Tough Love for Lazy Kids:

Dynamic Insurance and Equal Bequests

Ctirad Slavík[†] Kevin Wiseman[‡]

November 1, 2017

Abstract. This paper develops a dynamic insurance model to explain a central puzzle in intergenerational transfers: gifts partially compensate children for negative income shocks, but bequests are typically divided equally. In the model, parents use gifts (early in life) and bequests (later in life) to provide insurance against income shocks, but take into account that children would shirk if offered large transfers. We show in a simple model that parents can provide better incentives later in life by giving equal bequests. In a quantitative model, gifts are compensatory while bequests are nearly uncorrelated with income and approximately equal in most families.

JEL Classification: D19, D64, D82, E24, J1

Keywords: Intergenerational Transfers, Equal Bequest Division Puzzle, Private Information

[†] Ctirad Slavík, CERGE-EI. A joint workplace of the Center for Economic Research and Graduate Education, Charles University, and the Economics Institute of the Academy of Sciences of the Czech Republic. Politických vězňů 7, Praha 1, 111 21, Czech Republic. Email: ctirad.slavik@cerge-ei.cz.

[‡] Kevin Wiseman, International Monetary Fund. Strategy, Policy, and Review Department. 700 19th Street NW Washington, DC 20431, USA. Email: kwiseman@imf.org.

We would like to thank the editor, the associate editor and two anonymous referees as well as Alex Bick, Nicola Fuchs-Schündeln, Ioanna Grypari, Nezih Guner, Larry Jones, Marek Kapička, Tim Kehoe, Narayana Kocherlakota, Sang Yoon Lee, Fabrizio Perri, Chris Phelan, José-Víctor Ríos-Rull, Sigrid Röhrs, Michèle Tertilt, and participants at various conferences, workshops and seminars for their comments. We are grateful to Jun Li for excellent research assistance. All remaining errors are ours. The views expressed in this paper are those of the authors and do not necessarily represent those of the IMF or IMF policy.

1. Introduction

Understanding the transfers from parents to their children plays an important role in understanding the distribution of wealth, consumption inequality, and the effects of government redistribution policies. Yet the evidence about the ways parents distribute their money has defied explanation. In empirical studies, bequests are typically found to be divided equally among children, independently of their income. By contrast, parents give more to their poorer children in inter vivos transfers, or gifts, but not enough to fully compensate the differences in child income within the family.

Why are gifts to poorer children larger, while bequests are divided equally? This paper explains this behavior as a result of the timing of gifts and bequests in a dynamic moral hazard problem. In the model, children's income is a combination of exogeneous productivity, and endogeneous labor effort. Parents know their children's income but not their productivity or labor effort. An altruistic parent wants to help her unlucky low-productivity children, but she cannot distinguish between hardworking low-productivity children and lazy high-productivity children. If the parent gives too much money to low-income children, she encourages high-productivity children to slack off, earn a low income, and receive the transfer. The parent's best option is to provide partial insurance (i.e., give more money to her lower-income children), but not enough to fully compensate differences in income.

A parent could simply hand out the same gift and bequest to a given child. By the logic outlined above, poorer children would receive both a higher gift and a higher bequest than their richer siblings. However, the parent can do better by thinking about gifts and bequests separately. In the model, gifts are received early in a child's life, whereas bequests are received later in life. Children face uncertainty about their future productivity. The parent can take advantage of these two features when deciding about gifts and bequests.

The relative sizes of gifts and bequests are determined by productivity persistence. In our model, no productivity persistence corresponds to full mean reversion: all children face the same probability of becoming high- or low-productive in the future. In this case, the currently high-productive children are concerned about becoming low-productive in the future. Therefore, they care relatively more about bequests, which transfer resources to the future. In contrast, the currently low-productive children care relatively more about the gift. The parent takes advantage of these differences in time preferences and uses bequests to provide incentives and gifts to provide insurance: the bequest to the high types is larger than the bequest to the low types while the opposite is true for gifts.

On the other hand, with perfect persistence there is no mean reversion and no uncertainty: both types know that they will be of the same type in the future as they are today. In this case, both types care as much about the current transfer (gift) as they care about the future transfer (bequest). Therefore, both gifts and bequests are higher for the low types. For intermediate levels of persistence, bequests are about equal, and for some level they are exactly equal. Partial insurance implies that if bequests are equal, gifts are higher for the low types.

To summarize, transfers are weakly progressive in the sense that poorer children are given a higher transfer, but differences in income between siblings are not fully compensated. Gifts are distributed more progressively than bequests. Bequests are progressive when income is highly persistent, regressive when income is very impersistent, and equal or nearly equal for intermediate values. In a simple model with two productivity types, we demonstrate that there is a level of persistence at which bequests are exactly equal and gifts are weakly progressive, which is qualitatively what we observe in the data. This pattern cannot occur in the public information version of the model, in which the parent observes child productivity and effort and thus does not need to provide incentives.

To compare the performance of the model to the data, we build a richer version with many productivity types. We then pin down the children's productivity persistence in the model with U.S. data, solve the model numerically, and compute a number of transfer statistics. Certain features of the U.S. economy, such as the fact that equal bequest division is the default legal option, remain unmodelled. However, considering its simplicity, the model approximates the data reasonably well. In the benchmark parameterization, bequests are nearly uncorrelated with child income. As in the data, gifts are weakly progressive (albeit more progressive than in the data). Importantly, both gifts and bequests are substantially more progressive in the public information version of the model. Siblings' bequests are approximately equal (i.e., within 25% of the intrafamily mean) in a large number of families, as in the data. A rule of thumb in which parents divide their bequests equally results in a minimal loss in welfare compared to the optimal allocation. This welfare loss is more than 20 times smaller than the equivalent loss in the public information version of the model. Adding private information thus helps to bring the pattern of intergenerational transfers closer to the data.

We also examine the sensitivity of these results to alternative parameterizations and alternative specifications of the model. In all cases, gifts are weakly progressive but more progressive than bequests. Bequests are nearly uncorrelated with income and concentrated around equal division for a large number of families.

The rest of the paper is organized as follows. The next section provides a more detailed summary of the puzzle in the empirical literature. Section 3 then reviews papers that have tried to explain this puzzle, and discusses the place of this paper in the dynamic insurance literature. Section 4 introduces a simple version of the model, derives the main analytical results, highlights the forces at work in the model, and considers the model's robustness to a number of extensions. Section 5 evaluates the benchmark quantitative model with respect to the data on intergenerational transfers, compares it to a version of the model without incentive problems, and conducts sensitivity analysis. Section 6 concludes.

2. Empirical Evidence

This section first summarizes the evidence on the size of intergenerational transfers present in the Panel Study of Income Dynamics (PSID). Both bequests and gifts are substantial for those households that report a nonnegative bequest/gift. In a given year, approximately 2% of households report a bequest. The ratio of the average bequest relative to average annual labor earnings of households that report a positive bequest is 98.7% between 1988 - 2011. Taking all households in the sample, including those that report no bequest, the ratio of the average bequest relative to average annual labor earnings is 2.22% between 1988 and 2013. On average, 8.23% of households report a gift in a given year and this number has been rising. The average gift is substantial at approximately 15% relative to the recipient households' average annual labor earnings. Taking all households, including those that report no gift, the average gift between 1985 and 2013 is 0.65% relative to average annual labor earnings. The details of our empirical work with the PSID are contained in section 5 and in appendix A. In the PSID, however, one cannot uniquely match the source parent and recipient child. Therefore, for the within-family patterns of bequests and gifts, we rely on the existing literature.

As for the within-family pattern of bequests, Menchik (1980) finds in the Connecticut state tax records that in two-child families, 62.5% of bequests are divided exactly equally and 70.5% of bequests are divided almost equally (within 2% of the mean) across children. More recently, Wilhelm (1996) finds in a larger sample of estate tax records, which includes families of various sizes, that 68.6% of bequests are divided exactly equally, 76.6% almost equally (within 2% of the mean), and 88% approximately equally (within 25% of the mean) across children. He finds that even in families in which bequests are unequally divided, the difference in bequests does not vary with income differences (in a statistical sense). McGarry (1999) finds that 83% of respondents report that their wills treat all children about equally in the Asset and Health Dynamics Study survey (AHEAD). Light and McGarry (2004) find that 92.1% of mothers who have a will say that their estate will be divided equally among their children in the 1999 National Longitudinal Surveys (NLS) of Mature Women and Young Women. Other empirical work has documented the same pattern in other countries.[1]

Recent empirical work documents a very different pattern of transfers while parents are alive. McGarry and Schoeni find that lower income increases both the probability and size of a gift in Health and Retirement Survey (1994) and AHEAD data (1995). Hochguertel and Ohlsson (2009) analyze six waves (1992 - 2002) of the HRS and find that conditional on giving a gift, only 9.2% of parents divide gifts equally and only 10.9% of parents divide gifts so that they are within 20% of the intra-family mean. They also find that gifts are decreasing with children's income, but they do not fully compensate the differences in child income within the family (a 1$ difference in children's income is compensated by 2 cents in transfers, where the number is statistically different from zero). McGarry (2016) reports similar results using the HRS up until 2008. In a similar vein, Altonji, Hayashi, and Kotlikoff (1997) document partial insurance between parents and children in the Panel Study of Income Dynamics (PSID).

Dunn and Phillips (1997) find in the AHEAD data set that the probability of receiving a gift is decreasing with child income, while the probability of receiving a bequest does not depend on child income. Their paper is one of the first to emphasize the difference in

[1]Arrondel and Masson (2002) report that in France, unequal estate division, which concerns less than 8% of estate declarations, in 80% of the cases occurs only through unequal previous gifts, while bequests remain equally divided. Horioka (2009) reports that in Japan, 48.16% of respondents with two or more children plan to divide their bequest equally, 29.90% do not plan to leave a bequest, and 21.94% plan to divide their bequest unequally. Ohlsson (2007) reports that in Sweden, bequests are unequally divided in fewer than 15% of cases. Using a recent international survey concerning bequest plans, Horioka (2014) reports that equal bequest division occurs in 92.55% of cases in the United States, in 84.17% of cases in India, in 72.67% of cases in Japan and in 70.28% of cases in China.

bequest and gift behavior. Subsequent papers have tried to explain these differences, but to our knowledge this paper is the first to succeed on a qualitative level, meaning that it delivers equal bequests and weakly progressive gifts. In addition, this paper brings the model to the data and evaluates its performance quantitatively, a first for models in this literature. The next section proceeds with a more detailed discussion of the existing theories.

3. Related Literature

The related theoretical literature can be divided into four groups. The first one focuses on the patterns of intergenerational transfers *across* households, but not within households. The second group of papers studies within household patterns, but typically resorts to a behavioral explanation. The third group studies how moral hazard problems affect intergenerational transfers. These models are, however, static and, therefore, cannot distinguish between gifts and bequests. Finally, the most closely related papers to the present paper are those that study transfers in the context of dynamic moral hazard models. In what follows, these four groups of papers are discussed in more detail.

McGarry (1999) studies a one-parent, one-child model with pure altruism. As in our model, she assumes that gifts are received early in the child's life while bequests are received later in life. In her model, gifts are progressive as measured *across* families, while bequests may be equal or even regressive depending on parameters. However, this result does not explain the puzzle in the data that concerns children *within* the same family. If one adds a second child to McGarry's model, bequests will provide full insurance within the family because of altruism. This feature is also present in the model analyzed by Nishiyama (2002). He builds a rich overlapping generations model with gifts and bequests to account for the observed U.S. wealth distribution, but he does not analyze the distribution of gifts and bequests among children within a family.

Lundholm and Ohlsson (2000) build a model in which the parent is altruistic with respect to her children's consumption, but her utility is decreasing in the difference in bequests. They argue that this structure could be a result of privately observed gifts and publicly observed bequests, with the difference in bequests representing reputational concerns of the parent. Their model delivers equal bequests but perfectly progressive gifts, which is not quite in line with the data, in which gifts are only weakly progressive.

Bernheim and Severinov (2003) build a model of bequests with more structure with respect to the costs of unequal division. In the model, parents love some children more than others, and children care about how much their parents love them relative to their siblings. Even parents who love their children unequally may divide bequests equally under some parameterizations in order to avoid signaling their preferences. It is not clear whether gifts, which Bernheim and Severinov do not model, can be successfully incorporated into their framework. The authors suggest that if gifts were incorporated in their model, they would be perfectly progressive, unlike in the data.

The present paper models classic altruistic parents and delivers the basic features of the puzzle as a result of a standard moral hazard problem. Parents care about their children's utility with equal weight on each child. *Tough love* refers to the fact that in our model with private information, parents would treat highly productive kids harshly if they pretended to be less productive.[2] This paper is not the first to consider the role of moral hazard in intergenerational transfers. Our predecessors include Kotlikoff and Razin (1988), Chami (1996), Cremer and Pestieau (2001), and Fernandes (2011). Their models are static models, however, and therefore cannot explain differences between gift and bequest behavior, as we do in this paper.

[2]For an alternative notion of *tough love*, see Bhatt and Ogaki (2012).

Cremer and Pestieau (1996, 1998) build dynamic models with moral hazard, but their models deliver the opposite of our results – equal gifts and progressive bequests. They consider parents giving gifts to fund schooling and to relax credit constraints before children start working. In this framework, gifts are optimally divided equally, because parents have no information about their children's earnings. Bequests are delivered later, when they are less intrinsically useful, but can be conditioned on earnings information. Thus, they are distributed progressively and in many cases only the low type receives a bequest. These papers note the challenge of interpreting their results in light of the prevailing norm of equal bequest division. Nishiyama and Smetters (2002) build a dynamic moral hazard model as well, but in their model there is only one parent and one child. As a result, they do not analyze the distribution of gifts and bequests among children within a family.

The forces in our model are in fact common to a wide range of dynamic insurance models. Written recursively, these models imply that in the face of income or productivity uncertainty, unlucky types receive a larger transfer today, balanced by a reduction in expected future welfare (see, e.g., Thomas and Worrall, 1990). In other words, current transfers are progressive and expected future welfare is regressive. In our model, expected future welfare is determined by bequests and future productivity. For low levels of productivity persistence, expected future productivity is similar across children, and bequests must be regressive to generate regressive expected future welfare. For high productivity persistence, expected future welfare is regressive even with progressive bequests (which are desirable as an insurance device).

4. Model

This section sets up a model of intergenerational transfers and focuses on the basic forces behind the main results. It shows that if labor productivities are private information of the children, bequests are equal and gifts are weakly progressive for a particular level of productivity persistence. These forces are robust to a number of extensions considered in this section as well as a version of the model used in the quantitative analysis in section 5.

This section considers the problem of a parent who distributes gifts and bequests that depend on her children's income to maximize the children's welfare.[3] This section characterizes a stylized model of the parent's choice with two types of children to highlight the basic forces at work in the parent's decision and cleanly derive analytical results. Most of these results hold in a richer version of the model, which is compared to the data in section 5.

4.1. **Children.** A parent has a unit mass of children so that no child's individual choice of output affects the total family resources. The parent lives for one period, the children live for two periods. Each child draws a productivity type $z_i \in \{z_L, z_H\}$ in each period and can produce output from labor effort linearly, $y = z\ell$. Productivities are i.i.d. across children and half of the children are of each type in both periods.

A child's utility is additively separable in first- and second-period consumption and labor effort, $c^1, \ell^1(= y^1/z^1), c^2, \ell^2(= y^2/z^2)$: $u(c^1) - v(y^1/z^1) + E\left[u(c^2) - v(y^2/z^2)\right]$, where $u' > 0$, $u'' < 0$, $u'(\infty) = 0$, $u'(0) = \infty$, $v' > 0$, $v'' > 0$, $v'(0) = 0$, $v'(\infty) = \infty$. This section assumes for simplicity that children do not save on their own. The model is extended to account for children's savings in section 4.5. As a result of having no savings, the child enters the second period with bequest b and one can define his continuation utility as a function of the bequest b and the first-period productivity z^1: $W(b, z^1) := E\left[u(c^2) - v(y^2/z^2)|z^1\right]$.

We now characterize W. At the beginning of period two, the child realizes his second-period productivity z^2 and solves an autarky problem given z^2 and bequest b:

$$\mathcal{W}(b, z^2) := \max_{c^2, y^2} u(c^2) - v(y^2/z^2) \qquad \text{s.t.} \qquad c^2 \leq y^2 + b.$$

[3]We adopt the convention of a female parent and male children for clarity.

The probability of being the same productivity type in both periods is $\pi \geq .5$, and the probability of switching types is $1-\pi$, and, therefore, $W(b, z_i) = \pi \mathcal{W}(b, z_i) + (1-\pi)\mathcal{W}(b, z_j)$. $W(b, z)$ is strictly increasing and strictly concave in b, and $W_b(b, z)$ is strictly decreasing in z for $\pi > \frac{1}{2}$, as shown in appendix D.1.[4]

4.2. **Parent's Problem.** The parent's moral hazard problem stems from private information. We assume that the parent observes her children's income, but she does not observe their labor productivity (wage rate) and labor effort (labor hours).[5] Parents in the real world might have information about their children's productivities and how many hours they work. However, most of the time this information comes from the children themselves, as it does in equilibrium in our model. In the model, parents design a schedule of gifts and bequests so that children do not choose to lie about their productivity. The private information assumption can be restated as saying that parents do not have additional information with respect to their children's productivity and labor hours.[6]

We proceed with the description of the timing in the model. At the beginning of the first period, the parent has assets, A, to distribute among her children. The parent designs a schedule of gifts and bequests as a function of her children's first period income, $g(y)$ and $b(y)$, to maximize the sum of their utilities. The parent is assumed to be able to commit to the

[4]Observe that the child does not derive utility from the bequest directly, but rather from the consumption and leisure implied by the bequest. Similarly, when we define the parent problem below, the parent cares about children's utility from consumption and leisure and cares only *indirectly* about the bequest. Thus, there is no *joy-of-giving* motive in our model.

[5]The available empirical evidence suggests that the private information friction might actually be even stronger than what we assume in this paper. We find in the Health and Retirement Survey (HRS) that on average one fourth to one half of parents are not able to place their children's income into three broadly defined income brackets (see appendix B for more detail). In a similar vein, Doepke and Tertilt (2015) provide empirical evidence on imperfect information among couples.

[6]The main forces analyzed in detail here are also present in a more general version of our model in which the parent receives a signal about her children's productivities. This model is analyzed in appendix E. The public and private information models analyzed below are two special cases of that general model.

transfer schedule. This means that the parent cannot renege on her promises when the child-types have been revealed. This is a common assumption in the dynamic insurance literature. The parent dies at the end of the first period without observing second period productivities and allocations. Given the transfers, children decide how much output to produce in the first period. They then consume their first period output plus the gift they receive and carry the bequest into the next period, in which they are on their own and solve the second period problem defined above. Formally, the parent chooses functions $g, b : \Re_+ \longrightarrow \Re$ to solve:

$$\max_{g,b} \sum_{i\in\{H,L\}} \frac{1}{2}\left[u(g(y_i)+y_i) - v(y_i/z_i) + W(b(y_i), z_i)\right] \qquad \text{s.t.}$$

$$\frac{1}{2}\left[g(y_H)+b(y_H)\right] + \frac{1}{2}\left[g(y_L)+b(y_L)\right] \leq A,$$

$$\forall i : y_i \in \text{argmax}_y u(g(y)+y) - v(y/z_i) + W(b(y), z_i).$$

By a version of the revelation principle,[7] one can equivalently think of a parent picking consumption, output, and bequests directly, as long as her children are willing to truthfully report their types under this allocation. We focus on this equivalent problem because it can be characterized more directly. For two possible child-types, the parent's problem is:[8]

$$\max_{(c_i,y_i,b_i)_{i\in\{H,L\}}} \sum_{i\in\{H,L\}} \frac{1}{2}\left[u(c_i) - v(y_i/z_i) + W(b_i, z_i)\right] \qquad \text{s.t.}$$

$$\sum_{i\in\{H,L\}} \frac{1}{2}(c_i + b_i) \leq \sum_{i\in\{H,L\}} \frac{1}{2} y_i + A, \tag{4.1}$$

$$\forall i,j \quad u(c_i) - v(y_i/z_i) + W(b_i, z_i) \geq u(c_j) - v(y_j/z_i) + W(b_j, z_i). \tag{4.2}$$

[7]The proof is standard and therefore omitted. See Fudenberg and Tirole (1991, p. 253).
[8]For simplicity, we omit the time superscripts: c_i, y_i, z_i correspond to first-period allocations and productivity. Second period allocations are implied by the bequest and the shock realization as discussed above.

This model is similar, but not equivalent to the standard two-period Mirrlees model. In our model (unlike in the Mirrlees model) the parent does not directly choose children's consumption and labor effort/income in the second period. The parent only chooses bequests. As a result, the parent cannot distort children's second period consumption-labor choices.

4.3. **Public Information.** Before characterizing the solution to the parent's problem above, it is helpful to consider the solution to this problem with public information (i.e., without the incentive constraints (4.2)) as a benchmark. In this version of the model, the parent observes first period productivities when handing out gifts and bequests. The parent still does not observe the second period productivities, because she dies before they are realized. It is straightforward to show that the solution to the public information problem has the following properties: (i) $c_H = c_L$, (ii) $y_H > y_L$, (iii) $g_L > g_H$ since $g_i := c_i - y_i$, (iv) $b_L > b_H$ if $\pi > .5$.

The first three results highlight the fact that with altruistic preferences, gifts completely offset differences in income (full insurance). The last result follows from the fact that the parent provides full insurance with respect to bequests as well by equalizing the marginal utilities of bequests (i.e., $W_b(b_L, z_L) = W_b(b_H, z_H)$). Since second-period low-productivity types benefit more from bequests, productivity persistence implies more bequests to currently low-productive children. Therefore, the public information model is not able to qualitatively account for the equal bequests and weakly progressive gifts observed in the data.

4.4. **Private Information.** Only the incentive constraint preventing the high type from pretending to be the low type binds at the solution to the parent's problem.[9] At the optimum, the parent is doing just enough to keep the productive kid from slacking off. The first order conditions yield $u'(c_H) = \frac{\lambda}{1+\mu} < \frac{\lambda}{1-\mu} = u'(c_L)$, hence $c_H > c_L$, where λ is the Lagrange multiplier on the budget constraint (4.1) and $\frac{\mu}{2}$ is the Lagrange multiplier on the high type's

[9]This is a standard result. See appendix D.5 for the proof.

incentive constraint. μ represents the intensity with which the incentive constraint binds, a measure of the need to provide incentives. This imperative prevents full insurance with gifts in our model for any level of productivity persistence.

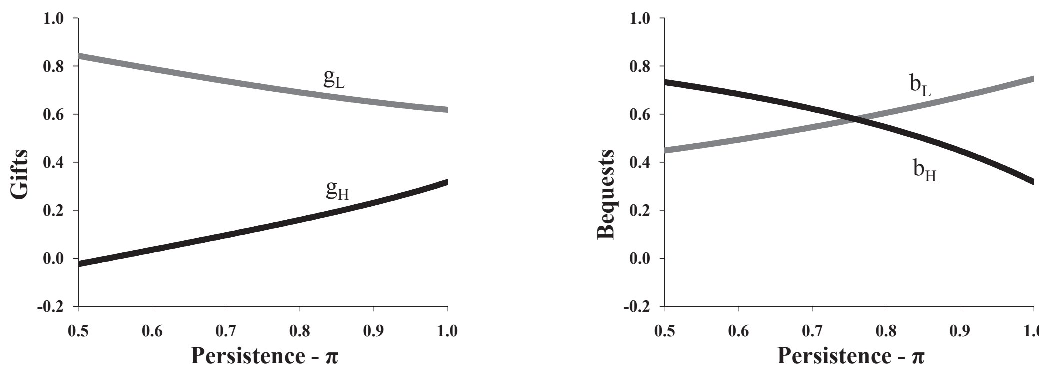
The next proposition presents the main result of this section. It shows that there is a level of productivity persistence for which the pattern of gifts and bequests in the model is qualitatively in line with the data. Gifts are progressive (but weakly, since $c_L < c_H$ as shown above) and bequests are equal.

Proposition 4.1. *Suppose u has nonincreasing absolute risk aversion, v has constant absolute risk aversion or constant relative risk aversion, and z_L is in the neighborhood of zero. Then there exists a level of persistence $\pi^* \in (.5, 1)$ under which $g_L > g_H$ and $b_L = b_H$.*

Discussion of proposition 4.1. In our model, gifts are received early in children's life, whereas bequests are received later in life. Children face significant uncertainty about their future productivity, which the parent can take advantage of. Productive children will be relatively more concerned about the size of their bequest, because they may be less productive in the future. This reasoning is reflected in our finding that $g_H < b_H$ for any level of persistence $\pi \in (.5, 1)$ (see the proof of lemma D.4 in appendix D.3 and figure 1 for a numerical example illustrating this result). Less productive children will be more concerned about the size of their gift, since they may become more productive in the future. Therefore, the parent can make the best use of her assets by giving relatively more to her poor children through gifts and relatively more to her rich children through bequests.

Proposition 4.1 highlights the importance of the uncertainty of the child's future productivity in our model. When there is no productivity persistence, $\pi = .5$, each type has the same chance of being a high or low type in the following period, and therefore both types feel the same way about bequests for tomorrow. We saw earlier that, since both types get the same utility from consumption, consumption is unevenly distributed. In a similar vein,

Figure 1. Persistence and the Distribution and Gifts and Bequests



with $\pi = .5$, both children get the same expected utility from bequests, so the parents use bequests to help provide incentives. They give more to the richer child but ask him to produce more in the first period.

When persistence is perfect, $\pi = 1$, both types know that they will be the same type next period as they are today. In this case, both types care as much about tomorrow as today. Incentives are not concentrated in either period, and therefore both gifts and bequests are progressive. For intermediate levels of persistence, both effects are at work and bequests are about equal, and for some level they are exactly equal.

We prove this result for u with nonincreasing absolute risk aversion (NIARA), v with constant absolute risk aversion (CARA) or constant relative risk aversion (CRRA), and z_L in the neighborhood of zero. Numerically, it appears to be true for a much broader class of utility functions, including nonseparable ones and any $z_L < z_H$. In the rest of this section, we state and provide intuition for a sequence of lemmas in service of the proposition above.

Figure 1 is a typical graph of each type's gifts and bequests, as a function of π for a particular set of parameters ($u(c) = c^{1-\sigma}/(1-\sigma)$ with $\sigma = 2$ and $v(l) = \ell^{1+\gamma}/(1+\gamma)$ with $\gamma = 2$, $z_H = 2$, $z_L = 1$, $A = 1$). These figures help illustrate the results proved in the lemmas below. Bequests to the high type are smaller than for the low type at $\pi = 1$, and larger at $\pi = .5$, but never enough to offset the gap in gifts, thus, total transfers are larger for the low type. In between the extremes the policy functions change smoothly, and for an intermediate value of π bequests are equal.

Lemma 4.2. *Total transfers are progressive:* $g_L + b_L > g_H + b_H$.

Proof: See appendix D.2. This result is a reflection of the parent's insurance motive. Incentive problems constrain but do not reverse this motive. □

Lemma 4.3. *Under no persistence,* $\pi = .5$*, bequests are regressive,* $b_L < b_H$.

Proof: No persistence implies that the expected value of bequests is independent of today's type, $W(b, z_L) = W(b, z_H) = W(b)$. Thus, one can use the same arguments as for consumption. The first order conditions yield $W_b(b_H) = \frac{\lambda}{1+\mu} < \frac{\lambda}{1-\mu} = W_b(b_L)$, which implies $b_H > b_L$. □

The right panel of figure 1 illustrates this result in our numerical example. A related result holds in benchmark dynamic insurance problems; see, e.g., Thomas and Worrall (1990). In their model, unlucky types receive a larger transfer today, balanced by a reduction in future welfare. These forces yield similar results in our model. If there is no income persistence in our model, less productive children receive more in gifts, $g_L > g_H$ and future welfare is determined by bequests only. Therefore, a reduction in future welfare corresponds to a smaller bequest, $b_L < b_H$.

Lemma 4.4. *Under perfect persistence,* $\pi = 1$*, and for* z_L *in the neighborhood of zero, bequests are progressive,* $b_L > b_H$.

Proof: Suppose the low type is disabled (i.e., $z_L = 0$). Then he consumes only what he is given, so that $c_L^1 = g_L$, $c_L^2 = b_L$. Since $\pi = 1$, $W(b, z_i) = \mathcal{W}(b, z_i)$, and $W_b(b_L, z_H) < W_b(b_L, z_L)$ as established in appendix D.1, we have the following result where superscripts refer to periods:

$$u'(c_L^1) = \frac{\lambda}{1-\mu} > \frac{\lambda}{1-\mu\frac{W_b(b_L,z_H)}{W_b(b_L,z_L)}} = W_b(b_L, z_L) = u'(c_L^2).$$

So $c_L^1 < c_L^2$, which implies $g_L < b_L$. The assumption that $z_L = 0$ implies that $c_L^1 = g_L$ and $c_L^2 = b_L$ which make it possible to characterize g_L and b_L using the first order conditions on consumption. If $z_L > 0$, one can show that $y_L^1 < y_L^2$ (see the proof of lemma D.7), making $c_L^1 - y_L^1 = g_L$ difficult to compare to $c_L^2 - y_L^2 = b_L$ (since $c_L^1 < c_L^2$).

The high type's first order conditions are undistorted, he smooths consumption and output, $c_H^1 = c_H^2$, $y_H^1 = y_H^2$, so $g_H = b_H$. Lemma 4.2 implies $2b_H = g_H + b_H < g_L + b_L < 2b_L$, which proves the lemma. Continuity of policies in z_L (lemma 4.5) guarantees that this will be true in the neighborhood of $z_L = 0$. □

In a large number of numerical simulations, we always find this neighborhood to be $[0, z_H)$. At perfect persistence, whenever $z_L < z_H$, the low type receives a larger bequest than the high type in all simulations we have run, encompassing a variety of utility functions (including nonseparable ones) and parameter values. Figure 1 illustrates that, with perfect persistence, $b_L > b_H$ in the numerical example in which $z_H = 2$ and $z_L = 1$. The forces driving this result are also present in simpler models. Imagine a version of the model in which the parent is alive in both periods. With perfect persistence, the two periods are symmetric and the static solution to the problem applies: transfers are larger to the low productivity type in both periods. These considerations suggest that the lemma is satisfied more generally.

Lemma 4.5. *Policy functions are continuous in π and z_L if u has NIARA and v has CARA or CRRA.*

Proof: See appendix D.4. The maximum theorem guarantees upper hemi-continuity. We use the assumptions on the utility functions to show that policies are single valued everywhere. One can relax the assumptions on v, and simply assume that it has nondecreasing relative risk aversion (NDRRA) if one also assumes that W has NIARA. The stronger assumptions on v are sufficient to show that W has NIARA. □

Proof of proposition 4.1. Combining the above results implies that $b_H - b_L > 0$ for $\pi = .5$ and $b_H - b_L < 0$ for $\pi = 1$. Since policy functions are continuous in π, there exists a π^* such that $b_H - b_L = 0$ by the intermediate value theorem. Lemma 4.2 then implies that gifts are progressive. This finishes the proof of the proposition. □

We have shown that our model can qualitatively explain a pattern of transfers in which gifts partially offset income differences but bequests are divided evenly, seemingly providing no insurance. This model delivers the stylized facts of the equal division puzzle with a standard altruistic parent facing a common moral hazard problem, one that recognizes concerns about providing the wrong incentives that real parents face.[10] The timing of the two types of transfers in this relatively standard parent's problem generates significant differences in the way gifts and bequests are distributed.

4.5. **Extensions.** The simple two-period, two-type, many-child model of the previous section highlights the basic forces driving apart gifting and bequesting behavior with a minimum of distractions. This section serves two purposes. First, it shows that the main results are robust to several extensions, which add realism to our model. Second, it presents the version of the model, which will be evaluated quantitatively in section 5.

Two-Child Model. To begin, we extend the model with the features which will be used in the quantitative evaluation in section 5. We extend the model to consider two-child families,

[10] In appendix E we extend this discussion to a version of the model in which parents receive a signal about their children's productivity. We find that the conclusions of this section are robust to adding signals.

which vary by parents' assets, parents' productivity, and children's productivity. This section considers the qualitative implications of these changes.

The parent has two children, indexed by 1 and 2. The state of the family (ij) defines productivity levels for the first child, z_i, and the second child, z_j. The parent also cares about her own utility from consumption and labor effort. The parent's productivity level is z_p, which is public information, her allocations are indexed by p, and she discounts the welfare of her children by η. Both children are weighted equally. They discount the future by β and the interest rate is R. A child's productivity is known neither by the parent nor by the other child. A version of the revelation principle can be proved for this environment as well. Therefore, we will consider a problem in which the parent assigns allocations as functions of the children's type. The full parent's problem is

$$V(A,z_p) = \max E\Big\{u(c_p(ij)) - v(\frac{y_p(ij)}{z_p}) + \eta\big[u(c_1(ij)) - v(\frac{y_1(ij)}{z_i}) + \beta W(b_1(ij),z_i) + u(c_2(ij)) - v(\frac{y_2(ij)}{z_j}) + \beta W(b_2(ij),z_j)\big]\Big|z_p\Big\} \quad \text{s.t.}$$

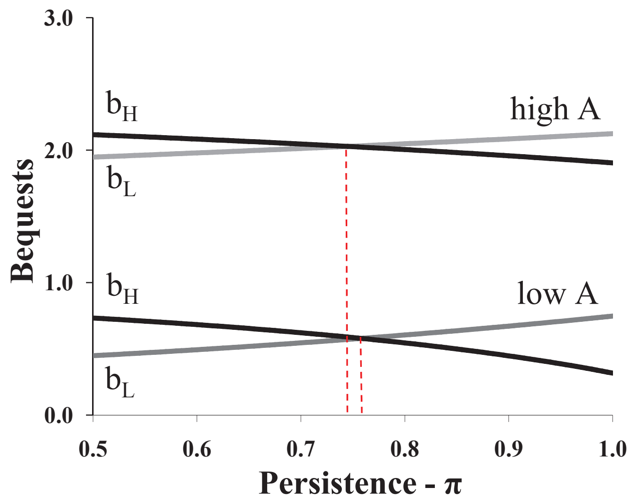
$$\forall i,j : c_1(ij) + c_2(ij) + c_p(ij) + \frac{b_1(ij)}{R} + \frac{b_2(ij)}{R} \leq y_1(ij) + y_2(ij) + y_p(ij) + A,$$

$$\forall i,k : E\big\{u(c_1(ij)) - v(\frac{y_1(ij)}{z_i}) + \beta W(b_1(ij),z_i)\big|z_p,z_i\big\} \geq E\big\{u(c_1(kj)) - v(\frac{y_1(kj)}{z_i}) + \beta W(b_1(kj),z_i)\big|z_p,z_i\big\},$$

$$\forall j,k : E\big\{u(c_2(ij)) - v(\frac{y_2(ij)}{z_j}) + \beta W(b_2(ij),z_j)\big|z_p,z_j\big\} \geq E\big\{u(c_2(ik)) - v(\frac{y_2(ik)}{z_j}) + \beta W(b_2(ik),z_j)\big|z_p,z_j\big\}.$$

Family-Level Heterogeneity. In the simple version of the model, parents vary by their asset level A. This version of the model adds two other sources of variation in family-wide resources. Parents have their own productivity z_p, and there are only two children in the family. These additions introduce heterogeneity across families in family-wide resources and uncertainty in children's incentive constraints. One can show that a version of proposition 4.1 still holds with two child-types. Parents with two high-type children or two low-type

Figure 2. Variation in Parent's Asset Level A and Equal Bequests



children will necessarily divide bequests evenly between these "identical twins," and there is a π such that parents with one child of each type divide bequests evenly.

With family heterogeneity this π may be family-specific. In figure 2, we simulate parent's bequests as in figure 1 for two different wealth levels. Richer parents are more generous, the children work less, and the incentive constraint does not bind as tightly. Thus, the π at which the insurance and incentive motives offset each other and bequests are equal is lower. In the quantitative section, we consider a large number of family types; hence, there will be no common level of persistence that sets bequests exactly equal within all families. There is, however, a range of persistence values for which most bequests will be close to equal.

More Than Two Productivity Types. Extending the model to more than two productivity types makes computations more difficult but does not change our basic quantitative findings. More child-types means more types of families, which contributes to the issue discussed above. It also implies that the number of incentive constraints increases to $n(n-1)$ for n

types. Numerically, we find that only the $n-1$ local downward constraints bind. It can be proven that this is the pattern of binding constraints in simpler insurance models (Thomas and Worrall, 1990), but we have not proven this result in our environment, so we cannot analytically establish the other results from our simple model.

Observed Child Savings. For simplicity, savings are excluded from the model presented in this section. Savings are, however, an important feature in the data, so we do account for them quantitatively in section 5. We assume that savings yield the same return R as bequests. With respect to timing, we assume that after production, children set aside savings s. The parent then hands out gifts and bequests as a function of both income and savings (i.e., $\hat{g}(y,s)$ and $\hat{b}(y,s)$), rather than just income (i.e., $g(y), b(y)$), as in the benchmark model.

Bequests and savings are perfect substitutes in this model. If children save, the parent can achieve identical consumption-labor allocations, as in the model without savings. The parent simply offsets the savings of the children with the appropriate transfer scheme: $\hat{g}(y,s) = g(y) + \frac{s}{R}, \hat{b}(y,s) = b(y) - s$. This transfer scheme leaves the children indifferent between saving and waiting for the bequest. In other words, from the children's perspective the optimal level of savings is indeterminate. To pin down the level of child savings in the quantitative exercises of section 5, we assume a savings level consistent with the data.

Unobserved Child Savings. In our model, parents observe the income and consumption of their children; thus, any child who tried to save secretly would be caught by a parent who calculates savings s as the difference between income y and consumption c: $s = y - c$. For savings to be unobserved, either consumption or income must be unobserved. If parents cannot observe their children's consumption and only local downward constraints bind (as in all versions of our model that we solve numerically), then the child's Euler equation is $u'(c_t) \geq W_b(b,z) = \beta RE\{u'(c_{t+1})\}$ with a strict inequality for all but the highest type. This means that children would like to deviate from the allocation offered by the parent by borrowing in the first period. If children were borrowing constrained, they would not deviate. Hence, the optimal allocation of our model is also optimal in a model with unobserved savings, unobserved consumption and borrowing constraints. Unobserved income has just the opposite effect on savings. The intertemporal first order condition in terms of labor is $v'(y_t/z_t) \leq W_b(b,z) = \beta R E\{v'(y_{t+1}/z_{t+1})\}$. At the solution to our model, children would prefer to work more in the first period and save. As a consequence, our model is not robust to making income unobserved in the presence of unobserved saving.

5. Quantitative Analysis

Section 4 shows that the opportunity to provide incentives dynamically (rather than just statically) offsets parents' desire to compensate low-income children through bequests. Parents distribute bequests equally across children in a model with two productivity types and a particular level of productivity persistence. This section pins down the productivity persistence with U.S. data and quantitatively analyzes the extended version of the model described in detail in section 4.5. The quantitative analysis highlights the role of incentives by comparing the private information model with its public information counterpart. To our knowledge, this is the first paper to evaluate a model addressing the equal division puzzle quantitatively.

5.1. **Parameterizing the Model.**

To compare the performance of the model to the data, as many statistics as possible are recovered directly from the Panel Study of Income Dynamics (PSID). The details of the data work with PSID are contained in appendix A. The statistics that are not available in PSID are collected from other sources in the literature.

Model Timing. To map the two-period model to the data, we begin by determining ages representing periods one and two, which are distinguished in the model by the receipt of a bequest. In the PSID, the average age of the household head when the household receives

Table 1. Age of Household Head When Receiving a Bequest

Age	<25	25-30	30-35	35-40	40-45	45-50	50-55	55-60	60-65	65-70	70-75	>75
Number	59	99	156	141	172	214	173	156	167	113	68	105
Fraction	3.6%	6.1%	9.6%	8.7%	10.6%	13.2%	10.7%	9.6%	10.3%	7.0%	4.2%	6.5%
Cum. Fraction	3.6%	9.7%	19.3%	28.0%	38.6%	51.8%	62.5%	72.1%	82.4%	89.3%	93.5%	100.0%

This table reports the number and percentage of people in the PSID who receive a bequest in a given age bracket between 1988 and 2013.

a bequest has been rising from 47 in 1988 (from this year onwards bequests are consistently reported) to 53 in 2013, with 50 being approximately the average over the years. The average median age over the years is also close to 50.[11]

We also look more closely at the distribution of household head's ages when the household receives a bequest (for the combined period 1988 - 2013). These results are summarized in table 1 and indicate that age 50 is a good approximation of the average age of the household head when the household receives a bequest. Therefore, we assume that period 1 in the model corresponds to children's ages 25-50 (at age 25 the majority of people have completed their education) and period 2 corresponds to ages 50-75 (according to the World Bank Open Data database, 75 years is a good approximation of the average age at which people in the United States die over the PSID sample period). Consistently, parents in period 1 of the model are interpreted as people between ages 50-75. These assumptions will be important for the estimation of the productivity process.

Utility Function Parameters. The utility functions take the following forms: $u(c) = \frac{c^{1-\sigma}}{1-\sigma}$ and $v(\ell) = \phi\frac{\ell^{1+\gamma}}{1+\gamma}$ for both parents and children. Both functions are CRRA and satisfy the assumptions of the proofs in the appendix. In the benchmark model $\sigma = \gamma = 1$.[12] Labor effort in the model is interpreted as labor hours in the data. Total hours are normalized to

[11]PSID reports only the receipt of a bequest by a household and the size of the bequest. We report the age of the household head at the time of any bequest receipt. Restricting the sample to first bequest receipts only or using the ages of wives would reduce this number by 3 to 5 years.

[12]Both values imply a constant relative risk aversion of 1, neutral in the sense of unit wage and price elasticities. $\sigma = 1$ also implies that the preferences are consistent with balanced growth and falls inside the range used in finance and macro literatures. $\gamma = 1$ implies a Frisch elasticity of labor supply also within the range considered in the literature.

TABLE 2. Internally Calibrated Parameters

Parameter	Symbol	Value	Target	Data & Model	Source
Disutility of labor	ϕ	6.80	Labor supply	1/3	
Parents' altruism	η	0.99	Average gift	0.65%	PSID
Discount factor	β	0.95^{25}	Average bequest	2.22%	PSID

This table reports the benchmark calibration procedure. Average gift and average bequest in the data are defined relative to total household labor earnings.

1 and the labor disutility parameter ϕ is calibrated so that labor supply in the first period is one-third on average. The annual interest rate is set to 4% and the discount rate β and the altruism parameter η are calibrated so that the model replicates the average size of bequests and gifts in the data. The average size of bequests between 1988 and 2013 is 2.22% relative to average household labor earnings. The average size of gifts between 1985 - 2013 is 0.65% relative to average labor earnings. The calibration procedure is summarized in Table 2. The rest of the parameters are set outside the model.

Productivity Process. Labor productivity in the model z_1 is interpreted as hourly wages in the data averaged over ages 25-50 and labor productivities z_2 and z_p as average hourly wages between ages 50-75. The joint distribution of parents' initial assets A and wages/productivities z_P is assumed to follow:

$$(5.1) \quad (\log z_P, \log A) \sim \mathcal{N}_2\left(\begin{pmatrix}\mu_p \\ \mu_A\end{pmatrix}, \begin{pmatrix}\sigma_p^2 & corr(A, z_p)\cdot\sigma_p\sigma_A \\ corr(A, z_p)\cdot\sigma_p\sigma_A & \sigma_A^2\end{pmatrix}\right).$$

Children's wages/productivities in the first period, z_1, and in the second period, z_2, evolve according to:

$$(5.2) \quad \log z_1 = \mu_1 + \rho_1(\log z_p - \mu_p) + \varepsilon_1, \qquad \varepsilon_1 \sim \mathcal{N}(0, \sigma_1^2),$$

$$(5.3) \quad \log z_2 = \mu_2 + \rho_2(\log z_1 - \mu_1) + \varepsilon_2, \qquad \varepsilon_2 \sim \mathcal{N}(0, \sigma_2^2).$$

To completely define the asset distribution and wage/productivity processes in equations (5.1) – (5.3), eleven parameters need to be specified: μ_A, μ_p, σ_A^2, σ_p^2, $corr(A, z_p)$, μ_1, ρ_1, σ_1^2, μ_2, ρ_2, σ_2^2. These parameters are estimated using the PSID.

The PSID's principle virtue is its length – it has been tracking participants since 1968 – a critical feature for estimating the parameters in equations (5.1) – (5.3) directly.[13] The long run persistence of income over one's lifetime, ρ_2, defined in equation (5.3) is central to our model. A direct estimation of equation (5.3) would mean calculating average wages from the age of 50 to 75 (z_2 in equation (5.3)) and regressing them on average wages from 25 to 50 (z_1 in equation (5.3)). This would require a 50-year panel, still beyond the current reach of the PSID (38 waves are available over a 45-year period from 1968 to 2013). Even with four additional years of data (and the years missing since 1997 when PSID became biennial), a simple estimate would rely exclusively on the cohort that was 25 years old in the 1968 sample and continued reporting through 2017, a small sample. To overcome these issues, equation (5.3) is estimated for less demanding time spans as follows.

We first interpolate wages for the missing years between 1997 and 2013 (taking averages of the adjacent years). Then, PSID participants for whom there are k consecutive wage observations between ages 25 and 50 and k consecutive wage observations exactly 25 years later are selected. Their average real wages z_1 and z_2 are then calculated over these periods of length k years. Finally, regression equation (5.3) is estimated using these measures of z_1 and z_2.[14]

The estimate of ρ_2 varies with k. For small k's, the number of observations is large, but the estimates of average wages when young and when old are imprecise. When k is large, the estimates of average wages are more precise, but the number of available observations

[13] The PSID does not, however, provide transfer data in which the source parent and recipient child can be uniquely identified. Thus the model outcomes must be compared to other results from the literature.
[14] If a person is included in the estimation multiple times, then these observations are weighted by the inverse of the number of these observations.

TABLE 3. Benchmark Parameters Estimated From PSID

Parameter	Symbol	Value
Average asset level	$exp(\mu_A)$	$75,000
Standard deviation of logged asset holdings	σ_A	1.6
Asset-productivity correlation	$corr(A, z_p)$	0.65
Average parent's hourly wage	$exp(\mu_p)$	$7.54
Standard deviation of parents' wages	σ_p	0.78
Intergenerational wage persistence	ρ_1	0.25
Intergenerational wage volatility	σ_1	0.47
Average child's hourly wage when young	$exp(\mu_1)$	$10.92
Average child's hourly wage when old	$exp(\mu_2)$	$9.18
Intertemporal wage persistence	ρ_2	0.78
Intertemporal wage volatility	σ_2	0.88

declines. In the benchmark, an intermediate $k = 12$ is used. This is the maximum k for which there are at least 500 people in the PSID for whom one can calculate the average wages z_1 and z_2 using the procedure described above. For this value of k, $\rho_2 = 0.78$. We use this number as a benchmark and analyze the sensitivity of our results to this number. We also recover the estimates of $\mu_1 = 2.39$, $\mu_2 = 2.22$ and $\sigma_2 = 0.88$.[15]

The details of the estimations of the remaining parameters of the asset distribution and the wage/productivity process are presented in appendix A. Table 3 summarizes these parameters which will be used as inputs to the model.[16]

Child Savings. Recall from section 4.5 that bequests and savings are perfect substitutes from the perspective of the parents; both transfer money from period 1 to period 2 with the

[15]The labor literature typically estimates versions of equation (5.3) using residuals from a Mincerian wage regression in which wages are first regressed on observables (see, e.g., Heathcote, Storesletten, and Violante (2010)). We chose to work with wages directly, because this is what appears in the model. Nevertheless, we also check the sensitivity of our estimates to including standard observables (year, age, education) in equation (5.3). We find that ρ_2 decreases by approximately 0.1 (again, this number varies with k). The decline is intuitive, since part of the persistence in wages is a result of observable characteristics. The estimates of σ_2^2 decrease by about 0.05, which is also intuitive. As shown in the quantitative analysis, using a smaller value for the persistence parameter ρ_2 brings the model even closer to matching the data. On the other hand, changes in the volatility parameter σ_2 do not affect the main results.

[16]To solve the model numerically, we discretize the continuous wealth-wage distributions specified in equations (5.1) - (5.3). In the benchmark simulations, we use 5 states for assets and 5 skill types for the parents, thus having a grid of 25 points. The distance on either axis is the same measured in probability of the unconditional distributions. For each z_P we have a grid of 5 values of z_1 and 5 values of z_2 with equal distance in conditional probabilities $z_1|z_P$ and $z_2|z_P$. Thus, we have 25 grid points for both z_1 and z_2.

same interest rate. The bequests that will be compared to the data are defined as the simple model (no savings) bequests b minus savings s pinned down by the data $\hat{b} = b - s$. Gifts that will be compared to the data are consequently $\hat{g} = g + \frac{s}{R}$. In our specification, the parent simply takes children's savings as a function of children's income as given, and adjusts her transfer schedule accordingly.

Children's savings s as a function of income y are approximated by estimating the following regression for people at the age of 50, using all the waves of PSID that report household wealth (i.e., the eight waves between 1999 and 2013): $\log s_i = \alpha + \beta \log y_i + \varepsilon_i$. In this regression equation, savings s_i is household wealth at the household head's 50th year of age (which is the result of past savings accumulated until the age of 50) and y_i are total household labor earnings between ages 25-50 as defined in appendix A. Total household earnings y_i thus correspond to the sum of 25 years of labor earnings. As expected, savings/wealth is increasing in earnings: $\log s_i = -9.86 + 1.55 \log y_i$. Given these savings/wealth, the parent then hands out an adjusted bequest and gift $\hat{b}(y, s(y)) = b(y) - s(y)$ and $\hat{g}(y, s(y)) = g(y) + \frac{s(y)}{R}$, respectively (recall that in the original model, bequests and gifts are functions of income y only).

5.2. **Benchmark Quantitative Results.** This section evaluates the performance of the quantitative model by constructing several uncalibrated statistics for gifts and bequests and comparing them with the data. Table 4 reports the statistics of interest under the benchmark parameterization. To highlight the role of dynamic incentives, the table also includes statistics from a model with public information (i.e. without incentive constraints) but otherwise identical in structure. The utility parameters ϕ, η and the discount rate β are recalibrated so that the public information model matches the same targets as in the private information model: average labor supply equals 1/3, average bequest equals 2.22% and average gift equals 0.65% relative to average labor earnings.

Table 4. Model Statistics for Benchmark Parameters

	Data	Private info	Public info
β_{by}	0	-0.0367	-0.1419
β_{gy}	-0.02	-0.4537	-0.9438
Bequests within 2% of mean	77%	20.00%	20.00%
Bequests within 25% of mean	88%	55.87%	22.74%
Equal gifts	9%	20.00%	20.00%
Gifts within 20% of mean	11%	24.89%	20.85%
κ_b		0.09%	1.94%

Progressivity of Gifts and Bequests. To compare the progressivity of gifts and bequests we run the following regressions on the within-family model-data:

$$\hat{b}^1 - \hat{b}^2 = \beta_{by}(y^1 - y^2), \qquad \hat{g}^1 - \hat{g}^2 = \beta_{gy}(y^1 - y^2).$$

Here, $\hat{b}^i, \hat{g}^i$, and y^i are the bequests, gifts, and income of child i, and therefore, $\hat{b}^1 - \hat{b}^2, \hat{g}^1 - \hat{g}^2, y^1 - y^2$ are within-family differences. Using HRS data, Hochguertel and Ohlsson (2009) estimate β_{gy} to be equal to -0.02 and significantly different from 0. As for bequests, Wilhelm (1996) runs the same regression using only families in which bequests are divided unequally (which is a small fraction of all families), and he finds that even with this sample, the coefficient is not significantly different from 0. In the other families, bequests are equal across children even if income is not, implying that $\beta_{by} = 0$ in this subsample.

In our benchmark private information model, differences in bequests are nearly independent of the differences in income, as in the data. The point estimate of β_{by} in our benchmark model implies slightly *progressive* bequests: a \$1.00 difference in income is compensated by a \$0.0367 difference in bequests. Bequests are much more progressive with public information. Gifts are more progressive than bequests in the benchmark private information model: an extra \$1.00 difference in income between children is compensated by \$0.4537 in gifts. Gifts are more progressive than in the data, but much less than with public information. As for the progressivity of transfers, the private information model thus fits the data better than the public information model.

These regression coefficients would be closer to zero if we added measurement error in income when calculating the model implied regression coefficients. Measurement error is present in the data; see, for instance, the discussion in Wilhelm (1996), on whose estimates we rely for the bequest coefficient. In this sense, our results can be regarded as conservative, since adding measurement error would bring the model implied coefficients closer to the data.

Dispersion of Gifts and Bequests. Recall that the model is calibrated so that the average bequest and gift relative to income match the data. This section discusses how the model performs with respect to bequest and gift dispersion which was not a calibration target. In the model, bequests will be exactly equal only in families in which both children are of exactly the same productivity type. We therefore also look at a broader measure of dispersion in bequests: the percentage of children who receive bequests within 2% (almost equal) and within 25% (approximately equal) of the mean bequest for their family. Wilhelm (1996) reports these numbers to be 77% for the 2% threshold and 88% for the 25% threshold. We do not have the exact counterpart of these measures for gifts in the data, but Hochguertel and Ohlsson (2009) report that conditional on giving a gift, 9% of parents give exactly equal gifts and 11% gifts that are within 20% of the intra-family mean.

As table 4 shows, the private information model accounts, to some extent, for the small within-family dispersion in bequests that we observe in the data. The public information model, in contrast, generates a much larger within-family dispersion of bequests than in the data. Gifts are more dispersed in the data than in both the private and the public information model (note that children are identical and are given the same allocation trivially in 20% of the families). In fact, the public information model is slightly closer to matching the data in this statistic.

Welfare Loss of Equal Division. To assess whether equal bequests can be explained as a simple rule of thumb, we compute the welfare loss associated with equal division. For a given parent with asset level A and productivity level z_P, we calculate $V(A, z_P)$, the value at the solution to her problem defined in section 4.5. This parent does not face any restrictions with respect to how she divides her gifts and bequests. We then compute $V_=(A, z_P)$, the value at the solution to the same problem with the added constraint $b_1(ij) = b_2(ij)$ in all states ij. This is the utility of a parent who must divide bequests equally among her two children. We then measure the welfare loss of equal division as the difference $V(A, z_P) - V_=(A, z_P)$ measured as a fraction of parents' consumption (holding other allocations fixed) and denote it as $\kappa_b(A, z_P)$. This number expresses how much consumption a parent who is forced to divide bequests equally is willing to sacrifice in order to be allowed to divide bequests unequally. It also measures the minimal costs of unequal division that would be needed to make the parent divide her bequests equally.

The average welfare loss of equal bequest division is equivalent to the loss of 0.09% of parents' consumption in the benchmark private information model.[17] This number is small and suggests that parents would be almost as happy to divide bequests equally. The welfare loss is much higher under public information, equivalent to 1.94% of the parents' consumption. This value is more than 20 times higher than in the private information model. These results show that a norm of equal division is nearly optimal in our benchmark model and that this result is generated by the dynamic incentive problem of the parent.[18]

[17]We also investigate the welfare loss if gifts, not bequests, were constrained to be equal across children in the private information model. The welfare loss in this case is significantly larger in terms of parents' income, at 5.47%. Parents want to front-load the insurance aspect of their transfers; thus, it is much more costly to constrain this type of transfer to be equal.

[18]Alternatively, one can think of equal division arising because of the costs of unequal division. This idea has been widely considered in the literature. For instance, Wilhelm (1996) argues that parents may be concerned that unequal division would cause strife in the family. Unequal bequests are more easily challenged, implying legal costs as well. The cost of unequal division that would be necessary to justify equal division is equivalent to the welfare loss of equal division discussed above. Our private information model thus provides a much

TABLE 5. Role of Wage Persistence ρ_2 in the Private Information Model

	Data	Benchmark		
		$\rho_2 = 0.9 \cdot 0.78$	$\rho_2 = 0.78$	$\rho_2 = 1.1 \cdot 0.78$
β_{by}	0	-0.0240	-0.0367	-0.0497
β_{gy}	-0.02	-0.4819	-0.4537	-0.4593
Bequests within 2% of mean	77%	25.94%	20.00%	20.00%
Bequests within 25% of mean	88%	66.30%	55.87%	44.21%
Equal gifts	9%	20.00%	20.00%	20.00%
Gifts within 20% of mean	11%	24.05%	24.89%	24.89%

Role of Wage Persistence. The life time wage persistence ρ_2 is important for the relative size of gifts and bequests, as shown theoretically in section 4.4. Table 5 documents how the pattern of intergenerational transfers in the benchmark private information model changes if the value of the wage persistence parameter ρ_2 is lower and higher by 10%, respectively. As discussed in section 4.4, lower persistence implies less progressive bequests (i.e., a larger β_{by}). Lower persistence also implies less dispersed bequests. Importantly, for a large set of values of ρ_2, β_{by} is relatively close to 0 (much closer to 0 than β_{gy}). The other model statistics are not affected much. We conclude that the model performs quite well for a range of ρ_2.[19]

5.3. **Extensions.** In the benchmark private information model, bequests are neither very progressive nor very regressive, while gifts are progressive, as in the data. In addition, bequests are approximately equal in a large number of families. This section considers two extensions to the benchmark model and analyzes to what extent the main quantitative results are affected.

Nonnegative Bequests. Negative bequests are allowed in the benchmark model to focus on the basic forces, insurance and incentives, at work in the model. However, most countries do not allow parents to pass debt on to their children. To evaluate the sensitivity of the

lower threshold for these costs to matter as explanatory factors than does the public information (pure altruism) model.

[19]Appendix C reports the sensitivity of our results to the volatility of the life time wage process, σ_2. We find that our results are robust to this parameter as well as other parameters of the asset-wage specification (not reported). Appendix C also reports sensitivity results to utility parameters σ and γ.

Table 6. Restricting $b \geq 0$ in the Private Information Model

	Data	Without $b \geq 0$	With $b \geq 0$
β_{by}	0	-0.0367	-0.0201
β_{gy}	-0.02	-0.4537	-0.4739
Bequests within 2% of mean	77%	20.00%	33.07%
Bequests within 25% of mean	88%	55.87%	63.00%
Equal gifts	9%	20.00%	20.00%
Gifts within 20% of mean	11%	24.89%	24.89%

main results to this type of restriction, we add a constraint that bequests be weakly positive to the benchmark model (recalibrating ϕ, η and β). Table 6 summarizes the results of this extension. The number of approximately equal bequests and the regression coefficient β_{by} are both larger than in the benchmark, bringing the model closer to the data. On the other hand, the gift statistics are virtually unchanged under this specification.

Children with a Bequest Motive. In the benchmark model, time ends at the end of the second period when children die. This section considers a simple extension in which the children themselves give a bequest (to their own children, who are not modelled explicitly) and analyzes to what extent the structure of intergenerational transfers is affected. The utility children derive from bequest b' is assumed to be: $\beta \cdot u(b' + \xi)$, where ξ controls the degree of non-homotheticity (the extent to which transfers are a luxury good). This specification is a special case of a more general specification used in DeNardi (2004). ξ is calibrated so that the average b' is the same fraction of the children's second-period income as the bequest they receive relative to their own parents' income in the first period.[20] ϕ, η and β are calibrated to match the same targets as in the benchmark. The results are reported in Table 7.

As the table shows, this extension brings the model closer to the data as well. Bequests are almost unrelated to income and the dispersion of bequests is also smaller relative to the

[20]The results are very similar if one instead requires that the average b' relative to second period income matches the average $b + g$ relative to parents' income.

TABLE 7. Children With a Bequest Motive

	Data	Benchmark	Bequest motive
β_{by}	0	-0.0367	0.0002
β_{gy}	-0.02	-0.4537	-0.5042
Bequests within 2% of mean	77%	20.00%	37.72%
Bequests within 25% of mean	88%	55.87%	91.13%
Equal gifts	9%	20.00%	20.00%
Gifts within 20% of mean	11%	24.89%	24.05%

benchmark model. Gifts, on the other hand, are more progressive. In this version of the model, high income children care even more about the bequest than in the benchmark. The parent takes advantage of this fact and uses bequests even more to provide incentives.

6. Conclusion

In the data, bequests are typically equally divided, but gifts partially compensate for differences in child income within a family. This paper shows that the difference between gifts and bequests can arise as a result of a moral hazard problem because of the way that the timing of gifts and bequests affects children's incentives. We first build a stylized model with two productivity types and no interfamily heterogeneity. In this model, there is a level of productivity persistence for which bequests are equal and gifts are weakly compensatory, as in the data. The forces behind this result are common in dynamic insurance models. In these models, a payout today comes at the cost of reduced future help. In our model, compensatory gifts are accompanied by less compensatory (equal) bequests.

We then build a richer model with interfamily heterogeneity and evaluate the model quantitatively, a novelty in the literature. In this model, no simple two-parameter productivity process common to all households makes bequests equal within all families. However, bequests are nearly uncorrelated with income and approximately equal in many families, which is consistent with the data. Gifts, in contrast, are weakly progressive. Both bequests and gifts are more progressive and bequests are more dispersed in the public information version of the model. Adding the private information component thus significantly helps to bring the pattern of intergenerational transfers closer to the data. Finally, the welfare loss of moving from the optimal policy to a policy of equal division is small in our model, but substantial (20 times larger) in its public information counterpart. Equal division of bequests is a nearly optimal policy in our model, possibly explaining its prevalence in the data.

The data on intergenerational transfer behavior pose a serious puzzle for which no consensus solution has arisen. Since this behavior is at the center of long-run intertemporal economic decisions, the problem is one not just for the economics of the family but also for many other areas of economics that are concerned with capital accumulation and distribution or that rely on models of those behaviors. Standard long-run macroeconomic models all use some form of perfect information in intergenerational transfers. In contrast, our work suggests that savings and transfer decisions may hinge on the dynamic properties of insurance models with private information.

Appendix A. PSID

A.1. **PSID Sample Selection and Variables of Interest.** We follow a procedure similar to Heathcote, Perri, and Violante (2010) and work with the representative SRC sample for the period 1968 - 2013. Households whose heads report positive labor income and no hours worked and vice versa, as well as households that report hourly wages less than 10% of the minimum federal wage and households whose heads do not report age, are excluded.

Wealth. In our sample, wealth statistics are reported in 1984, 1989, 1994, and then in every wave from 1999 onwards. Household wealth is defined as the sum of the value of the farm/business, checking/savings accounts, real estate, stocks, other assets, vehicles net of debt plus the home value net of mortgage. Wealth is reported within given thresholds (-10 million to +100 million), but these thresholds never bind. We adjust for inflation by using the CPI.

Wages. To estimate the wage process, we focus on the wages of male household heads (these are consistently reported throughout the sample period), which are constructed as annual labor earnings divided by annual hours, adjusted for inflation by using the CPI. No adjustment is made for top coding (a very small number of people are subject to top-coding in PSID). For the missing years between 1997 and 2013 when PSID becomes bi-annual, wages are constructed as simple averages of the two adjacent observations.

Earnings. (Total) household earnings are defined as the sum of the annual labor earnings of male household heads and their wives.

Bequests. Starting in 1988, PSID reports whether the household has received an inheritance (prior to 1988 PSID asks about transfers including inheritance). No adjustment is made for top-coding, since there are only a handful of such observations. It is not possible to identify who received the bequest (head or spouse) or who made the bequest.

Gifts. There is not enough information in PSID to determine the exact flows of gifts in a way that could be mapped to our model. We use the average (financial) help from relatives (of the head and the wife combined) as a measure of gifts. Since for wives this measure is only reported from 1985 onwards, we focus on the time period 1985 - 2013.

A.2. **Estimation of the Productivity Process.** This subsection describes the estimation of the wage/productivity process summarized in equations (5.1) - (5.3). The equations to be estimated are:

$$(\log z_P, \log A) \sim \mathcal{N}_2\left(\begin{pmatrix}\mu_p \\ \mu_A\end{pmatrix}, \begin{pmatrix}\sigma_p^2 & corr(A, z_p)\cdot\sigma_p\sigma_A \\ corr(A, z_p)\cdot\sigma_p\sigma_A & \sigma_A^2\end{pmatrix}\right), \tag{A.1}$$

$$\log z_1 = \mu_1 + \rho_1 (\log z_p - \mu_p) + \varepsilon_1, \qquad \varepsilon_1 \sim \mathcal{N}(0, \sigma_1^2), \tag{A.2}$$

$$\log z_2 = \mu_2 + \rho_2 (\log z_1 - \mu_1) + \varepsilon_2, \qquad \varepsilon_2 \sim \mathcal{N}(0, \sigma_2^2). \tag{A.3}$$

where z_p is the parent's productivity level and A is the parent's asset level. The child's productivity level in the first period is z_1 and in the second period z_2. Labor productivities z_1, z_2, z_p correspond to hourly wages in the data, averaged over ages 25-50 (z_1) and 50-75 (z_2 and z_p), respectively.

These equations/parameters are estimated separately using as much information as possible for each of them. As for equation (A.1), there is enough data in PSID to construct 25-year long wage histories for people aged 50-75 and hence these are used to compute the statistics. The average hourly log real wage between ages 50-75 is 2.02 (this corresponds to 7.54 dollars). To be used as model input μ_p, this number is adjusted, taking into account that one period corresponds to 25 years. The other parameters are estimated to be $\sigma_p = 0.78, corr(A, z_p) = 0.65, \sigma_A = 1.60$, and $\mu_A = 11.23$, which corresponds to approximately 75,000 dollars. Wealth A is calculated at age 50.

To estimate equation (A.2), first father-son pairs are constructed. There is enough data in PSID to construct 25-year wage histories for fathers aged 50-75 and for their sons aged 25-50. The intergenerational wage persistence parameter ρ_1 is found to be 0.25 and the error variance in equation (A.2), σ_1, is estimated to be 0.47 (these values do not change substantially when age and time effects are included). The means μ_1 and μ_p are recovered during the estimation of equations (A.3) and (A.1), respectively. The estimation of equation (A.3) is explained in the main text.

Appendix B. Private Information in HRS

Table 8 reports the fraction of parents in several waves of the Health and Retirement Survey (HRS), who report that they do not know their children's income (provided that they do not refuse to respond to this question). In the HRS, this means that they do not know whether their child's income is lower than $10,000, between $10,000 - $35,000 or above $35,000. An additional fraction of parents report not knowing the exact income of their child when the subsequent questions narrow the income brackets. Clearly, if parents do not have perfect information about their children's income, they cannot know its individual components: labor productivity and labor hours, which is what we assume in this paper.

Table 8. Private Information in HRS

2010	2008	2006	2004	2002
27%	37%	46%	38%	38%

This table reports the percentage of parents in HRS, who report that they do not know their children's income.

Appendix C. Benchmark Private Information Model Sensitivity Results

Table 9. Sensitivity to Wage Volatility σ_2

	Data	Benchmark		
σ_2		$0.9 \cdot 0.88$	0.88	$1.1 \cdot 0.88$
β_{by}	0	-0.0384	-0.0367	-0.0353
β_{gy}	-0.02	-0.4719	-0.4537	-0.4633
Bequests within 2% of mean	77%	20.00%	20.00%	20.00%
Bequests within 25% of mean	88%	56.95%	55.87%	55.17%
Equal gifts	9%	20.00%	20.00%	20.00%
Gifts within 20% of mean	11%	24.89%	24.89%	24.89%

Table 10. Sensitivity to Risk Aversion Parameter σ

	Data	Benchmark		
σ		0.5	1	2
β_{by}	0	-0.0412	-0.0367	-0.0203
β_{gy}	-0.02	-0.4087	-0.4537	-0.5530
Bequests within 2% of mean	77%	20.00%	20.00%	32.15%
Bequests within 25% of mean	88%	42.85%	55.87%	74.21%%
Equal gifts	9%	20.00%	20.00%	20.00%
Gifts within 20% of mean	11%	24.05%	24.89%	24.89%

Table 11. Sensitivity to Labor Supply Elasticity Parameter γ

	Data	Benchmark		
γ		0.5	1	2
β_{by}	0	-0.0069	-0.0367	-0.0670
β_{gy}	-0.02	-0.4241	-0.4537	-0.5605
Bequests within 2% of mean	77%	27.45%	20.00%	20.00%
Bequests within 25% of mean	88%	77.99%	55.87%	74.21%%
Equal gifts	9%	20.00%	20.00%	20.00%
Gifts within 20% of mean	11%	27.25%	24.89%	21.69%

Appendix D. The Benchmark Model without Family-Level Uncertainty

D.1. Characterization of W and $\mathcal{W}$.

Lemma D.1. *$\mathcal{W}(b,z), \mathcal{W}_b(b,z), \mathcal{W}_z(b,z) \in \mathcal{C}^1$ on $\Re \times \Re_{++}$.*

Proof: First we show that the policy function $c(b,z) \in \mathcal{C}$. Recall that $\mathcal{W}(b,z) = \max_c u(c) - v(\frac{c-b}{z})$ s.t. $c \geq b$. Standard arguments (strict concavity of the objective function, convexity of the constraint set, and the Inada conditions) imply that this problem has a unique interior solution $\forall b, z$. Consumption solves $zu'(c) - v'(\frac{c-b}{z}) = 0$. The left-hand side is C^1, and the derivative with respect to c is strictly negative: $zu''(c) - v''(\frac{c-b}{z}) < 0$, hence invertible for all c, b, z. Thus, $c(b,z)$ is a continuously differentiable function by the implicit function theorem, and hence $\mathcal{W}(b,z) = u(c(b,z)) - v(\frac{c(b,z)-b}{z})$ is a continuously differentiable function as well. By the envelope theorem, $\mathcal{W}_b(b,z) = u'(c(b,z))$ and $\mathcal{W}_z(b,z) = u'(c(b,z))\frac{c(b,z)-b}{z}$. Since $c(b,z)$ is $\mathcal{C}^1$, so are $\mathcal{W}_b$ and $\mathcal{W}_z$. □

Lemma D.2. *$\mathcal{W}_b > 0 \quad \& \quad \mathcal{W}_{bb} < 0 \quad \& \quad \mathcal{W}_b$ is decreasing in z.*

Proof: (i) The Envelope Theorem implies that $\mathcal{W}_b = \frac{1}{z}v'(\frac{c-b}{z}) = u'(c) > 0$. Clearly, $\mathcal{W}(b,z)$ is increasing in b. (ii) Differentiation with respect to b yields $\frac{\partial}{\partial b}\mathcal{W}_b = \frac{1}{z^2}v''(\frac{c-b}{z})(\frac{\partial c}{\partial b} - 1)$. Then it is enough that $\frac{\partial c}{\partial b} < 1$, which must be true since more money must reduce output somewhat. We can establish that by noting that $\frac{\partial}{\partial b}\mathcal{W}_b = \frac{\partial u'(c)}{\partial b} = u''(c)\frac{\partial c}{\partial b}$. Setting the two expressions equal, we get $\frac{\partial c}{\partial b} = \frac{v''(\frac{c-b}{z})}{v''(\frac{c-b}{z}) - z^2 \cdot u''(c)} < 1 \Rightarrow \frac{\partial}{\partial b}\mathcal{W}_b < 0$. (iii) Differentiation with respect to z yields $\frac{\partial}{\partial z}\mathcal{W}_b = u''(c)\frac{\partial c}{\partial z}$. $\frac{\partial c}{\partial z} > 0$ which we obtain by differentiating the first order condition $zu'(c) - v'(\frac{c-b}{z}) = 0$ to get $\frac{\partial c}{\partial z} = \frac{\frac{c-b}{z}v''(\frac{c-b}{z}) + z \cdot u'(c)}{v''(\frac{c-b}{z}) - z^2 u''(c)} > 0 \Rightarrow \frac{\partial}{\partial z}\mathcal{W}_b < 0$. □

Lemma D.3. *$W_b > 0, W_{bb} < 0$. $\pi \geq \frac{1}{2} \Longrightarrow W_b(b, z_L) > W_b(b, z_H)$.*

Proof: The first two follow from the previous lemma and the definition: $W(b, z_i) := \pi\mathcal{W}(b, z_i) + (1-\pi)\mathcal{W}(b, z_j)$. Since we have $\mathcal{W}_b(b, z_L) > \mathcal{W}_b(b, z_H)$ and $\pi \geq \frac{1}{2}$, we have $W_b(b, z_L) = \pi\mathcal{W}_b(b, z_L) + (1-\pi)\mathcal{W}_b(b, z_H) > (1-\pi)\mathcal{W}_b(b, z_L) + \pi\mathcal{W}_b(b, z_H) = W_b(b, z_H)$. □

D.2. **Proof of Lemma 4.2.** Total transfers are progressive: $g_L + b_L > g_H + b_H$.

Proof: Consider an artificial problem of a high type who is given a transfer x_H. We let him decide how much he wants to work, consume, and save. His problem is the following

$$\max_{c,y,b} u(c) - v(\frac{y}{z_H}) + W(b, z_H) \qquad \text{s.t.} \qquad c + b \leq y + x_H.$$

Notice that the first order conditions for this problem are the same as those for the high type in the parent's problem without IC_L (i.e., relaxed problem). Thus, by setting $x_H = c_H + b_H - y_H$, we guarantee that the unique solution (the objective function is strictly concave, the constraint set is convex) to the maximization problem above is $(c^*, y^*, b^*) = (c_H, y_H, b_H)$, the high type's allocations in the relaxed problem.

Since the IC_H is binding in the relaxed problem, c_L, y_L, b_L give the high type the same utility as c_H, y_H, b_H. Thus, c_L, y_L, b_L cannot be in the constraint set of the artificial problem above (otherwise a convex combination of c_L, y_L, b_L and c_H, y_H, b_H would give a strictly higher value of the objective function while still being in the constraint set).

Thus $c_L + b_L > y_L + x_H$ and hence $c_L + b_L > y_L + y_L + A - c_L - b_L$. This implies $(c_L - y_L) + b_L = g_L + b_L > \frac{A}{2}$ and by feasibility $(c_H - y_H) + b_H = g_H + b_H < \frac{A}{2}$. □

D.3. **High-type bequests are higher than high-type gifts.**

Lemma D.4. *$g_H < b_H$ wherever IC_H binds except for $\pi = 1$ where they are equal.*

Proof: High type allocations are undistorted:

$$u'(c_H^1) = v'(y_H^1/z_H)/z_H = W_b(b_H, z_H) = \pi\mathcal{W}_b(b_H, z_H) + (1-\pi)\mathcal{W}_b(b_H, z_L).$$

For $\pi = 1$ we have

$$u'(c_H^1) = v'(y_H^1/z_H)/z_H = \mathcal{W}_b(b_H, z_H) = u'(c_H^2) = v'(y_H^2/z_H)/z_H$$

$$\Rightarrow c_H^1 = c_H^2 \quad y_H^1 = y_H^2 \quad \Rightarrow g_H = c_H^1 - y_H^1 = c_H^2 - y_H^2 = b_H.$$

For $\pi \in [.5, 1)$, using $\mathcal{W}_b$ decreasing in z and $\pi < 1$ we have

$$\mathcal{W}_b(b_H, z_H) < W_b(b_H, z_H) < \mathcal{W}_b(b_H, z_L) \qquad \text{hence:}$$

$$u'(c_H^1) = v'(y_H^1/z_H)/z_H > \mathcal{W}_b(b_H, z_H) = u'(c_{HH}^2) = v'(y_{HH}^2/z_H)/z_H,$$

where the subscript XY indicates (second period's) allocations for a child that was type X in period one and type Y in period two. The last expression implies $c_H^1 < c_{HH}^2, y_H^1 > y_{HH}^2$ and hence $g_H = c_H^1 - y_H^1 < c_{HH}^2 - y_{HH}^2 = b_H$. □

D.4. **Proof of Lemma 4.5.** Policy functions are continuous in π and z_L if u has non-increasing absolute risk aversion, and v is CARA or CRRA.

Proof: The proof has two steps. In the first, we show that under our assumptions, the policy correspondence is single valued. In the second that it is upper hemi-continuous in π and z. Combined, these properties imply that the policy correspondence is in fact a continuous function.

Step 1. The policy is a function (i.e., a single valued correspondence). We can convexify the constraint set by having the parent choose utility values instead of real values. In this setup, the parent chooses utility from consumption, output, and welfare for the high type (u_H, v_H, w_H), and for the high type pretending to be the low type (u_L, v_L, w_L). Here, we use capital letters to denote the functions U, V, and W. The parent solves

$$\max_{u,v,w} \quad u_H - v_H + w_H + u_L - V(\frac{z_H}{z_L}V^{-1}(v_L)) \quad + \quad W_L(W_H^{-1}(w_L)) \qquad \text{s.t.}$$

$$U^{-1}(u_H) + U^{-1}(u_L) + W_H^{-1}(w_H) + W_H^{-1}(w_L) \leq z_H V^{-1}(v_H) + z_H V^{-1}(v_L) + 2A,$$

$$u_H - v_H + w_H \geq u_L - v_L + w_L.$$

The incentive constraint is linear. Since $U(\cdot)$ and $W_i(\cdot)$ are strictly increasing and strictly concave, their inverses are strictly convex. Similarly, the inverse of $V(\cdot)$ is strictly concave. The convex functions are on the lesser side of the inequality and the concave ones are on the greater side, so this constraint is convex as well.

To show the uniqueness of the solution, it remains to demonstrate that the objective function is weakly concave. The first four terms are linear. We will show that our assumptions guarantee that V and W are weakly concave. Both of the functions are of the form $h(x) = f(g^{-1}(x))$. Recall that $\frac{\partial}{\partial x} g^{-1}(x) = \frac{1}{g'(g^{-1}(x))}$. Then: $h'(x) = \frac{f'(g^{-1}(x))}{g'(g^{-1}(x))}$ and $h''(x) = \left[g'(g^{-1}(x)) \frac{f''(g^{-1}(x))}{g'(g^{-1}(x))} - f'(g^{-1}(x)) \frac{g''(g^{-1}(x))}{g'(g^{-1}(x))} \right] / (g'(g^{-1}(x)))^2$. The sign of the second derivative is the sign of the numerator, which can be written as $\left(-\frac{g''(g^{-1}(x))}{g'(g^{-1}(x))} \right) - \left(-\frac{f''(g^{-1}(x))}{f'(g^{-1}(x))} \right)$. We want to show that the second derivative of W_L as a function of w_L is negative and that it is positive for V as a function of v_L, since this function is subtracted, so we need it to be convex. Let $RR_V(l)$ denote the relative risk aversion of V and similarly the absolute risk aversion $AR_W(b; z)$ where the derivatives are taken with respect to b. Plugging into the formulas above we get $\frac{\partial^2 V(\frac{z_H}{z_L} V^{-1}(v_L))}{\partial v_L^2} \geq 0 \iff RR_V(\frac{y}{z_H}) \geq RR_V(\frac{y}{z_L})$ and $\frac{\partial^2 W_L(W_H^{-1}(w_L))}{\partial w_L^2} \leq 0 \iff AR_W(b; z_H) \leq AR_W(b; z_L)$. Thus, v needs to be NIRRA, which covers CARA. For the last term in the objective function, we prove the following lemma.

Lemma D.5. *Assume u is NIARA (covers CRRA) and v is CARA or CRRA. Then $-\frac{\mathcal{W}_{bb}}{\mathcal{W}_b}$ is decreasing in z and $\mathcal{W}_{bb}$ is increasing in z.*

Once we prove this claim, it is easy to show that $-\frac{W_{bb}(b,z_L)}{W_b(b,z_L)} \geq -\frac{W_{bb}(b,z_H)}{W_b(b,z_H)}$ and $W_{bb}(b, z_L) \leq W_{bb}(b, z_H)$ using $\pi \geq \frac{1}{2}$ in a way similar to lemma D.3.

Proof: From above we have that:

$$\frac{\mathcal{W}_{bb}}{\mathcal{W}_b} = \frac{u''(c)}{u'(c)}\frac{\partial c}{\partial b} = \frac{\frac{1}{z}v'(\frac{c(z)-b}{z})}{\frac{1}{z^2}v''(\frac{c(z)-b}{z})} + \frac{u'(c(z))}{-u''(c(z))}.$$

So now, $\frac{\mathcal{W}_{bb}}{\mathcal{W}_b}$ increasing in $z \iff \frac{\frac{1}{z}v'(\frac{c(z)-b}{z})}{\frac{1}{z^2}v''(\frac{c(z)-b}{z})} + \frac{u'(c(z))}{-u''(c(z))}$ increasing in z. The second term is increasing in z for u NIARA since c is increasing in z. We claim that the first term is increasing in z for v CRRA. To see that, note that the first term can be rewritten as

$$\frac{y(z)\cdot v'(\frac{y(z)}{z})}{\frac{y(z)}{z}\cdot v''(\frac{y(z)}{z})} = \frac{y(z)}{-RR_v(y(z))}.$$

Remember that $RR_v(y(z))$ is negative so that $-RR_v(y(z))$ is positive. Assuming CRRA $-RR_v(y(z))$ is constant, we get that $\frac{\partial}{\partial z}\frac{y(z)}{-RR_v(y(z))} = \frac{\frac{\partial y}{\partial z}}{-RR_v(y(z))} > 0$ since y is increasing in z (by c is increasing in z which was established above). Note that it is not straightforward to find a sufficient condition in terms of DRRA or IRRA, because we donot know the sign of $\frac{\partial \ell(z)}{\partial z}$. The argument is similar for CARA (again we need not worry which way $\ell(z)$ goes). Another sufficient condition would be $\ell(z)$ constant in z. The functions that we are using do not have this property, so we do not include this sufficient condition in the statement of the lemma. Finally, the fact that $\mathcal{W}_{bb}$ is (strictly) increasing in z follows from the fact that $\mathcal{W}_b$ is (strictly) decreasing in z.

Step 2. The policy correspondences are upper hemi-continuous in π and z.

Proof: for upper hemi-continuity in π, define $f(\pi; c_H, c_L, y_H, y_L, b_H, b_L) := u(c_H) - v(\frac{y_H}{z_H}) + \pi\mathcal{W}(b_H, z_H) + (1-\pi)\mathcal{W}(b_H, z_L) + u(c_L) - v(\frac{y_L}{z_L}) + \pi\mathcal{W}(b_L, z_L) + (1-\pi)\mathcal{W}(b_L, z_H), \Gamma(\pi) := \{(c_H, c_L, y_H, y_L, b_H, b_L) \in \Re^4_+ \times \Re^2 : c_H + c_L + b_H + b_L \leq y_H + y_L + A, u(c_H) - v(\frac{y_H}{z_H}) + \pi\mathcal{W}(b_H, z_H) + (1-\pi)\mathcal{W}(b_H, z_L) \geq u(c_L) - v(\frac{y_L}{z_H}) + \pi\mathcal{W}(b_L, z_H) + (1-\pi)\mathcal{W}(b_L, z_L)\}$ and $h(\pi) := \{(c^*_H, c^*_L, y^*_H, y^*_L, b^*_H, b^*_L) \in \Re^4_+ \times \Re^2 : f(\pi; c^*_H, c^*_L, y^*_H, y^*_L, b^*_H, b^*_L) = \max_{(c_H, c_L, y_H, y_L, b_H, b_L)\in\Gamma(\pi)} f(\pi; c_H, c_L, y_H, y_L, b_H, b_L)\}$. We have shown above that $\mathcal{W}(b, z)$ is continuous in b. This means that f is a continuous mapping $[0,1] \times \Re_+^4 \times \Re^2 \longrightarrow \Re$. Clearly, $\Gamma : [0,1] \longrightarrow \Re_+^6$ is a non-empty valued correspondence. Γ is also a continuous correspondence, since all the functions are continuous and π enters linearly. One can show that $\exists L$ large enough and $\exists B$ small enough s.t. $\forall \pi : l_H, l_L < L, b_H, b_H > B$ at a solution to the relaxed problem. Then WLOG Γ can be made compact valued, since these bounds imply an upper bound on c_H, c_L, y_L, y_H as well. Thus, $h(\pi)$ is a non-empty, compact, upper hemi-continuous correspondence. The proof of upper hemi-continuity of the policy correspondences in z_L and z_H on $\Re_{++}$ with $z_L \leq z_H$ is similar. For $z_L = 0$, one can show that policies are right continuous.

D.5. **Relaxed problem valid.**

Lemma D.6. *The constraint preventing the high type from reporting the low productivity level, IC_H, is sufficient (i.e. at the solution to the relaxed problem where IC_L is not included, IC_L will be satisfied for $\pi \in \{\frac{1}{2}, \pi^*, 1\}$). For $\pi = 1$ we have the result as long as $b_L > b_H$, which has been established for z_L small enough and v is CRRA or CARA.*

Proof: Clearly, at the solution to the relaxed problem, the IC_H binds. To prove the lemma, we need to characterize the solutions in more detail. We will proceed case by case. Note that for $z_L = 0$, the validity of the relaxed problem is clear, because the low type cannot pretend to be the high type.

(i) $\pi = \frac{1}{2}$. First, note that $y_H > y_L$. This is because, as we established above, $c_H > c_L, b_H > b_L$ and the IC_H binds. Moreover $W(b_i, z) = W(b_i)$. Denote $W(b_i)$ as w_i. Then we want to show that if $u(c_H) - v(\frac{y_H}{z_H}) + w_H = u(c_L) - v(\frac{y_L}{z_H}) + w_L$ then $u(c_L) - v(\frac{y_L}{z_L}) + w_L \geq u(c_H) - v(\frac{y_H}{z_L}) + w_H$. This is equivalent to $u(c_H) - u(c_L) + w_H - w_L = v(\frac{y_H}{z_H}) - v(\frac{y_L}{z_H}) \Rightarrow v(\frac{y_H}{z_L}) - v(\frac{y_L}{z_L}) \geq u(c_H) - u(c_L) + w_H - w_L$. Thus it is enough to show that $v(\frac{y_H}{z_L}) - v(\frac{y_L}{z_L}) \geq v(\frac{y_H}{z_H}) - v(\frac{y_L}{z_H})$, which is equivalent to $\int_{y_L}^{y_H} \frac{v'(\frac{y}{z_L})}{z_L} dy \geq \int_{y_L}^{y_H} \frac{v'(\frac{y}{z_H})}{z_H} dy$. This is true by convexity of v, $y_H > y_L$, and $z_H > z_L$.

(ii) $\pi = 1$. Showing $y_H > y_L$ is more complicated here. It deserves a separate lemma.

Lemma D.7. *Suppose v is NIRRA. Then $y_L^1 < y_H^1$.*

Proof: First, we will show that $y_L^1 < y_L^2$. We use superscripts for periods. We have $\pi = 1$ so the types are constant over time. Define the output that will be chosen in the second period by a high type who misreported in the first period as a function of b_L:$\tilde{y} := u'(\tilde{y}(b_L) + b_L) = v'(\frac{\tilde{y}(b_L)}{z_H}) \cdot \frac{1}{z_H}$. The properties of u and v imply that y^2 is strictly increasing in z and therefore $\forall b : y_L^2(b) < \tilde{y}(b)$. Now, output in the first period is determined by $v'(\frac{y_L^1}{z_L})\frac{1}{z_L} - \mu v'(\frac{y_L^1}{z_H})\frac{1}{z_H} = W_1(b_L, z_L) - \mu W_1(b_L, z_H)$. Using $W_1(b_L, z_H) = v'(\frac{\tilde{y}(b_L)}{z_H}) \cdot \frac{1}{z_H}$ and $W_1(b_L, z_L) = v'(\frac{y_L^2(b_L)}{z_L})\frac{1}{z_L}$ we can rewrite this as $v'(\frac{y_L^1}{z_L})\frac{1}{z_L} - \mu v'(\frac{y_L^1}{z_H})\frac{1}{z_H} = v'(\frac{y_L^2(b_L)}{z_L})\frac{1}{z_L} - \mu v'(\frac{\tilde{y}(b_L)}{z_H}) \cdot \frac{1}{z_H}$. Since $y_L^2(b) < \tilde{y}(b)$, this implies (dropping the arguments for simplicity) $v'(\frac{y_L^1}{z_L})\frac{1}{z_L} - \mu v'(\frac{y_L^1}{z_H})\frac{1}{z_H} < v'(\frac{y_L^2}{z_L})\frac{1}{z_L} - \mu v'(\frac{y_L^2}{z_H}) \cdot \frac{1}{z_H}$.

Now to prove the claim, we will show that $f(y) := v'(\frac{y}{z_L})\frac{1}{z_L} - \mu v'(\frac{y}{z_H})\frac{1}{z_H}$ is increasing in y, which implies that $y_L^1 < y_L^2$. Taking the derivative yields $f'(y) = v''(\frac{y}{z_L})\frac{1}{z_L^2} - \mu v''(\frac{y}{z_H})\frac{1}{z_H^2}$. By NIRRA we have that $-\frac{v''(\frac{y}{z_L})}{v'(\frac{y}{z_L})}\frac{y}{z_L} \leq -\frac{v''(\frac{y}{z_H})}{v'(\frac{y}{z_H})}\frac{y}{z_H} \Longrightarrow v''(\frac{y}{z_L})\frac{1}{z_L^2} > v''(\frac{y}{z_H})\frac{1}{z_H^2}$. The reasoning comes from $v', v'' > 0$ and $z_L < z_H$. Combine the last expression with $\mu \in (0, 1)$ to get $f'(y) > 0$. Thus, $y_L^1 < y_L^2$. To prove the lemma, suppose by way of contradiction that $y_H^1 \leq y_L^1$. Then either $b_H \geq b_L$ or $b_H < b_L$. The first is inconsistent with the IC_H binding because $c_H > c_L$. For the second we would have $y_L^2 < y_H^2$ and hence $y_L^1 < y_L^2 < y_H^2 = y_H^1$, a contradiction. Thus, $y_L^1 < y_H^1$. □

To show that the relaxed problem is valid, we want to show that $u(c_H) - v(\frac{y_H}{z_H}) + W(b_H, z_H) = u(c_L) - v(\frac{y_L}{z_H}) + W(b_L, z_H) \Rightarrow \quad u(c_L) - v(\frac{y_L}{z_L}) + W(b_L, z_L) \geq u(c_H) - v(\frac{y_H}{z_L}) + W(b_H, z_L)$. Thus, it is enough to show that $v(\frac{y_H}{z_L}) - v(\frac{y_L}{z_L}) + W(b_L, z_L) - W(b_H, z_L) \geq v(\frac{y_H}{z_H}) - v(\frac{y_L}{z_H}) + W(b_L, z_H) - W(b_H, z_H)$. Since we have $b_L > b_H, y_L < y_H$ we can rewrite the inequality as $\int_{y_L}^{y_H} \frac{v'(\frac{y}{z_L})}{z_L} dy + \int_{b_H}^{b_L} W_b(b, z_L) db \geq \int_{y_L}^{y_H} \frac{v'(\frac{y}{z_H})}{z_H} dy + \int_{b_H}^{b_L} W_b(b, z_H) db$, which is true by the convexity of v, $z_H > z_L$, and $W_b(b, z_L) \geq W_b(b, z_H)$. □

(iii) $\pi = \pi^*$. The Ws in the ICs cancel out by $b_L = b_H$. Moreover, IC_H binding implies that $y_H > y_L$. Thus, this reduces to showing the following, which has been already shown above: $v(\frac{y_H}{z_L}) - v(\frac{y_L}{z_L}) \geq v(\frac{y_H}{z_H}) - v(\frac{y_L}{z_H})$. □

Appendix E. Model with Two States and Signals

In this section, we analyze a variant of the model described in section 4. The only difference is that now the parent first receives a signal about her children's productivity. This model subsumes the private and public information models we analyze in the main body of this paper as special cases. We show that in the two-type version of the model with informative signals, there is no level of persistence such that all children within a family are given an equal bequest. However, we also show that our model with private information is robust to adding signals in the sense that the welfare loss of equal division does not increase much as we add signals. In this sense, the private information model without signals is a good approximation of the model with signals (as long as the information content is not "too high").

E.1. **Model.** The basic setup of the model is the same as in section 4. A parent has a continuum of children. One-half of them is more productive (type H) than the other half (type L). In the second period, children remain of the same type with probability π and switch type with probability $1 - \pi$. The parent does not know which is which. However (and this is how this model differs from the one described in section 4), the parent receives a signal on the children's productivity $\{\theta_H, \theta_L\}$. This is the children's probability of becoming the high type, with $\theta_H \geq \theta_L$. Without loss of generality we assume that one-half of children draw a signal θ_H and one-half of children draw θ_L, with the appropriate consistency condition: $\frac{1}{2}\theta_H + \frac{1}{2}\theta_L = \frac{1}{2}$ (i.e., $\theta_H + \theta_L = 1$). The parent's problem (for the allocations, the first subscript denotes the signal, the second the realized type) is as follows:

$$\max_{c,y,b} \sum_i \theta_i \left[u(c_{iH}) - v(\frac{y_{iH}}{z_H}) + W(b_{iH}, z_H) \right] \quad + \quad (1-\theta_i) \left[u(c_{iL}) - v(\frac{y_{iL}}{z_L}) + W(b_{iL}, z_L) \right] \qquad \text{s.t.}$$

$$\sum_i \theta_i [c_{iH} - y_{iH} + b_{iH}] + (1-\theta_i)[c_{iL} - y_{iL} + b_{iL}] \quad \leq \quad A,$$

$$\forall i,j : u(c_{ij}) - v(\frac{y_{ij}}{z_j}) + W(b_{ij}, z_j) \quad \geq \quad u(c_{ij^c}) - v(\frac{y_{ij^c}}{z_j}) + W(b_{ij^c}, z_j).$$

Here j^c defines the complement of j in the set(L, H). There are two special cases: (i) $\theta_H = \theta_L = \frac{1}{2}$ (i.e., signals are not informative at all), and (ii) $\theta_H = 1$ (i.e. full information). In the main body of this paper, we focus on these two special cases. In the model with informative signals, the parent is solving two separate problems for children of each signal that are connected only though the budget constraint. We denote $\theta := \theta_H = 1 - \theta_L$ and focus on the relaxed problem with downward constraints only (the expressions on the left denote the Lagrange multipliers we will use):[21]

$$(\theta\mu_{HH}) : u(c_{HH}) - v(\frac{y_{HH}}{z_H}) + W(b_{HH}, z_H) \quad \geq \quad u(c_{HL}) - v(\frac{y_{HL}}{z_H}) + W(b_{HL}, z_H),$$

$$((1-\theta)\mu_{LH}) : u(c_{LH}) - v(\frac{y_{LH}}{z_H}) + W(b_{LH}, z_H) \quad \geq \quad u(c_{LL}) - v(\frac{y_{LL}}{z_H}) + W(b_{LL}, z_H).$$

E.2. **Characterization.** We are now ready to show that, with informative signals (i.e., $\theta > \frac{1}{2}$), there no measure zero set in the parameter space in which all four types (i.e. HH, HL, LH, LL) receive the same bequest.

Theorem E.1. *For any parameters, there is no $\theta > \frac{1}{2}$ s.t. $b_{HH} = b_{HL} = b_{LH} = b_{LL}$.*

[21] The relaxed problem is trivially valid for $z_L \in N_\varepsilon(0)$, for which we have established our main result in section 4. In the quantitative examples we report below, we check and find that the relaxed problem is always valid as well.

Proof: The first order conditions on bequests from this problem are

$$\begin{aligned}\lambda &= (1+\mu_{HH})W_1(b_{HH},z_H) = (1+\mu_{LH})W_1(b_{LH},z_H)\\ &= W_1(b_{HL},z_L) - \frac{\theta}{1-\theta}\mu_{HH}W_1(b_{HL},z_H) = W_1(b_{LL},z_L) - \frac{1-\theta}{\theta}\mu_{LH}W_1(b_{LL},z_H).\end{aligned}$$

$b_{HH} = b_{LH}$ requires $\mu_{HH} = \mu_{LH}$. This implies that $b_{HL} = b_{LL}$ requires $\frac{\theta}{1-\theta} = \frac{1-\theta}{\theta}$, which is equivalent to $\theta = \frac{1}{2}$. □

Note that this proof does not rely on the transition matrix between states in the first and second period being symmetric. We show in section 4.4 that with symmetric transition matrices and uninformative signals $\theta = \frac{1}{2}$, there is a level of persistence π such that $b_{HH} = b_{HL} = b_{LH} = b_{LL}$. This follows from the fact that with uninformative signals, children with different signals are treated the same way. Note that when signals are perfectly informative (i.e., $\theta = 1$) and productivity shocks are i.i.d. (i.e., $\pi = \frac{1}{2}$), then $b_{HH} = b_{LL}$ as shown in section 4.3. However, the two measure zero types (i.e., LH and HL) do not get the same bequest: $b_{HH} = b_{LL} \neq b_{LH} \neq b_{HL}$.

Next we show that positive surprises (low signal and high realization) are rewarded and that negative surprises (high signal and low realization) are punished. This seems a fairly intuitive way to provide incentives.

Theorem E.2. $c_{HH} < c_{LH}, y_{HH} > y_{LH}, b_{HH} < b_{LH}, c_{HL} < c_{LL}, y_{HL} > y_{LL}, b_{HL} < b_{LL}$.

This result follows from the following lemma, which we provide without proof.

Lemma E.3. $\theta > \frac{1}{2}, \pi \geq \frac{1}{2}$ *implies* $\mu_{LH} > \mu_{HH}$ *and* $\mu_{LH}\frac{1-\theta}{\theta} < \mu_{HH}\frac{\theta}{1-\theta}$.

Next, we prove that the welfare loss of equal division κ at $\theta = .5$ is robust to adding signals in the sense that $\frac{\partial \kappa}{\partial \theta}|_{\theta=.5} = 0$. κ is defined as the welfare loss of equal division in consumption units. Specifically, we use the following approximation: $\frac{V(A,\theta)-V_{=}(A,\theta)}{\frac{\partial V_{=}(A,\theta)}{\partial A}}$.[22] Here, $V(A,\theta)$ is the solution to the parents' problem above and $V_{=}(A,\theta)$ is a solution to the same problem with added constraints: $b_{HH} = b_{HL} = b_{LH} = b_{LL}$. Note that the theorem also applies to the model with multiple types and family-level uncertainty that we use in section 5.

Theorem E.4. *Assume that $V(A,\theta)$ and $V_{=}(A,\theta)$ are twice differentiable. Then* $\frac{\partial \kappa}{\partial \theta}|_{\theta=.5} = 0$.

Proof: Both value functions are symmetric around $\theta = .5$. Therefore, their derivatives $\frac{\partial V(A,\theta)}{\partial A}$ and in particular $\frac{\partial V_{=}(A,\theta)}{\partial A}$ are symmetric around $\theta = .5$. This implies that $\partial \left(\frac{\partial V_{=}(A,\theta)}{\partial A}\right)/\partial\theta = 0$. Using this fact, we can express the derivative of κ with respect to θ as:

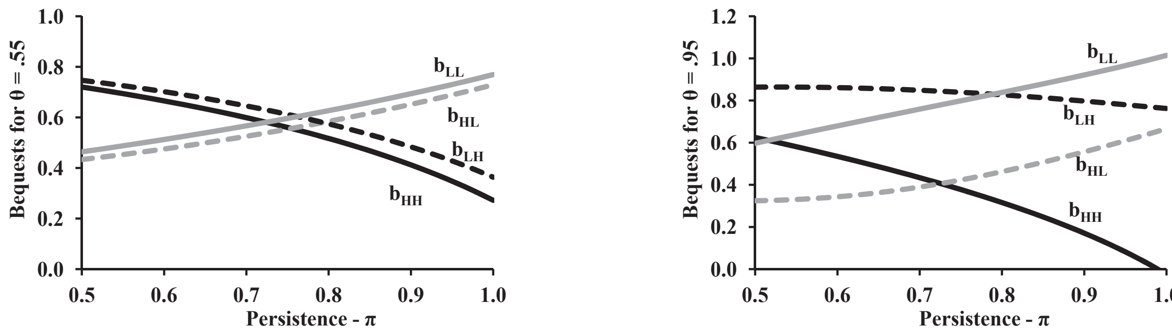
$$\frac{\partial \kappa}{\partial \theta} = \frac{\partial}{\partial \theta}\left[\frac{V(A,\theta)-V_{=}(A,\theta)}{\frac{\partial V_{=}(A,\theta)}{\partial A}}\right] = \frac{\partial V(A,\theta)/\partial\theta - \partial V_{=}(A,\theta)/\partial\theta}{\left[\frac{\partial V_{=}(A,\theta)}{\partial A}\right]^2}.$$

Since the value functions are symmetric around $\theta = .5$, we have that $\partial V(A,\theta)/\partial\theta = \partial V_{=}(A,\theta)/\partial\theta = 0$ and thus $\frac{\partial \kappa}{\partial \theta}|_{\theta=.5} = 0$. □

E.3. **Quantitative Example.** In this section, we illustrate the theorems above. We set the parameter values to the same ones as in section 4.4: $z_H = 2, z_L = 1, A = 1$. The transition matrix is assumed to be symmetric. Figure 3 provides a good representative picture of the effect of signals. Nearly uninformative signals ($\theta = .55$) result in allocations close to the model without signals, with both productivity types doing better when they received a low signal. As the signal θ approaches 1, so that productivity types are nearly public, the HH and LL allocations converge to their perfect information equivalents, while the HL and LH converge to wherever the incentive constraints push them as the other types move to full information.

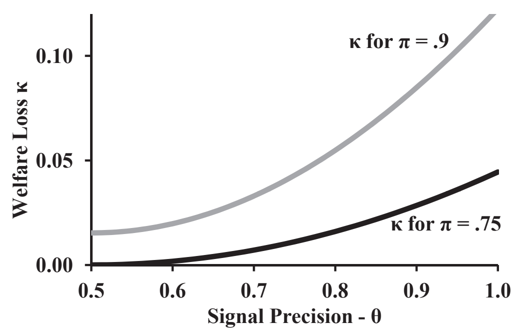
[22] One can also prove theorem E.4 for an alternative approximation: $\kappa = [V(A,\theta) - V_{=}(A,\theta)] / \frac{\partial V_{=}(A,\theta)}{\partial A}$ and the simpler approximation associated with log utility, which we use in section 5.

FIGURE 3. Bequests for Various Levels of Persistence π and Signal Precision θ



(A) Bequests for $\theta = .55$ (B) Bequests for $\theta = .95$

FIGURE 4. Welfare Loss κ for Various Levels of π and θ



We also compute the welfare loss of equal division associated with various values of the signal precision parameter θ and the persistence parameter π. These results are reported for two particular values of π in figure 4. The case of $\pi = .75$ is the one in which bequests are nearly equal with uninformative signals, as in our quantitative analysis of section 5. We add the case of higher persistence $\pi = .9$ and unequal bequests for comparison. We see that the welfare loss of equal division as a function of signal precision θ is flat for small values of θ. This implies that in terms of the welfare loss, the uninformative signal case of $\theta = .5$ is a good approximation of the cases when the signals have some small information content. For this reason, it makes sense to focus on the two extreme cases of public and private information. In addition, we do not see a straightforward way to use the data to determine the precision of the signals that parents get that would enable us to use a model with signals in our quantitative analysis.

References

Altonji, J. G., F. Hayashi, and L. J. Kotlikoff (1997): "Parental Altruism and Inter Vivos Transfers: Theory and Evidence," *Journal of Political Economy*, 105(6), 1121–66.

Arrondel, L., and A. Masson (2002): "Altruism, Exchange or Indirect Reciprocity: What Do the Data on Family Transfers Show?," DELTA Working Paper 2002-18.

Bernheim, B. D., and S. Severinov (2003): "Bequests as Signals: An Explanation for the Equal Division Puzzle," *Journal of Political Economy*, 111(4), 733–64.

Bhatt, V., and M. Ogaki (2012): "Tough Love and Intergenerational Altruism," *International Economic Review*, 35(3), 791–814.

Chami, R. (1996): "King Lear's Dilemma: Precommitment Versus the Last Word," *Economics Letters*, 52(2), 171–76.

Cremer, H., and P. Pestieau (1996): "Bequests as a Heir 'discipline device'," *Journal of Population Economics*, 9(4), 405–14.

——— (1998): "Delaying *Inter Vivos* Transmissions Under Asymmetric Information," *Southern Economic Journal*, 65(2), 322–30.

——— (2001): "Non-linear taxation of bequests, equal sharing rules and the tradeoff between intra- and inter-family inequalities," *Journal of Public Economics*, 79(1), 35–53.

DeNardi, M. (2004): "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, 71(3), 743–768.

Doepke, M., and M. Tertilt (2015): "Asymmetric Information in Couples," Discussion paper.

Dunn, T. A., and J. W. Phillips (1997): "The Timing and Division of Parental Transfers to Children," *Economics Letters*, 54(2), 135–37.

Fernandes, A. (2011): "Altruism, labor supply and redistributive neutrality," *Journal of Population Economics*, 24(4), 1443–1469.

Fudenberg, D., and J. Tirole (1991): *Game Theory*, vol. 1 of *MIT Press Books*. MIT Press.

Heathcote, J., F. Perri, and G. L. Violante (2010): "Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967–2006," *Review of Economic Dynamics*, 13(1), 15–51.

Heathcote, J., K. Storesletten, and G. L. Violante (2010): "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, 118(4), 681–722.

Hochguertel, S., and H. Ohlsson (2009): "Compensatory Inter Vivos Gifts," *Journal of Applied Econometrics*, 24(6), 993–1023.

Horioka, C. Y. (2009): "Do Bequests Increase or Decrease Wealth Inequalities?," *Economics Letters*, 103(1), 23–25.

——— (2014): "Are Americans and Indians More Altruistic than the Japanese and Chinese? Evidence from a New International Survey of Bequest Plans," *Review of Economics of the Household*, 12(3), 411–437.

Kotlikoff, L. J., and A. Razin (1988): "Making Bequests Without Spoiling Children: Bequests as an Implicit Optimal Tax Structure and the Possibility That Altruistic Bequests Are Not Equalizing," NBER Working Paper 2735.

Light, A., and K. McGarry (2004): "Why Parents Play Favorites: Explanations for Unequal Bequests," *American Economic Review*, 94(5), 1669–81.

Lundholm, M., and H. Ohlsson (2000): "Post Mortem Reputation, Compensatory Gifts and Equal Bequests," *Economics Letters*, 68(2), 165–71.

McGarry, K. (1999): "Inter Vivos Transfers and Intended Bequests," *Journal of Public Economics*, 73(3), 321–51.

——— (2016): "Dynamic aspects of family transfers," *Journal of Public Economics*, 137(C), 1–13.

McGarry, K., and R. F. Schoeni (1994): "Transfer Behavior: Measurement and the Redistribution of Resources within the Family," NBER Working Paper 4607.

——— (1995): "Transfer Behavior within the Family: Results from the Asset and Health Dynamics Survey," NBER Working Paper 5099.

Menchik, P. L. (1980): "Primogeniture, Equal Sharing, and the U.S. Distribution of Wealth," *Quarterly Journal of Economics*, 94(2), 299–316.

Nishiyama, S. (2002): "Bequests, Inter Vivos Transfers, and Wealth Distribution," *Review of Economic Dynamics*, 5(4), 892–931.

Nishiyama, S., and K. Smetters (2002): "Ricardian Equivalence with Incomplete Household Risk Sharing," NBER Working Paper 8851.

Ohlsson, H. (2007): "The Equal Division Puzzle - Empirical Evidence on Intergenerational Transfers in Sweden," Working Paper Series 2007:10, Uppsala University, Department of Economics.

Thomas, J., and T. Worrall (1990): "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem," *Journal of Economic Theory*, 51(2), 367–90.

Wilhelm, M. O. (1996): "Bequest Behavior and the Effect of Heirs' Earnings: Testing the Altruistic Model of Bequests," *American Economic Review*, 86(4), 874–92.