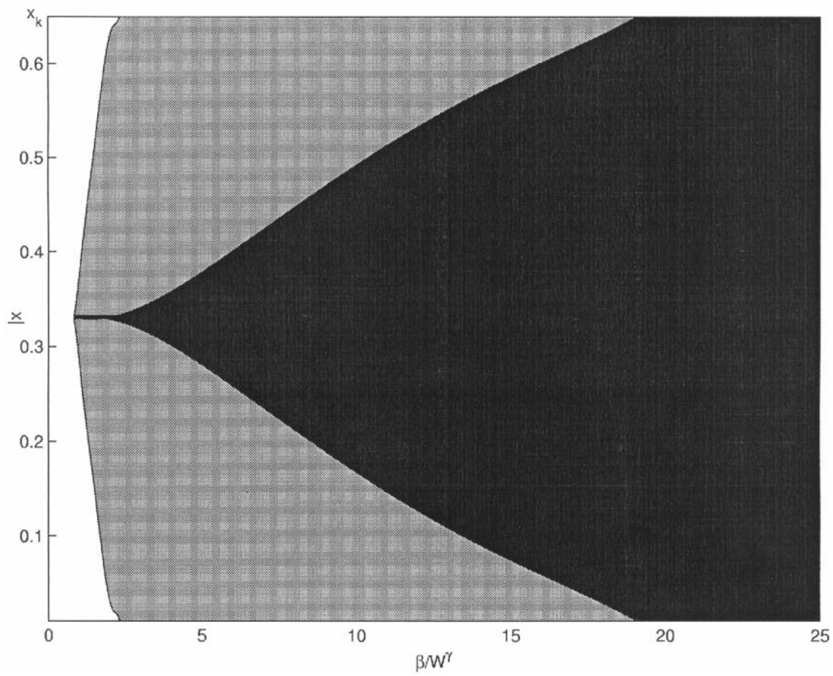

FIG. 3.—Existence of equilibria with an equal division norm: the effects of β and W

ues for which a central pooling equilibrium with equal division actually exists. Note that the shaded areas widen as β/W^{γ} increases. For $\beta/W^{\gamma} > 20$, equigeniture emerges in equilibrium essentially irrespective of how children's endowments are divided.[18] The figure also implies that an increase in the family's resources (with proportional endowments held fixed) can make equal division more or less feasible, depending on whether γ is, respectively, negative or positive.

Figure 4 shows that an equilibrium norm of equal division can be consistent with either greater or lesser inequality between children's endowments when the parent is wealthier relative to the children. The figure consists of three panels, corresponding to different values of β/W^{γ}. We consider all possible values of $w_1/w_k \in [0, 1]$ (where $w_k \equiv w_1 + w_2$, so $w_1/w_k = 0.5$ signifies equality of the children's endowments) and $w_P/W \in [0, 1]$ (w_P/W denotes the parent's endowed share of the family's resources). The shaded areas (light and dark) are defined as before.

For small values of w_P/W, an equilibrium norm of equal division is

[18] We say "essentially" any division of children's endowments because our computational approach encounters boundary problems as $\underline{x}$ approaches zero or x_k.

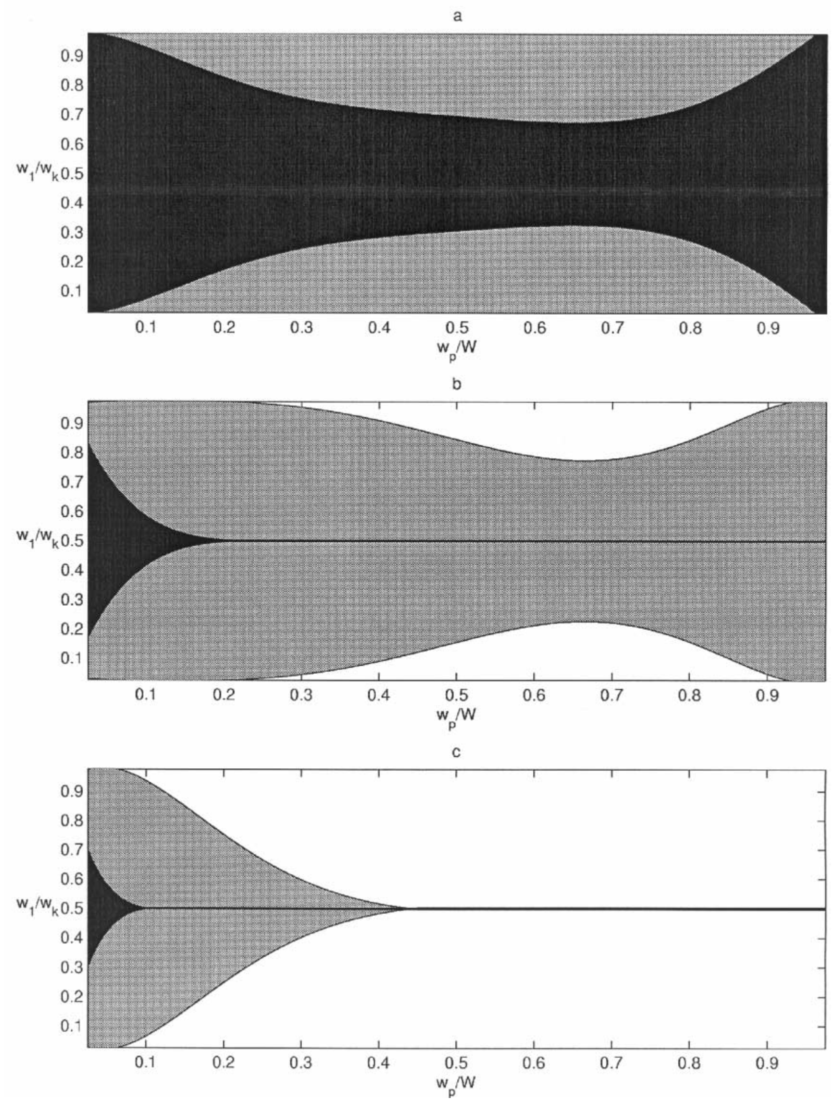

FIG. 4.—Existence of equilibria with an equal division norm: the effects of w_p/W. *a*, $\beta/W^{\gamma} = 10$; *b*, $\beta/W^{\gamma} = 2$; *c*, $\beta/W^{\gamma} = 1$.

always consistent with substantial inequality of children's endowments.[19] This is intuitive. When the parent has little wealth, the necessary condition for an equilibrium with equal division is always satisfied,[20] and the costs of giving offense to a less loved child are potentially large relative to the benefits of giving the favored child a larger share. As the parent's share of the family's wealth rises, the parent has more to gain from dividing bequests unequally (particularly if the degree of inequality between the children's endowments is great); consequently, the existence of an equal division norm requires greater equality of children's endowments. Note, however, that this relationship reverses in figure 4*a* when the parent is sufficiently wealthy relative to the children. This occurs for two reasons: (i) for high values of β/W^{γ}, an equilibrium with an equal division norm always exists when $w_1/w_k = \frac{1}{2}$, and (ii) when w_P/W is close to unity, the children are endowed with so few resources that there is no substantive difference between $w_1/w_k = \frac{1}{2}$ and any other value of w_1/w_k.[21] For intermediate values of w_P/W, x^E varies substantially with w_1/w_k; consequently, x^E may be inconsistent with the equilibrium conditions for central pooling when the value of w_1/w_k is sufficiently extreme.

This analysis generates at least one robust and potentially testable implication: when the parent's share of the family's resources is small, the likelihood of equal division rises as this share shrinks. However, since the relationship between the frequency of equal division and w_P/W may be nonmonotonic and since family resources are often difficult to measure, one must interpret pertinent empirical evidence with caution.

C. *The Role of Assumptions Concerning Observability*

1. The Observability of Endowments and Preferences

So far, we have assumed that parents and children are asymmetrically informed only about the parent's preferences (the parameter t). It is perhaps equally plausible to assume that the parent is unable to observe some aspect of a child's preferences. In practice, parents and children also have limited abilities to observe each other's resources. The intro-

[19] We suspect (but have not proved) that, for sufficiently small values of w_P/W, any value w_1/w_k is consistent with the existence of an equal division equilibrium. This is not apparent in figs. 4*b* and *c* because we do not use a sufficiently fine grid for w_P/W.

[20] This follows from the fact that $\underline{\mu}(\frac{1}{2})$ must exceed x^E (x^E must exceed $\bar{\mu}(\frac{1}{2})$) if $\underline{\mu}(0)$ is close enough to x^E ($\bar{\mu}(1)$ is close enough to x^E).

[21] For this reason, we suspect (but have not established) that, for sufficiently large values of w_P/W, any value w_1/w_k is consistent with the existence of an equal division equilibrium in fig. 4*b* as well. Figure 4*b* may obscure this property because we have not used a sufficiently fine grid for w_P/W.

duction of asymmetric information with respect to these additional variables does not, however, significantly alter our analysis.

To illustrate, suppose that the parent is uncertain about the value of the preference parameter β. Also imagine that each party has private information concerning his or her own endowment (but not concerning the endowment of any other party) and that it is impossible for the parent to infer the children's endowments from observed expenditures. Since bequests are observable, children can always compute the parent's terminal resources prior to inferring the parent's preferences. Consequently, we can continue to treat the parent's endowment as though it is publicly observable.

Given the parent's uncertainty about β, w_1, and w_2, we replace equation (1), our expression for the parent's utility, with the following (where E_P is the expectations operator corresponding to the parent's beliefs):

$$U_P = [tE_P u(w_1 + xw_p) + (1 - t)E_P u(w_2 + (1 - x)w_p)]$$
$$+ E_P(\beta)[tv(\hat{t}) + (1 - t)v(1 - \hat{t})].$$

Though one can no longer think in terms of the parent dividing total family resources W, few other changes in our analysis are required, and our central results are unaffected. The reason is that the informativeness of the parent's choices depends on the parent's beliefs about the parameters β, w_1, and w_2 rather than on their actual values.

2. The Observability of Cheap Talk

We have assumed that the cheap-talk signal m is observed by both children. With minor modifications, the model can accommodate the possibility that the parent can also speak privately to either child. Formally, the parent selects a triplet of pure messages, (m, m_1, m_2), where both children receive m, and child i privately observes m_i. It is easy to demonstrate that this changes nothing of substance since the private signals m_i are necessarily uninformative. Assume on the contrary that, for a given (x, m), m_i' leads to a different inference than m_i''. Changing m_i affects only the utility of child i. Moreover, every parent type t has the same preference ranking over child i's inference. Thus, if m_i' leads to a more favorable inference than m_i'' for one parent type, it does so for all parent types. Consequently, there can be no separation of types along the m_i dimension.

3. Gifts versus Bequests: The Observability of Transfers

Next, imagine that, contrary to our assumptions, a transfer to one child is not observable by the other child. If everything else is observable (aside from t), each child can compute the magnitude of the transfer to his or her sibling from endowments and expenditures, so our results are unchanged. However, in more realistic cases in which endowments or expenditures are also imperfectly observable, the preceding analysis is inapplicable.

In our model, a norm of equal division emerges because transfers serve as a signal of a parent's preferences that is common to both children. In other words, there is a single signal and two audiences. When transfers are neither directly observable nor perfectly inferable from other public information, they cannot serve as a common signal. Instead, each transfer provides a private signal to each child. If neither child is in a position to verify that the parent's resources have been divided equally, then the equilibrium inference function cannot systematically link the children's beliefs about the parent's preferences to the equality of transfers. Without such a link, an equal division norm cannot survive.[22]

As we mentioned in Section I, the available evidence suggests that a norm of equal division applies to bequests, but not to gifts. Our analysis suggests that the key difference between gifts and bequests relates to observability: the division of bequests is perfectly observable by all concerned parties, whereas the division of gifts need not be. Since a parent can give gifts to a favored child without revealing this to another child, gifts cannot serve as a common signal of the parent's preferences, so the analysis of the previous section does not apply.

IV. Imitation toward the Extremes

Having completed our analysis of condition 1, we briefly consider the implications of condition 2. When children have complete information concerning parents' preferences, any parent with $t \leq X^{-1}(\underline{x})$ or $t \geq X^{-1}(\bar{x})$ bequeaths everything to the most preferred child. This practice, known as unigeniture, is of interest because it serves as a behavioral norm in a number of societies outside of the United States. Existing theories of unigeniture (e.g., Chu 1991; Guinnane 1992) provide reasons to believe, in effect, that the population distribution of parental action bliss points, $X(t)$, is skewed toward the extremes. The mechanism outlined in this section should be viewed as a complement to these

[22] Imagine a candidate equilibrium in which a positive fraction of parent types divide their resources equally between their children conditional on each realization of the parent's resources. Provided that the parent transfers to each child i some amount b_i (possibly different for each child) such that $2b_i$ lies in the support of the probability distribution for the parent's endowment, the child must assume that he or she is observing an equilibrium choice rather than a deviation. Consequently, the parent has the ability to deviate from an equilibrium by giving more to one child and less to the other without encountering undesirable inferences associated with out-of-equilibrium beliefs.

theories in the following sense: for any distribution of $X(t)$, incomplete information about parents' preferences enlarges the set of individuals practicing unigeniture.

Formally, the problem is similar to one that we have already studied. Refer once again to figure 2. Note that the parent's perception bliss point functions, $p(t)$, have the same shape on the half interval $[0, \frac{1}{2}]$ for condition 1 and on the half interval $[\frac{1}{2}, 1]$ for condition 2. One can therefore analyze behavior for the half interval $[\frac{1}{2}, 1]$ under condition 2 analogously to our treatment of the half interval $[0, \frac{1}{2}]$ under condition 1.[23] In this instance, each parent discourages imitation by those who care about their children more equally by leaning toward greater inequality. This enlarges the set of parents selecting $\bar{x}$ relative to the case of complete information.[24] An analogous argument implies that incomplete information also enlarges the set of parents selecting $\underline{x}$.

Often (but not always), unigeniture takes the form of primogeniture, which means that the oldest child typically receives the parent's estate. Our model cannot explain a preference for older children. However, any other consideration that favors transfers to the oldest child would skew the distribution of parental action bliss points to one side of the parent type space (the side that represents favoritism toward the oldest child). The mechanism considered here would then enlarge the set of individuals practicing primogeniture.

V. Conclusion

In this paper, we have studied environments in which parental choices concerning the division of bequests provide children with information about the parent's preferences and in which children are directly affected by their perceptions of parental affection. Under conditions identified in the text, the model gives rise to equilibria that support norms of equal division, and these equilibria have attractive properties that

[23] Despite the obvious similarities, there are some important technical differences between the problems considered here and in Sec. III. Equation (4) still defines a dynamic system governing the evolution of the separating action function from any initial condition. However, since $t = \frac{1}{2}$ is now, in effect, the lowest type for both halves of the type space, $\mu(\frac{1}{2}) = \frac{1}{2}$ (rather than $\mu(0) = \underline{x}$ or $\mu(1) = \bar{x}$) is the natural initial condition. This means that, in a separating equilibrium, the lowest type receives both its action bliss point and its perception bliss point. It follows that the initial condition is a stationary point of the dynamic system described by (4). To obtain separating functions, one must therefore examine the stability properties of the dynamic system around the stationary point $(\frac{1}{2}, \frac{1}{2})$. It turns out that the system is saddle-point stable and that the unstable arm corresponds to a separating equilibrium with the desired properties. We omit the proof to conserve space (see Bernheim and Severinov 2000).

[24] Thus we obtain an extremal pool as in Cho and Sobel (1990) rather than a central pool. The analysis of pooling equilibria is actually much simpler than in the preceding section: one must make sure that nonimitation constraints are satisfied on both sides of a central pool but on only one side of an extremal pool.

argue in favor of their selection. Since these results depend critically on the assumption that transfers are necessarily observable by all children, the theory applies to bequests, but not to gifts given prior to death. Consequently, our model not only provides an explanation for the equal division of bequests but also reconciles this pattern with the unequal division of gifts. Under an alternative set of conditions, the model gives rise to equilibria that support norms of unigeniture. This is of interest because unigeniture is a common pattern outside of the United States.

Appendix

A. *Relationships with Other Signaling Models*

Standard signaling models (e.g., Spence 1974) assume that all senders wish to be perceived as the "highest" type, and imitation occurs in only one direction. In our framework, this corresponds to the assumption that $p(t) = 1$ for all $t \in [0, 1]$.

In Crawford and Sobel (1982), $p(t)$ is strictly increasing (as in our model), but (as in the standard setting) $p(t) > t$ for all t. Consequently, imitation occurs in only one direction. Signaling is permitted through pure communication (cheap talk), but senders cannot take costly (and therefore potentially discriminatory) actions. Crawford and Sobel demonstrate that a limited degree of separation through cheap talk is usually possible: in some equilibria, senders segment themselves into a finite number of groups.[25]

In Banks (1990) and Bernheim (1994), $p(t) = \frac{1}{2}$ for all $t \in [0, 1]$. Imitation occurs toward the center of the type space (and is therefore multidirectional, as in the current setting). However, $p(t)$ does not vary with t (as in the standard setting), and cheap talk is not permitted. Both Banks and Bernheim exhibit signaling equilibria with central pools.

Our model combines features of the settings studied by Crawford-Sobel and Banks-Bernheim. Because the perception bliss point is increasing in type, there is a role for cheap talk. However, since imitation is multidirectional (the perception bliss point function crosses the 45-degree line), we also obtain central pooling with respect to the costly action.

B. *Equilibrium Refinements and Central Pooling*

Intuitively, central pooling equilibria are plausible because they satisfy the following two properties: (1) the action function is monotonic,[26] and (2) the set of types choosing any pooling action x includes $t = \frac{1}{2}$. Since preferences are monotonic (those with higher values of t prefer higher actions x), the first property strikes us as natural. The desirability of the second property requires further explanation.

[25] Austen-Smith and Banks (2000) have extended the model of Crawford and Sobel by allowing costly signaling (burning money). They assume that the marginal cost of action is type-independent. In their model, a set of types take costly actions, whereas other types segment into a number of pools through cheap talk.

[26] Even though separating equilibria are necessarily monotonic (theorem 1), there are usually nonmonotonic pooling equilibria.

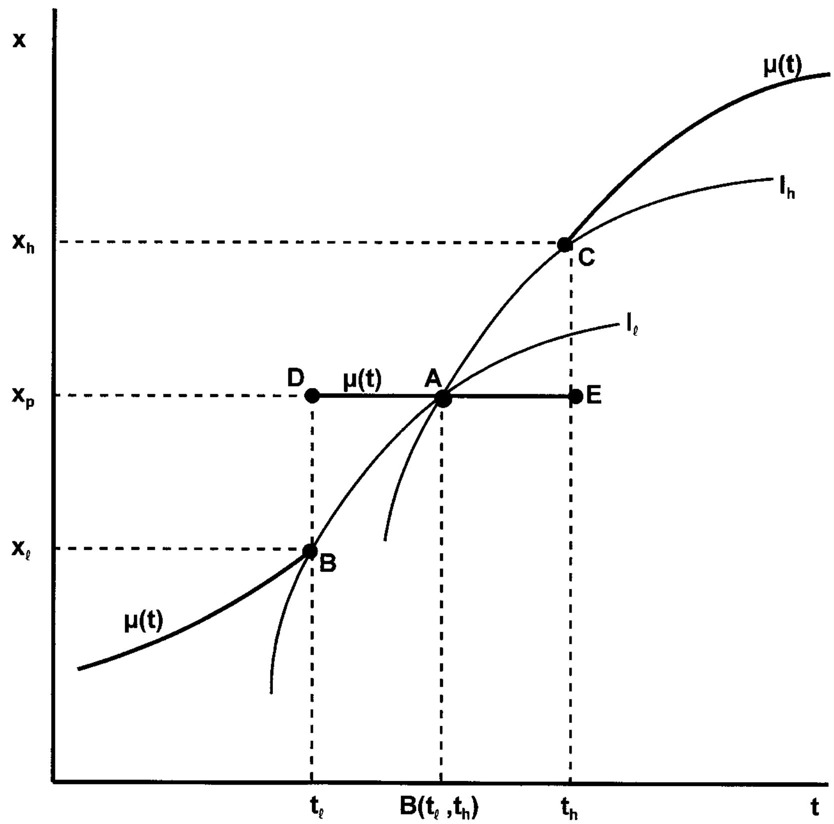

FIG. A1.—Pooling and equilibrium refinements

Standard equilibrium refinements for signaling games usually eliminate most types of pooling equilibria. To understand the nature of the argument, consider figure A1. The three dark line segments represent the equilibrium action function, $\mu(t)$. There is a single pool consisting of all types between t_l and t_h. Since equilibrium prevails, type t_l is indifferent between point *A,* the outcome if it joins the pool, and point *B,* the outcome if it separates; I_l represents a type t_l indifference curve passing through these points. Likewise, type t_h is indifferent between points *A* and *C;* I_h represents a type *h* indifference curve passing through these points. To obtain the equilibrium inference function, we invert the action function where possible. Consequently, inferences are given by $\mu^{-1}(x)$ for $x \leq x_l$ and $x \geq x_h$ (the lowest and highest segments of μ in the graph) and by $B(t_l, t_h)$ for $x = x_p$ (point *A*). The action function does not tie down inferences for actions between x_l and x_h other than x_p since these actions are not taken by any type in equilibrium.

Suppose that receivers attribute all actions just below the pooling action (those in (x_l, x_p)) to t_l, the lowest type in the pool, and all actions just above the pooling action (those in (x_p, x_h)) to t_h, the highest type in the pool. Graphically, this amounts to supplementing the inference function with the open intervals $\overline{BD}$

and $\overline{CE}$. With these inferences, type t_l would not deviate to any x in (x_l, x_p), but t_h would deviate to some x in (x_p, x_h). Thus the pool unravels from the top. Various equilibrium refinements eliminate pooling equilibria by establishing criteria that isolate the aforementioned inferences (see Banks and Sobel 1987; Cho and Kreps 1987; Cho and Sobel 1990). For example, the D1 criterion of Cho and Kreps attributes choices in (x_l, x_p) to t_l (and choices in (x_p, x_h) to t_h) because no other type is willing to take the action in question for as wide a range of possible inferences.

In standard settings, a pooling equilibrium can survive the application of such a refinement only if the action space is bounded and if the pool is located at the highest action. Since the highest type in the pool cannot deviate to a higher action, the pool cannot unravel from the top (see Cho and Sobel 1990). In the context of our model, similar reasoning suggests that one should confine attention to pools that include $t = \frac{1}{2}$. Recall that the direction of imitation is always toward $t = \frac{1}{2}$. Thus, for a pooling interval $[t_l, t_h]$ with $t_h < \frac{1}{2}$, t_h is the top, whereas for a pooling interval $[t_l, t_h]$ with $t_l > \frac{1}{2}$, t_l is the top. In either case, one might expect the pool to unravel from the top in the usual fashion. In contrast, a pooling interval $[t_l, t_h]$ with $t_l < \frac{1}{2} < t_h$ has, in effect, two bottoms and no top. This renders the associated equilibrium highly robust with respect to reasonable assumptions about inferences for actions not chosen in equilibrium.

We formalize this analysis as follows. Consider an equilibrium $(\mu(\cdot), \gamma(\cdot), \phi(\cdot))$. Let $\hat{b}(x, m) = \hat{B}(\phi(\cdot, x, m))$ and $b(t) = \hat{b}(\mu(t), \gamma(t))$. For any x not chosen in equilibrium ($x \neq \mu(t)$ for all $t \in [0, 1]$), let $\underline{b}(x) = \sup_{t:\mu(t)<x} b(t)$ and $\bar{b}(x) = \inf_{t:\mu(t)>x} b(t)$.

Definition. We shall say that an equilibrium satisfies the *monotonic D1* criterion if the following conditions are satisfied: (i) The function $\mu(t)$ is weakly increasing in t. (ii) For all x and m, $\hat{b}(x, m) \geq \sup_{x'<x, m' \in [0,1]} \hat{b}(x', m')$. (iii) Consider some x' such that $\mu(t) \neq x'$ for all $t \in [0, 1]$. Suppose that there exist t', t_1, and t_2 with $t_1 < t_2$ and $t' \notin [t_1, t_2]$ such that if

$$U(x', t) + \beta V(b, t) \geq U(\mu(t), t) + \beta V(b(t), t)$$

for some $t \in [t_1, t_2]$ and $b \in [\underline{b}(x'), \bar{b}(x')]$, then

$$U(x', t') + \beta V(b, t') > U(\mu(t'), t') + \beta V(b(t'), t').$$

Then, for all m, $\phi(t_1, x', m) = \phi(t_2, x', m)$.

Parts i and ii, respectively, simply require monotone action and inference functions. If one substituted $b \in [0, 1]$ for $b \in [\underline{b}(x'), \bar{b}(x')]$, then part iii would be equivalent to the D1 criterion, which requires that a receiver not attribute a deviation to a particular type if there is another type that is willing to make the deviation for a strictly larger set of inferences. For the monotone D1 criterion, we restrict the possible inferences in advance to make sure that they respect monotonicity.

Theorem 5. Let $\mathcal{E}$ denote the set of all response functions, $(\mu(t), \gamma(t))$, associated either with some central pooling equilibrium or with the separating action function (if it exists) for which $\mu(0) = \underline{x}$ and $\mu(1) = \bar{x}$. Let $\mathcal{D}$ be the set of all response functions associated with monotonic D1 equilibria. Then $\mathcal{E} = \mathcal{D}$.

C. *Sketches of Proofs*

At the request of the editors, we provide abbreviated sketches of proofs. Complete proofs are available at http://www.stanford.edu/~bernheim/. Though a number of details have changed, the reader can also consult the working paper version of this article (Bernheim and Severinov 2000).

Proof of Theorem 1 (Sketch)

First observe that $\mu(0) \geq \underline{x} > 0 = X(0)$ and $\mu(1) \leq \bar{x} < 1 = X(1)$. Consequently, as t moves from zero to one, $\mu(t)$ must cross over $X(t)$.

We argue that $\mu(t)$ cannot coincide with $X(t)$, except at $t = \frac{1}{2}$. If $\mu(t) = X(t)$ for some $t < \frac{1}{2}$, then some type $t' = t - \epsilon$ (for small $\epsilon > 0$) would imitate t. Intuitively, for such t', imitation yields a first-order gain from the change in inference, but at most a second-order loss from the change in action. A symmetric argument rules out $\mu(t) = X(t)$ for $t > \frac{1}{2}$.

Next we argue that $\mu(t)$ cannot jump discontinuously over $X(t)$. If there was such a discontinuity, then, for all $\epsilon > 0$, one could find t', t'' with $\Delta t = t'' - t' = \epsilon > 0$ but $\Delta\mu = \mu(t') - \mu(t'')$ positive and bounded away from zero. As ϵ gets small, the difference between the loss in utility for types t' and t'' occurring when the inference changes from t'' to t' is on the order of $(\Delta t)^2$. However, the difference between the gain in utility for types t' and t'' occurring when the action changes from $\mu(t'')$ to $\mu(t')$ is on the order of Δt. Since type t'' values an increase in action more than t', if t' is not hurt by a switch from $(\mu(t''), t'')$ to $(\mu(t'), t')$, then, for Δt sufficiently small, t'' must strictly benefit from the same switch.

From the preceding discussion, we conclude that $\mu(t) > X(t)$ for $t < \frac{1}{2}$, $\mu(t) < X(t)$ for $t > \frac{1}{2}$, and $\mu(\frac{1}{2}) = X(\frac{1}{2}) = \frac{1}{2}$.

To establish continuity, suppose that there is a sequence t_k converging to some t_d with $\mu(t_k)$ converging to $\bar{\mu} \neq \mu(t_d)$. Since $\mu(t) > X(t)$, if $\bar{\mu} > \mu(t_d)$, then t_k would imitate t_d for large enough *k;* if $\bar{\mu} < \mu(t^d)$, then t_d would imitate t_k for large enough k.

To establish monotonicity, note that $\mu(t) < \mu(t')$ for all t and t' with $t < t' < p(t) < \frac{1}{2}$; if not, then $\mu(t) > \mu(t') > X(t') > X(t)$, which implies that t would strictly prefer the action and inference assigned to t'. The argument for monotonicity on $(\frac{1}{2}, 1]$ is analogous. Q.E.D.

Proof of Theorem 2 (Sketch)

First we claim that there exists a unique separating function $\underline{\mu}(t)$ on $t \in [0, \frac{1}{2})$ satisfying $\underline{\mu}(t) = \underline{x}$. (A similar statement holds for the upper half of the type space.) From theorem 1, we know that, in our search for separating functions, we can confine attention to $\underline{\mu}(t)$ that are strictly increasing and continuous and that satisfy $\underline{\mu}(t) > X(t)$ for $t \in [0, \frac{1}{2})$. One can show that any $\underline{\mu}(t)$ with these properties is a separating function if and only if it satisfies the following condition:

$$\underline{\mu}'(t) = -\frac{\beta V_1(t, t)}{U_1(\underline{\mu}(t), t)} \tag{A1}$$

at all points of differentiability. Note that (A1) is a first-order differential equation for $\underline{\mu}(t)$. Unfortunately, since $U_1(X(t), t) = 0$ for all t, the right-hand side of (A1) is not Lipschitz on $(t, x) \in [0, \frac{1}{2}) \times [\underline{x}, 1]$. A slight modification of standard arguments nevertheless establishes existence and uniqueness.

Next we claim that there exists $\underline{\beta}^*(\underline{x})$ such that (i) for $\beta \leq \underline{\beta}^*(\underline{x})$, $\lim_{t\uparrow 1/2}\underline{\mu}(t) = \frac{1}{2}$, and for $\beta > \underline{\beta}^*(\underline{x})$, $\lim_{t\uparrow 1/2}\underline{\mu}(t) > \frac{1}{2}$; and (ii) $\underline{\beta}^*(\underline{x}) > 0$ iff $\underline{x} < \frac{1}{2}$.

(Since the problem is completely symmetric apart from endowments, it also follows that, for $\beta \le \underline{\beta}^*(\bar{x})$, $\lim_{t\downarrow 1/2}\bar{\mu}(t) = \frac{1}{2}$, and for $\beta > \underline{\beta}^*(\bar{x})$, $\lim_{t\downarrow 1/2}\bar{\mu}(t) < \frac{1}{2}$.) The case of $\underline{x} \ge \frac{1}{2}$ is trivial, so consider $\underline{x} < \frac{1}{2}$. For large enough β, nonimitation of type $t' = p(t)$ by type t requires t' to select an action greater than $\frac{1}{2}$. Since $\underline{\mu}$ is monotonic, $\lim_{t\uparrow 1/2}\underline{\mu}(t) > \frac{1}{2}$ for large β. From an inspection of the differential equation (A1), one can see that the slope of the solution through any given point is steeper for higher β. Therefore, for all $t \in [0, \frac{1}{2})$, $\underline{\mu}(t)$ (and hence $\lim_{t\uparrow 1/2}\underline{\mu}(t)$) is increasing in β. This establishes the existence of $\underline{\beta}^*(\underline{x})$ satisfying property i. Now consider a straight line L in the (t, x) plane passing through the point $(\frac{1}{2}, \frac{1}{2})$ and intersecting the vertical axis above $\underline{x}$. By taking $\beta > 0$ sufficiently small, one can make the slope of $\underline{\mu}(t)$ evaluated at any point on L less than the slope of L (this follows from [A1], though some work is required to show that the same value of β will work for all points close to $(\frac{1}{2}, \frac{1}{2})$). For such $\beta > 0$, $\underline{\mu}(t)$ can never cross L; hence $\lim_{t\uparrow 1/2}\underline{\mu}(t) = \frac{1}{2}$. This implies that $\underline{\beta}^*(\underline{x})$ satisfies property ii.

To complete the proof, simply define $\beta^*(\underline{x}, \bar{x}) \equiv \min\{\underline{\beta}^*(\underline{x}), \underline{\beta}^*(\bar{x})\}$. $\beta \le \beta^*(\underline{x}, \bar{x})$ is necessary and sufficient for $\lim_{t\uparrow 1/2}\underline{\mu}(t) = \frac{1}{2} = \lim_{t\downarrow 1/2}\bar{\mu}(t)$, which in turn is necessary and sufficient for the existence of a separating equilibrium. It is easy to verify that $\underline{\mu}(t)$ is increasing in $\underline{x}$, from which it follows that $\underline{\beta}^*(\underline{x})$ is decreasing in $\underline{x}$. Consequently, $\min\{\underline{\beta}^*(\underline{x}), \underline{\beta}^*(\bar{x})\} = \underline{\beta}^*(\max\{\underline{x}, \bar{x}\})$. Q.E.D.

Proof of Theorem 3 (Sketch)

Define $t^* \in (0, \frac{1}{2})$ such that $\underline{\mu}(t^*) = \bar{\mu}(1 - t^*)$. Also, for all $t \in [0, t^*]$, define $x_p(t) \in [\underline{\mu}(t), \bar{\mu}(1 - t)]$ to be the value of x that solves

$$U(\underline{\mu}(t), t) - U(x, t) = U(\bar{\mu}(1 - t), 1 - t) - U(x, 1 - t). \tag{A2}$$

It is easy to verify the existence and uniqueness of both t^* and $x_p(t)$.

Now consult figure A2. By definition, $(t^*, x_p(t^*))$ lies on $\underline{\mu}(t)$. Define $\tilde{t} < t^*$ as the type that is indifferent between $(\tilde{t}, \underline{\mu}(\tilde{t}))$ and $(p(\tilde{t}), x_p(\tilde{t}))$ (graphically, note the tangency between the horizontal line at $x_p(\tilde{t})$ and the type $\tilde{t}$ indifference curve through $(\tilde{t}, \underline{\mu}(\tilde{t}))$, labeled $\tilde{I}$). Again, existence and uniqueness are easy to establish.

Consider $t' \in (\tilde{t}, t^*)$. As indicated in the figure, there are two crossings between the horizontal line at $x_p(t')$ and the type t' indifference curve through $(t', \underline{\mu}(t'))$, labeled I'. If we attempt to construct a pooling equilibrium by setting $t_1 = t'$, then, applying the indifference condition (6), we know that there are two choices for the highest type in the first segment of the central pool (t_2). In the figure, these choices correspond to the horizontal coordinates of the points $s_1^L(t')$ and $s_1^U(t')$. Next, applying the indifference condition (7), we know that each of these choices maps into a unique choice for the highest type in the second segment of the central pool (t_3). In the figure, these choices correspond to the horizontal coordinates of the points $s_2^L(t')$ (under the assumption that we began with $s_1^L(t')$) and $s_2^U(t')$ (under the assumption that we began with $s_1^U(t')$). With repeated applications of (7), we generate two sequences of points. For large enough k, $s_k^U(t')$ must cross the vertical line at $t = \frac{1}{2}$ (in the figure, this occurs for $k = 2$).

Note that when $t = \tilde{t}$, $s_k^L(\tilde{t}) = s_k^U(\tilde{t})$ (the two crossings converge as $t \to \tilde{t}$). Moreover, when $t = t^*$, $s_k^L(t^*) = s_{k-1}^U(t^*)$ (in the case of $s_1^L(t)$, the first segment of the central pool shrinks as t approaches t^* and becomes degenerate in the limit).

As we have drawn the figure, the horizontal coordinate of $s_2^U(t^*)$ exceeds $\frac{1}{2}$. As we reduce t continuously from t^* to $\tilde{t}$, $s_2^U(t)$ traces out a continuous path,

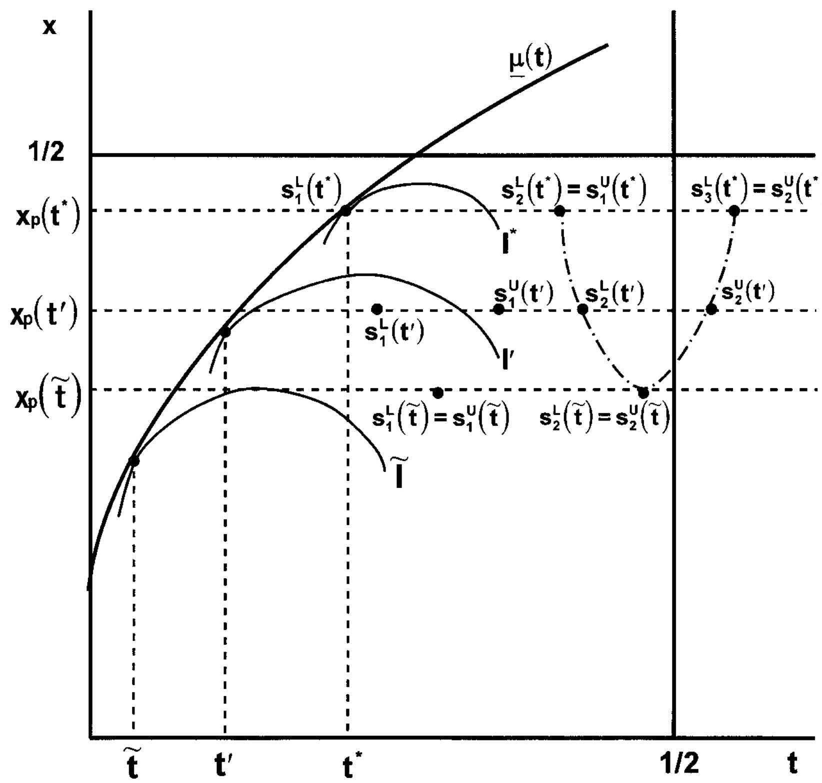

FIG. A2.—Existence of a central pooling equilibrium

shown as the dashed curve in the figure. Once we get to $\tilde{t}$, we switch to $s_2^L(\tilde{t})$ (which is, of course, the same point). As we increase t continuously from $\tilde{t}$ to t^*, $s_2^L(t)$ also traces out a continuous path, shown as the continuation of the dashed curve. Because the path from $s_2^L(t^*)$ to $s_2^L(t^*)$ is unbroken, it must intersect the vertical line at $\frac{1}{2}$ for at least one value of t, say t''. Construct a central pooling equilibrium by setting $t_l = t''$, $t_h = 1 - t''$, and $x_p = x_p(t'')$ and partitioning the central pool into segments as dictated by the indifference condition. By construction of $x_p(t)$, the segments are symmetric on both sides of $\frac{1}{2}$, and they meet in the middle, at $\frac{1}{2}$. Symmetry of the segments guarantees that type $\frac{1}{2}$, which lies at the boundary between two segments, is indifferent between them. Q.E.D.

Proof of Theorem 4 (Sketch)

The first step of the proof is to construct a function $\sigma(\tau, x, x_p)$, defined for τ in some interval $[0, \bar{\tau}]$, that describes configurations of the central pool that are consistent with equilibrium starting from one side of the parent type set. In particular, for $\tau \in [0, 1]$, the values $t_1 = \sigma(\tau, x, x_p)$ and $t_2 = \sigma(\tau + 1, x, x_p)$ satisfy equation (6), and for $\tau \in [1, \bar{\tau} - 1]$, the values $t_{n-1} = \sigma(\tau - 1, x, x_p)$, $t_n = \sigma(\tau, x, x_p)$, and $t_{n+1} = \sigma(\tau + 1, x, x_p)$ satisfy equation (7). Using a construction related to the one used in the proof of theorem 3 (for $s_1^L(t)$ and $s_1^U(t)$), one can guarantee that $\sigma(\tau, x, x_p)$ is differentiable with bounded derivatives. As τ varies from zero to one, $(\sigma(\tau, x, x_p), \sigma(\tau+1, x, x_p), \ldots)$ maps out various configurations of possible endpoints for segments of the central pool.

For any pooling action x_p and pair of endowments $\underline{x}$ and $\bar{x}$, define

$$\xi(\tau_1, \tau_2, \underline{x}, \bar{x}, x_p) \equiv \begin{bmatrix} \sigma(\tau_1, \underline{x}, x_p) + \sigma(\tau_2+1, \bar{x}, 1-x_p) - 1 \\ \sigma(\tau_1+1, \underline{x}, x_p) + \sigma(\tau_2, \bar{x}, 1-x_p) - 1 \end{bmatrix}.$$

Note that if $\xi(\tau_1, \tau_2, \underline{x}, \bar{x}, x_p) = 0$, then the existence of a solution to equations (6)–(8) is guaranteed, and a central pooling equilibrium exists. If $\det D_\tau \xi(\tau_1, \tau_2, \underline{x}, \bar{x}, x_p) \neq 0$ at any (τ_1, τ_2) for which $\xi(\tau_1, \tau_2, \underline{x}, \bar{x}, x_p) = 0$, then, by the implicit function theorem, for any $(\underline{x}', \bar{x}', x_p')$ in some neighborhood of $(\underline{x}, \bar{x}, x_p)$, there exists (τ_1', τ_2') satisfying $\xi(\tau_1', \tau_2', \underline{x}', \bar{x}', x_p') = 0$.

Using an argument similar to the one given in the proof of theorem 3, one can show that, when $\underline{x} = \bar{x} = x$ (equal endowments), there is always a central pooling equilibrium corresponding to the value of τ^* satisfying $\sigma(\tau^*, x, \frac{1}{2}) = 1 - \sigma(\tau^*+1, x, \frac{1}{2})$ (i.e., equal division and a symmetric partition, where the central segment is centered at $\frac{1}{2}$). Consequently, if $\det D_\tau \xi(\tau^*, \tau^*, x, x, \frac{1}{2}) \neq 0$ at some τ^* for which $\sigma(\tau^*, x, \frac{1}{2}) = 1 - \sigma(\tau^*+1, x, \frac{1}{2})$, then there exists a central pooling equilibrium with equal division for all endowments sufficiently close to (x, x).

Note that

$$\det D_\tau \xi(\tau^*, \tau^*, x, x, \tfrac{1}{2}) = [\sigma_1(\tau^*, x, \tfrac{1}{2})]^2 - [\sigma_1(\tau^*+1, x, \tfrac{1}{2})]^2.$$

Thus, as long as $|\sigma_1(\tau^*, x, \frac{1}{2})| \neq |\sigma_1(\tau^*+1, x, \frac{1}{2})|$, we have $D_\tau \xi(\tau^*, \tau^*, x, x, \frac{1}{2}) \neq 0$, which implies the existence (locally) of central pooling equilibria with equal division and asymmetric endowments.

Define Λ_x as the set of functions $v \in \Lambda$ for which $|\sigma_1(\tau, x, \frac{1}{2})| \neq |\sigma_1(\tau+1, x, \frac{1}{2})|$ at all τ with $\sigma(\tau, x, \frac{1}{2}) = 1 - \sigma(\tau+1, x, \frac{1}{2})$. We claim that Λ_x is open dense in Λ.

We demonstrate that Λ_x is open by showing that $\Lambda \backslash \Lambda_x$ is closed. First we define the function $\psi(\tau) = \sigma(\tau, x, \frac{1}{2}) + \sigma(\tau+1, x, \frac{1}{2}) - 1$ and show that the correspondence mapping Λ into values of τ satisfying $\psi(\tau) = 0$ is upper hemicontinuous. Using the fact that a function v is in $\Lambda \backslash \Lambda_x$ iff there is a value τ with $\psi(\tau) = 0$ and $|\sigma_1(\tau, x, \frac{1}{2})| = |\sigma_1(\tau+1, x, \frac{1}{2})|$, one can then show that there exists a value of τ with the same properties for any function in the closure of $\Lambda \backslash \Lambda_x$.

Finally, we argue that Λ_x is dense. This requires two steps. The first step is to define Λ_x^R as the set of functions v for which zero is a regular value of $\psi(\tau)$ and to show that Λ_x^R is dense in Λ. This is accomplished by introducing a perturbation function, parameterized by a scalar q (with $q = 0$ indicating no perturbation) such that v depends on q only in a small neighborhood of $\frac{1}{2}$ and is increasing in q within this neighborhood. One can verify that $\psi_q(\tau) > 0$ at any τ satisfying $\psi(\tau) = 0$. The transversality theorem then implies that zero is a regular value of $\psi(\tau)$ for almost all (in the sense of full measure) q. But then one can find q arbitrarily close to zero (and hence a function arbitrarily close to v) for which zero is a regular value of $\psi(\tau)$.

The second step is to show that Λ_x^R lies in the closure of Λ_x. Take any $v \in \Lambda^R$. We consider a perturbation that leaves the value of $v(t)$ unchanged at $t = \frac{1}{2}$ and outside of a small neighborhood of $\frac{1}{2}$ and that changes $v'(\frac{1}{2})$. By construction, this perturbation leaves unchanged the set of τ such that $\psi(\tau) = 0$. It is easy to verify that changing $v'(\frac{1}{2})$ alters $\sigma_1(\tau + 1, x, \frac{1}{2})$, but not $\sigma_1(\tau, x, \frac{1}{2})$. Since the set of τ such that $\psi(\tau) = 0$ is finite (zero is a regular value of $\psi(\tau)$ and the range of τ is compact), one can always choose an arbitrarily small perturbation such that $|\sigma_1(\tau, x, \frac{1}{2})| \neq |\sigma_1(\tau + 1, x, \frac{1}{2})|$ for all τ with $\psi(\tau) = 0$. Q.E.D.

Proof of Theorem 5 (Sketch)

The main part of the proof involves establishing that $\mathcal{D} \subseteq \mathcal{E}$. We demonstrate this by showing that any equilibrium with monotonic actions and an action pool not containing $\frac{1}{2}$ cannot satisfy the monotonic D1 criterion.

Since actions are monotonic in type, an action pool is a nondegenerate interval, $[t_l, t_h]$. Since $V_{12} > 0$, every segment within $[t_l, t_h]$ is also an interval. It is easy to show that t_h must belong to a nondegenerate segment, which implies $b(t_h) < t_h$.

Let $\mu' = \lim_{t \downarrow t_h} \mu(t)$ and $b' = \lim_{t \downarrow t_h} b(t)$. Clearly, $b' \geq t_h > b(t_h)$. It is straightforward to show that, in equilibrium, t_h must be indifferent between $(\mu(t_h), b(t_h))$ and (μ', b') and that $\mu' > \mu(t_h)$.

Choose some $\mu_d \in (\mu(t_h), \mu')$; note that this action is not chosen in equilibrium by any type. Let m be an arbitrary message. We claim that, if the monotonic D1 criterion is satisfied, $\hat{b}(\mu_d, m) = t_h$. But then t_h would deviate to μ_d (provided that we have chosen μ_d sufficiently close to $\mu(t_h)$), which overturns the equilibrium.

To establish the claim, we need to demonstrate that, for all $b \in [\underline{b}(\mu_d), \bar{b}(\mu_d)]$ and $t \neq t_h$, $U(\mu_d, t) + \beta V(b, t) \geq U(\mu(t), t) + \beta V(b(t), t)$ implies $U(\mu_d, t_h) + \beta V(b, t_h) > U(\mu(t_h), t_h) + \beta V(b(t_h), t_h)$. Consider $t < t_h$. Suppose that t weakly prefers (μ_d, b) to $(\mu(t), b(t))$. Then t must also weakly prefer (μ_d, b) to $(\mu(t_h), b(t_h))$. But then, since $b \geq \underline{b}(\mu_d) = b(t_h)$, $\mu(t_h) < \mu_d$, $V_{12} > 0$, and $U_{12} > 0$, t_h must strictly prefer (μ_d, b) to $(\mu(t_h), b(t_h))$. The argument for $t > t_h$ is similar, except that it uses $b \leq \bar{b}(\mu_d) = b'$ and $\mu_d < \mu'$.

To complete the proof, one must show that $\mathcal{E} \subseteq \mathcal{D}$. The separating equilibrium, if it exists, obviously satisfies the monotonic D1 criterion (since actions are monotonic and all actions are taken by some type in equilibrium). Consider any central pooling equilibrium characterized by the three parameters t_l, t_h, and x_p. The set of out-of-equilibrium actions is $(\underline{\mu}(t_l), x_p) \cup (x_p, \bar{\mu}(t_h))$. Construct out-of-equilibrium beliefs as follows: for $x \in (\underline{\mu}(t_l), x_p)$, let $\phi(t_l, x, m) = 1$ and $\phi(t, x, m) = 0$ for $t < t_l$, and for $x \in (x_p, \bar{\mu}(t_h))$, let $\phi(t_h, x, m) = 1$ and $\phi(t, x, m) = 0$ for $t < t_h$. It is easy to check that this is an equilibrium. One establishes that it satisfies the monotonic D1 criterion through an argument analogous to that used in the first part of this proof. Q.E.D.

References

Austen-Smith, David, and Banks, Jeffrey S. "Cheap Talk and Burned Money." *J. Econ. Theory* 91 (March 2000): 1–16.

Bank, Stephen P., and Kahn, Michael D. *The Sibling Bond.* New York: Harper Collins, 1997.

Banks, Jeffrey S. "A Model of Electoral Competition with Incomplete Information." *J. Econ. Theory* 50 (April 1990): 309–25.

Banks, Jeffrey S., and Sobel, Joel. "Equilibrium Selection in Signaling Games." *Econometrica* 55 (May 1987): 647–61.

Barro, Robert J. "Are Government Bonds Net Wealth?" *J.P.E.* 82 (November/December 1974): 1095–1117.

Bednar, Richard L., and Peterson, Scott R. *Self-Esteem: Paradoxes and Innovations* *in Clinical Theory and Practice.* 2d ed. Washington: American Psychological Assoc., 1995.

Bernheim, B. Douglas. "How Strong Are Bequest Motives? Evidence Based on Estimates of the Demand for Life Insurance and Annuities." *J.P.E.* 99 (October 1991): 899–927.

———. "A Theory of Conformity." *J.P.E.* 102 (October 1994): 841–77.

Bernheim, B. Douglas, and Bagwell, Kyle. "Is Everything Neutral?" *J.P.E.* 96 (April 1988): 308–38.

Bernheim, B. Douglas, and Severinov, Sergei. "Bequests as Signals: An Explanation for the Equal Division Puzzle." Working Paper no. 7791. Cambridge, Mass.: NBER, July 2000.

Bernheim, B. Douglas; Shleifer, Andrei; and Summers, Lawrence H. "The Strategic Bequest Motive." *J.P.E.* 93 (December 1985): 1045–76.

Brody, Gene H.; Stoneman, Zolinda; and McCoy, J. Kelly. "Contributions of Family Relationships and Child Temperaments to Longitudinal Variations in Sibling Relationship Quality and Sibling Relationship Styles." *J. Family Psychology* 8 (September 1994): 274–86.

Cho, In-Koo, and Kreps, David M. "Signaling Games and Stable Equilibria." *Q.J.E.* 102 (May 1987): 179–221.

Cho, In-Koo, and Sobel, Joel. "Strategic Stability and Uniqueness in Signaling Games." *J. Econ. Theory* 50 (April 1990): 381–413.

Chu, C. Y. Cyrus. "Primogeniture." *J.P.E.* 99 (February 1991): 78–99.

Coopersmith, Stanley. *The Antecedents of Self-Esteem.* San Francisco: Freeman, 1967.

Crawford, Vincent P., and Sobel, Joel. "Strategic Information Transmission." *Econometrica* 50 (November 1982): 1431–51.

Dunn, Thomas, and Phillips, John. "Do Parents Divide Resources Equally among Children? Evidence from the AHEAD Survey." Aging Studies Program Paper no. 5. Syracuse, N.Y.: Syracuse Univ., 1997.

Gale, William G., and Scholz, John Karl. "Intergenerational Transfers and the Accumulation of Wealth." *J. Econ. Perspectives* 8 (Fall 1994): 145–60.

Guinnane, Timothy W. "Intergenerational Transfers, Emigration, and the Rural Irish Household System." *Explorations Econ. Hist.* 29 (October 1992): 456–76.

Hurd, Michael D. "Savings of the Elderly and Desired Bequests." *A.E.R.* 77 (June 1987): 298–312.

Lundholm, Michael, and Ohlsson, Henry. "Post Mortem Reputation, Compensatory Gifts and Equal Bequests." *Econ. Letters* 68 (August 2000): 165–71.

McGarry, Kathleen. "Inter Vivos Transfers and Intended Bequests." Manuscript. Los Angeles: Univ. California, Dept. Econ., 1998.

Menchik, Paul L. "Primogeniture, Equal Sharing, and the U.S. Distribution of Wealth." *Q.J.E.* 94 (March 1980): 299–316.

———. "Unequal Estate Division: Is It Altruism, Reverse Bequests, or Simply Noise?" In *Modelling the Accumulation and Distribution of Wealth,* edited by Denis Kessler and André Masson. New York: Oxford Univ. Press, 1988.

Schelling, Thomas. *The Strategy of Conflict.* Cambridge, Mass.: Harvard Univ. Press, 1960.

Spence, A. Michael. *Market Signaling: Informational Transfer in Hiring and Related Screening Processes.* Cambridge, Mass.: Harvard Univ. Press, 1974.

Tesser, Abraham. "Self-Esteem Maintenance in Family Dynamics." *J. Personality and Soc. Psychology* 39 (July 1980): 77–91.

Tomes, Nigel. "The Family, Inheritance, and the Intergenerational Transmission of Inequality." *J.P.E.* 89 (October 1981): 928–58.

Unur, A. Sinan. "The Role of Incentives in Parents' Transfers to Children." Manuscript. Ithaca, N.Y.: Cornell Univ., 1998.

Wilhelm, Mark O. "Bequest Behavior and the Effect of Heirs' Earnings: Testing the Altruistic Model of Bequests." *A.E.R.* 86 (September 1996): 874–92.