ARTICLE IN PRESS

YJETH:4157

ELSEVIER

Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal of Economic Theory ••• (••••) •••–•••

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Optimal contracting with dynastic altruism: Family size and per capita consumption ☆

Roozbeh Hosseini [a,*], Larry E. Jones [b], Ali Shourideh [c]

[a] *Arizona State University, United States*
[b] *University of Minnesota, United States*
[c] *Wharton School, University of Pennsylvania, United States*

Received 9 March 2012; final version received 21 November 2012; accepted 2 April 2013

Abstract

We use a Barro–Becker model of endogenous fertility, in which parents are subject to idiosyncratic shocks that are private information (either to labor productivity or taste for leisure), to study the efficient degree of consumption inequality in the long run. The planner uses the trade-off between family size and future consumption and leisure, to provide incentives for workers to reveal their shocks. We show that in this environment, the optimal dynamic contract no longer features immiseration in consumption. We also discuss the implications of the model on the long run properties of family size in the optimal contract and show that the long run trend in dynasty size can be either positive or negative depending on parameters.

JEL classification: D31; D82; D86; H21; J13

Keywords: Private information; Risk sharing; Long run inequality; Endogenous fertility; Altruism

☆ We are indebted to Alice Schoonbroodt for all of her helpful comments and suggestions at various stages of this project. We would also like to thank Laurence Ales, V.V. Chari, Bob Lucas, Mike Golosov, Greg Kaplan, Chris Phelan, Richard Rogerson, Maxim Troshkin, Aleh Tsyvinski, Ariel Zetlin-Jones and seminar participants at ASU, Columbia, Carnegie Mellon, Iowa, Ohio State, St. Louis Fed, Texas Austin, Western Ontario, Wharton, Yale, 2009 SED summer meeting, 2009 Minnesota Macro Workshop and 2010 Cowles Summer Conference for comments. We would like to also thank two anonymous referees and Christian Hellwig (the editor) for helpful comments and suggestions. Larry E. Jones thanks NSF for financial support, NSF Grant No. SES-0962432.

* Corresponding author.
E-mail addresses: rhossein@asu.edu (R. Hosseini), lej@umn.edu (L.E. Jones), shouride@wharton.upenn.edu (A. Shourideh).

0022-0531/$ – see front matter
http://dx.doi.org/10.1016/j.jet.2013.04.022

Please cite this article in press as: R. Hosseini et al., Optimal contracting with dynastic altruism: Family size and per capita consumption, J. Econ. Theory (2013), http://dx.doi.org/10.1016/j.jet.2013.04.022

1. Introduction

A common feature of efficient contracts in dynamic settings with private information is the presence of a negative drift in consumption of the agent over time as the contract evolves – immiseration (see [16,25,3,22] as examples). This arises due to the desire, by both parties, to use future payoffs as a means to provide current incentives. Having a downward drift on average also makes it cheaper to provide incentives in the future and hence, allows for the provision of better insurance in the short run.

This creates a problem for trying to use these models to study some questions of interest in welfare economics. For example: What is the optimal amount of consumption inequality to balance incentives and social insurance? When the contracting problem features immiseration as a feature of the optimal contract, this question has no well-defined answer – inequality should grow without bound as time goes on and almost every member of society should have a time path of utility that is declining over time.

One solution to this dilemma has been provided in [23,11,12] (and implicitly earlier in [4]) where different periods in the contracting problem correspond to successive generations in a dynastic model. In that setting, these authors show that if the social planner puts higher weight on subsequent generations than parents' themselves do, the optimal contract features mean reversion and hence, there is a non-degenerate stationary distribution over consumption. Implicit in this formulation is the assumption that family size is fixed – each agent is 'replaced' in the subsequent period by exactly one agent – no population growth (or shrinkage) is allowed.

In this paper, we study the form of the optimal incentive contract in a dynamic setting when fertility choice is added to the problem through dynastic altruism. This is a natural extension of the standard model when a period is a generation. There are two, complementary, reasons for making this change. First, as de la Croix and Doepke [9] have shown, including family size in overlapping generations models is crucial to match the empirical relationship between income inequality and mean output growth. Second, it is of interest to know how this extra margin for incentive provision affects the long run features of the optimal contract – e.g., consumption and labor supply.

We show that in this case, the optimal dynamic contract no longer features immiseration even when private and social weights on children are aligned. Indeed, with Barro and Becker's style of dynastic preferences ([5] and [6]) and *i.i.d.* shocks, we show that there is a stationary distribution over consumption, continuation utilities, etc., under a variety of assumptions about the nature of private information and the costs of children.

In some cases – when the cost of children is purely in terms of consumption goods – the result is particularly stark: there is a unique continuation utility that is given to all children independent of the history of shocks in the family. Further, consumption, labor supply and family size are all *i.i.d.* That is, all future incentives for parents are provided through family size and none through children's utility.

In our model, we get an extreme version of the [11] mean reversion result: When the cost of raising children is in terms of consumption goods continuation utility is *i.i.d.* even when social and private discounting factors are identical. The reason for this is that the planner has two instruments to vary future promised utility to parents: the number of children and the promised utility to each child. Because of a homotheticity property of the Barro and Becker dynastic formulation, it turns out that per child future utility is held fixed while the number of children is moved up and down only as a function of the parent's shock to provide incentives. Equivalently, under a natural implementation of the optimal contract, total bequests for the next generation and the number of children are moved up and down proportionally leaving per capita bequests independent of parent's wealth and/or history of past shocks. This is reminiscent of an analogous result in the Barro–Becker model without private information – that per child bequests are independent of parental wealth (see also [5] and [2]) – continues to hold here with private information because both consumption and family size are observable and enter separately from the shock.[1]

We go on to study the long run properties of dynasty size under the optimal contract. We show that family size can either shrink or grow over time depending on the parameters of the model. We study two cases, the first when the rate of return on capital is exogenous, the second where it is endogenous. In the first case, when the discount factor is less than the inverse of the rate of return on capital, we show that family size shrinks over time (i.e., fertility is below replacement) in the optimal contract. Here then, there is immiseration in family size but not in per capita variables. Conversely, if the discount factor is larger than the rate of return on capital (as seems to be true empirically), family size increases without bound as long as the amount of uncertainty about type is not too large. In this case, there is no immiseration in either family size or per capita utility. When the rate of return on capital is endogenous an additional key factor determines the long run properties of population size: the curvature of utility. To address this, we study a special case of the model (goods costs for children and a taste shock to the value of leisure) and show that if either the curvature of utility is small or the level of uncertainty is small, family size grows without bound. Thus, in this case there is neither immiseration in continuation utility nor family size.[2]

Finally, we study the contracting problem in the more realistic case when the costs of children are a mixture of both goods and time. In this case, the properties of the optimal contract are more nuanced – a mix of family size and children's utility are used to provide incentives and the *i.i.d.* property no longer holds. Rather, a limited form of history independence is shown to hold. Formally, continuation utility for all children of a parent with the highest shock is independent of history while for other shocks a similar property holds when parents' continuation utility is sufficiently low. These properties are enough to show that a stationary distribution over per capital variables exists under some additional, technical, assumptions.

Our paper is related to the large literature on dynamic contracting including [16,25,3,23,11,12] (among many others). These papers established the basic way of characterizing the optimal allocation in endowment economies where there is private information. They also showed that, in the long run, inequality increases without bound, i.e. the immiseration result. Phelan [22] shows that this result is robust to many variations in the assumptions of the model. Moreover, Khan and Ravikumar [18] establish numerically that in a production economy, the same result holds and although the economy grows, the detrended distribution of consumption has a negative drift. We contribute to this literature by extending the model to allow for endogenous choice of fertility. We employ the methods developed in the aforementioned papers to analyze this problem.

The paper is organized as follows. In Section 2 we study our basic dynamic contracting problem when family size is included and the costs of children are in terms of goods. In Section 3 we

[1] Since both consumption and family size are observable and they enter separately from the shock, this margin is undistorted by the planner. When time is also an input into child rearing (see Section 4), this homotheticity argument only works for undistorted types – i.e., those with the highest shock.

[2] Hence, overall, the model is consistent with the observation that there has not been immiseration in per capita consumption, but that historically, the wealthy, proxied here by high continuation utility, have had more surviving offspring. See [8] for a thorough discussion of this observation in early England.

study population dynamics implied by the optimal contract. In Section 4, we allow for costs to be in terms of parents' time. Finally, Section 5 concludes.

2. The model and basic results

In this section, we lay out the basic model that we will analyze in an infinite horizon setting and provide some first results concerning the efficient allocation. Our model is an extension of the intergenerational interpretation of [3,11,12] with private information about preferences over leisure. The novelty of our approach is that we include fertility choice in the model with dynastic altruism.

Time is discrete from 0 to ∞. At date 0, the economy is populated by N_{-1} agents. Each agent lives for one period. Agents, when alive, draw labor productivity shocks, can consume and have children. The cost of raising a child is in terms of the consumption good. We assume that it takes a units of consumption good to raise each child. We will report the results with a 'time cost' included in Section 4.[3]

Preferences. Following [5] and [6], we assume that the family head has preferences over own flow utility, u_0, fertility, n_0, and per child utility, U_1 given by:

$$U_0 = u_0 + \beta g(n_0) U_1,$$

in which β is the intergenerational discount factor. Assuming that children are also altruistic toward their children (i.e., the grandchildren of the family head), and so on, we obtain by substitution:

$$U_0 = u_0 + \beta g(n_0) u_1 + \beta^2 g(n_0) g(n_1) u_2 + \cdots.$$

As in [5] and [6], this simplifies when it is assumed that $g(\cdot)$ is iso-elastic, $g(n) = n^\eta$. In this case, it follows that $g(n_0)g(n_1) = g(n_0 n_1)$, etc. Under this assumption, the utility of the family head is given by:

$$U_0 = u_0 + \beta n_0^\eta u_1 + \beta^2 (n_0 n_1)^\eta u_2 + \cdots.$$

To ensure concavity of the utility function we make the following assumption:

Assumption 1. $\eta < 1$.

We will also assume that utility is additively separable in consumption and disutility from work[4] and that the disutility from work depends on both the level of output produced, y, and an idiosyncratic 'shock', θ:

$$u(c, y) = u(c) + h(y, \theta).$$

We assume $h(y, \theta)$ is strictly decreasing and strictly concave in y, and $h_y \equiv \frac{\partial h}{\partial y}$ is strictly increasing in θ.

[3] Goods and time costs of this form are best thought of as short cuts for situations in which parents care directly about consumption of their children (the goods cost) and the amount of time they spend with them (the time cost). This formulation is standard and simplifies the analysis.

[4] This is a standard assumption in dynamic Mirrleesian models, see for example [1,12–15,19,20] among many others.

Written in this way, θ can be interpreted as either a productivity shock or as a shock to the value of leisure. For example, assuming that each agent is endowed with one unit of time to work in each period and assuming that θ is a productivity shock,

$$u(c) + h(y, \theta) = u(c) + \hat{h}\left(1 - \frac{y}{\theta}\right),$$

where $\hat{h}(\cdot)$ is the utility from leisure. Alternatively, an example where θ is a taste shock is:

$$u(c) + h(y, \theta) = u(c) + \frac{1}{\theta}\hat{h}(1 - y).$$

Information. The set of possible types is given by $\Theta = \{\theta_1 < \cdots < \theta_I\}$. We assume that θ_t is drawn from Θ in each period according to the probability distribution $\pi(\cdot)$ and that these draws are independent across time. A history of shocks up to and including period t is represented by $\theta^t = (\theta_0, \ldots, \theta_t)$. We assume that the planner can observe the amount of output for each individual (y), but cannot observe θ. Furthermore, for analytical tractability, we assume that a strong law of large numbers holds so that if an individual has n children, a fraction $\pi(\theta)$ of them is of type θ.

Technology. There is a constant return to scale production function $F(K_t, L_t)$ where K_t is aggregate capital and $L_t = \sum_{\theta^t} \pi(\theta^t) Y_t(\theta^t)$ is aggregate effective labor hours. The initial level of capital is given by K_0. For now, we assume that the rate of return on capital, R, is fixed, i.e., the production function is given by

$$F(K_t, L_t) = RK_t + L_t.$$

We consider the case of endogenous R below when we discuss the implications of the model for population dynamics in Section 3.

Allocation. An *aggregate* allocation is a vector $\{C_t, Y_t, N_t\}_{t=0}^{\infty}$ where $C_t : \Theta^t \to \mathbb{R}_+$, $Y_t : \Theta^t \to \mathbb{R}_+$ and $N_{t+1} : \Theta^t \to \mathbb{R}_+$. Here, $C_t(\theta^t)$ is total consumption of all members of the generation with history θ^t; $Y_t(\theta^t)$ is the total amount of effective labor units that all members of this generation supplies; $N_{t+1}(\theta^t)$ is the population of the generation that immediately follows them. We also define *per capita* consumption and labor supply of each member of the generation with history θ^t as $c_t(\theta^t) \equiv C_t(\theta^t)/N_t(\theta^{t-1})$ and $y_t(\theta^t) \equiv Y_t(\theta^t)/N_t(\theta^{t-1})$. The fertility of each member of the generation with history θ^t is $n_t(\theta^t) \equiv N_{t+1}(\theta^t)/N_t(\theta^{t-1})$.

2.1. The contracting problem

We follow Alvarez [2] and present the problem in terms of aggregate allocations to ensure that the planner's objective is convex. Consider a social planner who offers a contract $\{C_t(\theta^t), Y_t(\theta^t), N_{t+1}(\theta^t)\}_{t=0}^{\infty}$ to each dynasty in period $t = 0$. Each dynasty chooses a reporting strategy $\sigma = \{\sigma_t\}$, which is a sequence of functions $\sigma_t : \Theta^{t+1} \to \Theta$ that maps a history of shocks, θ^t, into a current report $\hat{\theta}_t \in \Theta$. Any strategy σ induces a history of reports $\sigma^t = (\sigma_0(\theta_0), \ldots, \sigma_t(\theta^t))$ with $\sigma^t : \Theta^{t+1} \to \Theta^{t+1}$. We denote the set of all possible reporting strategies by Σ_0. Each person in generation t in a dynasty which has a reporting strategy σ receives $c_t(\sigma_t(\theta^t)) = C_t(\sigma_t(\theta^t))/N_t(\sigma_{t-1}(\theta^{t-1}))$ units of consumption, will supply $y_t(\sigma_t(\theta^t)) = Y_t(\sigma_t(\theta^t))/N_t(\sigma_{t-1}(\theta^{t-1}))$ units of efficiency labor units and will have $n_t(\sigma_t(\theta^t)) = N_{t+1}(\sigma_t(\theta^t))/N_t(\sigma_{t-1}(\theta^{t-1}))$ kids.

Consider the problem of a planner that minimizes the cost of delivering utility W_0 to the first generation parents and has to provide incentives for truthful reporting of types. We write the contracting problem as[5]:

$$\min_{\{C_t(\theta^t), Y_t(\theta^t), N_{t+1}(\theta^t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \sum_{\theta^t} \frac{1}{R^t} \pi(\theta^t) \left[C_t(\theta^t) + aN_{t+1}(\theta^t) - Y_t(\theta^t)\right] \tag{1}$$

subject to

$$\sum_{t=0}^{\infty} \sum_{\theta^t} \beta^t \pi(\theta^t) N_t(\theta^{t-1})^{\eta} \left[u\left(\frac{C_t(\theta^t)}{N_t(\theta^{t-1})}\right) + h\left(\frac{Y_t(\theta^t)}{N_t(\theta^{t-1})}, \theta_t\right)\right] \geqslant W_0 \tag{2}$$

and

$$\begin{aligned} &\sum_{t=0}^{\infty} \sum_{\theta^t} \beta^t \pi(\theta^t) N_t(\theta^{t-1})^{\eta} \left[u\left(\frac{C_t(\theta^t)}{N_t(\theta^{t-1})}\right) + h\left(\frac{Y_t(\theta^t)}{N_t(\theta^{t-1})}, \theta_t\right)\right] \\ &\quad \geqslant \sum_{t=0}^{\infty} \sum_{\theta^t} \beta^t \pi(\theta^t) N_t(\sigma_{t-1}(\theta^{t-1}))^{\eta} \left[u\left(\frac{C_t(\sigma_t(\theta^t))}{N_t(\sigma_{t-1}(\theta^{t-1}))}\right)\right. \\ &\qquad \left. + h\left(\frac{Y_t(\sigma_t(\theta^t))}{N_t(\sigma_{t-1}(\theta^{t-1}))}, \theta_t\right)\right] \quad \text{for all } \sigma \in \Sigma_0. \end{aligned} \tag{3}$$

Using standard arguments for the *i.i.d.* shock case, this problem can be reformulated as a dynamic programming problem with state variable W, the total promised utility to current parents, and N, the current population. Thus, we can show that the above problem is equivalent to the following functional equation:

$$\begin{aligned} V(N, W) = \min_{C(\theta), Y(\theta), N'(\theta), W'(\theta)} \sum_{\theta} \pi(\theta) \Big[& C(\theta) + aN'(\theta) - Y(\theta) \\ & + \frac{1}{R} V\big(N'(\theta), W'(\theta)\big)\Big] \end{aligned} \tag{P}$$

subject to

$$\sum_{\theta} \pi(\theta) \left[N^{\eta}\left\{u\left(\frac{C(\theta)}{N}\right) + h\left(\frac{Y(\theta)}{N}, \theta\right)\right\} + \beta W'(\theta)\right] = W \tag{4}$$

and

$$\begin{aligned} &N^{\eta}\left\{u\left(\frac{C(\theta)}{N}\right) + h\left(\frac{Y(\theta)}{N}, \theta\right)\right\} + \beta W'(\theta) \\ &\quad \geqslant N^{\eta}\left\{u\left(\frac{C(\hat{\theta})}{N}\right) + h\left(\frac{Y(\hat{\theta})}{N}, \theta\right)\right\} + \beta W'(\hat{\theta}) \quad \forall \theta, \hat{\theta} \in \Theta, \end{aligned} \tag{5}$$

[5] This problem may still not be convex. We need extra assumptions on preferences to guarantee concavity of the period utility in consumption, efficiency units of labor and population. Also, it is standard and well known that incentive constraints may not be convex. In Appendix A we give a set of sufficient conditions on preferences that guarantee the convexity of the problem.

where $W'(\theta)$ and $N'(\theta)$ are future promised utility and dynasty population for a parent who receives shock θ in the current period.

It can be shown that the value function from this problem has a homogeneity property and so this functional equation can be equivalently formulated in per capita terms.[6] That is, if we define $\hat{v}(N,w)=\frac{V(N,N^{\eta}w)}{N}$, $\hat{v}(N,w)=v(w)$ will not depend on N and satisfies the following functional equation (note that all choices are, once again, per capita variables):

$$v(w)=\min_{c(\theta),y(\theta),n(\theta),w'(\theta)}\sum_{\theta}\pi(\theta)\left[c(\theta)+an(\theta)-y(\theta)+\frac{1}{R}n(\theta)v\big(w'(\theta)\big)\right] \tag{P$'$}$$

s.t.

$$\sum_{\theta}\pi(\theta)\Big(u\big(c(\theta)\big)+h\big(y(\theta),\theta\big)+\beta n(\theta)^{\eta}w'(\theta)\Big)=w, \tag{6}$$

$$\begin{aligned}&u\big(c(\theta)\big)+h\big(y(\theta),\theta\big)+\beta n(\theta)^{\eta}w'(\theta)\\&\quad\geqslant u\big(c(\hat{\theta})\big)+h\big(y(\hat{\theta}),\theta\big)+\beta n(\hat{\theta})^{\eta}w'(\hat{\theta})\quad\forall\theta,\hat{\theta}\in\Theta.\end{aligned} \tag{7}$$

Here, $n(\theta)$ is the number of children for a parent who receives shock θ and $w'(\theta)$ is the expected promised utility to each of his children.

2.2. *The stationary distribution of per capita variables*

In this section, we show that the per capita variables from the contracting problem of the previous section have a stationary distribution. This is in marked contrast to the situation without fertility where all endogenous variables trend to their lower bounds (see [16,25,3]) – immiseration. Moreover, from any initial condition (N_0, W_0) we jump to the stationary distribution in one period. This follows because of a simple characterization for continuation utilities – a property that we call 'resetting'. This is that the expected promised utility for every child is the same, independent of the history of the parent. Because of this property, it also follows that consumption, hours and efficiency units of labor are *i.i.d.* Thus, when the cost of raising children is in terms of consumption goods and shocks are *i.i.d.*, the optimal contract has an extreme lack of history dependence.

This result depends on one additional (strong) assumption on the solution to the contracting problem from the previous section. This is that the solution can be derived from the first order conditions from problem (P) (and that the solution to the FOC's is unique). Of course, this will follow automatically if the value function $V(\cdot,\cdot)$ is both strictly convex and differentiable. In Appendix A, we show that when θ is a shock to the utility value of leisure, these assumptions on V can be derived from first principles.

Under these assumptions, we can use the first order conditions from problem (P$'$) to characterize the properties of the optimal contract.

Proposition 1. *Assume that $V(\cdot,\cdot)$ is strictly convex and differentiable and that the solution to problem* (P) (*and* (P$'$)) *is interior. Then,*

6 See Online Appendix, Section 4, for a formal proof of this result. Available at http://www.public.asu.edu/~rhossein.

Please cite this article in press as: R. Hosseini et al., Optimal contracting with dynastic altruism: Family size and per capita consumption, J. Econ. Theory (2013), http://dx.doi.org/10.1016/j.jet.2013.04.022

1. *$n(\theta^t)$, $c(\theta^t)$, and $y(\theta^t)$, $t = 1, 2, \dots$ are i.i.d.*
2. *$n(\theta) = n(\theta, w)$, $c(\theta) = c(\theta, w)$, and $y(\theta) = y(\theta, w)$ are all monotone increasing in θ.*[7]

As a first step in proving part 1, we will give a partial characterization of the solution in terms of continuation utilities, w. Let λ and $\mu(\theta, \hat{\theta})$ be the multipliers on constraints (6) and (7), respectively. Taking first order conditions with respect to $n(\theta)$ and $w'(\theta)$ in problem (P′) gives us the following two equations:

$$\pi(\theta)\left(a + \frac{1}{R}v\big(w'(\theta)\big)\right) = \beta\eta n(\theta)^{\eta-1}w'(\theta)\left(\lambda\pi(\theta) + \sum_{\theta\neq\hat{\theta}}\mu(\theta,\hat{\theta}) - \sum_{\hat{\theta}\neq\theta}\mu(\hat{\theta},\theta)\right),$$

$$\pi(\theta)\frac{1}{R}n(\theta)v'\big(w'(\theta)\big) = \beta n(\theta)^{\eta}\left(\lambda\pi(\theta) + \sum_{\theta\neq\hat{\theta}}\mu(\theta,\hat{\theta}) - \sum_{\hat{\theta}\neq\theta}\mu(\hat{\theta},\theta)\right).$$

If we divide these two equations we obtain:

$$\frac{a + v(w'(\theta))/R}{n(\theta)v'(w'(\theta))/R} = \eta\frac{w'(\theta)}{n(\theta)}.$$

The right hand side is the marginal rate of substitution between number of children and promised utility to each child. The left hand side is the marginal rate of transformation. In other words, it is the ratio of marginal costs of having one more child, $Ra + v(w'(\theta))$, and marginal cost of raising promised utility of each child, $n(\theta)v(w'(\theta))$. This condition simplifies to:

$$Ra + v\big(w'(\theta)\big) = \eta w'(\theta)v'\big(w'(\theta)\big). \tag{8}$$

For intuition, imagine that the planner wants to deliver a certain level of utility to the parent in the future. This is done by choosing the best – i.e., cost minimizing – combination of the number of children and promised utility to each child. Increasing the number children by one costs $Ra + v(w'(\theta))$ in terms of the future consumption good. To keep future utility of the parent constant implies that promised utility to each child has to go down by $\eta\frac{w'(\theta)}{n(\theta)}$. This change reduces costs by $n(\theta)v(w'(\theta)) \times \eta\frac{w'(\theta)}{n(\theta)}$. At the optimal contract, these two must exactly offset – i.e., (8) must hold.

From our assumption that the value function in problem (P) is strictly convex and differentiable, it follows that $\eta wv'(w) - v(w)$ is strictly increasing.[8] Thus, there is a unique solution to Eq. (8). Notice that in this equation, the only endogenous variable that appears is the continuation value for the children of an adult who experienced the shock θ in the current period, $w'(\theta)$.

There are two important observations that result from this: First, since promised utility to the parent, w, does not appear in the equation, it follows that promised utility for each child, $w'(\theta)$, is independent of w – formally, $w'(\theta, w) = w'(\theta)$ for all w. Thus, promised utility of future generations is 'reset' each period independent of the history of past shocks to the family. Second, this equation does not depend on θ. From this it follows there is a level of promised utility w^* such that $w'(\theta) = w^*$ for all $\theta \in \Theta$. We state this result as the following lemma:

[7] In the Online Appendix, Section 3, we show that this result holds without assuming convexity and differentiability by analyzing the homotheticity properties of the sequence problem. We focus on the recursive formulation here for ease of exposition.

[8] See Online Appendix, Section 4.

Lemma 1. *Assume that value function in problem (P), V, is strictly convex and continuously differentiable and that the solution is interior. Then, there is a degenerate distribution over per child continuation utilities at w^* where w^* is the unique solution to (8).*

If follows immediately from this lemma that the per capita allocations are only a function of current shock to parents and not on the history of shocks. Formally, let $c(\theta, w)$, $y(\theta, w)$, $n(\theta, w)$ and $w'(\theta, w)$ be policy functions that are solutions to the problem (P′). We construct allocations $c_t(\theta^t)$, $y_t(\theta^t)$ and $n_t(\theta^t)$ using these policy functions. First, recall that $w'(\theta, w) = w^*$ for some w^* and therefore $w_t(\theta^t) = w'(\theta_t, w_{t-1}(\theta^{t-1})) = w^*$ for all θ^t, all t and any initial promised utility w_0. Therefore, the per capita allocations are:

$$\begin{aligned}
c_t(\theta^t) &= c\big(\theta_t, w_{t-1}(\theta^{t-1})\big) = c\big(\theta_t, w^*\big), \\
y_t(\theta^t) &= y\big(\theta_t, w_{t-1}(\theta^{t-1})\big) = y\big(\theta_t, w^*\big), \\
n_t(\theta^t) &= n\big(\theta_t, w_{t-1}(\theta^{t-1})\big) = n\big(\theta_t, w^*\big).
\end{aligned}$$

Hence, the current allocation depends only on the current shock realization, θ_t, which is assumed to be *i.i.d.*

Note that, as a result of Lemma 1, if we know the value of w^* and $v(w^*)$, then the efficient allocation is obtained from the solution to the following static problem:

$$v\big(w^*\big) = \min_{c(\theta), y(\theta), n(\theta)} \sum_\theta \pi(\theta)\left[c(\theta) + an(\theta) - y(\theta) + \frac{1}{R} n(\theta) v\big(w^*\big)\right]$$

s.t.

$$\sum_\theta \pi(\theta)\big(u\big(c(\theta)\big) + h\big(y(\theta), \theta\big) + \beta n(\theta)^\eta w^*\big) = w^*,$$

$$u\big(c(\theta)\big) + h\big(y(\theta), \theta\big) + \beta n(\theta)^\eta w^* \geqslant u\big(c(\hat\theta)\big) + h\big(y(\hat\theta), \theta\big) + \beta n(\hat\theta)^\eta w^* \quad \forall \theta, \hat\theta \in \Theta.$$

This problem is equivalent to a static Mirrlees problem with two consumption goods, c and n. Note that in this problem, the margin of choice between consumption and fertility is undistorted. Therefore, if $c(\theta)$ is monotone in θ, so is $n(\theta)$. If the utility function $u(c) + h(y, \theta)$ satisfies the single crossing property, using standard arguments, we can establish that $y(\theta)$ and $c(\theta)$ are monotone increasing functions of θ. This concludes the proof of Proposition 1.

The results of Proposition 1 and Lemma 1 imply that the long run distribution of future promised utilities is concentrated on w^*. That is, in contrast to the case where fertility is not included (e.g. [16,25,3]), there is not a downward drift in future utilities (at least not in per capita terms). Note that this does not require any particular assumptions on the social discount factor (as in [11,12]).

Unlike standard models of dynamic contracting, where the period t effect of future incentives enters only through promised utility from period $t + 1$ on, in our environment family size is a second instrument that the planner can exploit. What the proposition shows is that the planner will only use fertility to provide for future incentives – per child continuation utility is the same for everyone. This is due to a homotheticity property that is present in the Barro–Becker model of dynastic altruism.[9] To see this, note that, using aggregate variables, we can write $U(C, Y, N, \theta) = N^\eta(u(C/N) + h(Y/N, \theta))$, hence:

[9] This property is also behind the well-known feature of the Barro and Becker model that bequests are independent of wealth, cf., [2].

$$U(\lambda C, \lambda Y, \lambda N, \theta) = (\lambda N)^\eta \left(u\left(\frac{\lambda C}{\lambda N}\right) + h\left(\frac{\lambda Y}{\lambda N}, \theta\right)\right)$$
$$= \lambda^\eta N^\eta \left(u\left(\frac{C}{N}\right) + h\left(\frac{Y}{N}, \theta\right)\right) = \lambda^\eta U(C, Y, N, \theta),$$

i.e., U is homogeneous of degree η. Now consider the problem beginning with Eq. (1). The objective in that problem is linear in C, N and Y. This, together with the homotheticity property discussed above, implies that the aggregate allocations are homogeneous of degree one in W_0 (promised utility to the head of the dynasty). Hence, the ratio of aggregate consumption $C_t(\theta^t)$ (and $Y_t(\theta^t)$), to the dynasty population, $N_t(\theta^{t-1})$, is independent of W_0 after any history θ^t – in particular, this holds after one period. Therefore, the per capita allocations in the first period are independent of W_0. The same intuition applies to future periods.[10]

The fact that the optimal allocations in our setting are *i.i.d.* is an extreme result – it says that there is no history dependence of 'wealth' at all[11] – i.e., promised utility of children is independent of the shock received by the parent. This is something that is unique to the particular example where cost of raising children is in terms of goods and is independent of the type of the parent. To see this, consider two alternatives:

1. First suppose that the cost of children is in terms of goods, but that this cost depends on the type of the parent and denote this cost by $a(\theta)$. In this case, it can be shown that continuation utility per child depends on the type of the parent, but not on the promised utility to the parent. Hence, there is a kind of type specific resetting in the optimal contract. Because of this, there is some history dependence – w' and the other per capita variables are $MA(1)$ – but it is limited.
2. When the costs of children are completely in terms of time, things are even more complicated. There is resetting for the highest shock and at the bottom of the continuation utility distribution only. Thus, there is full history dependence until either the highest shock is realized or continuation utility drifts down to a low level. See Section 4 for a more complete discussion of this case.

In sum, the key equation here is the one relating to the optimal choice of family size:

$$\eta v'(w_{t+1})w_{t+1} - v(w_{t+1}) = R \times MC_t$$

where MC_t is the marginal cost of one additional child including the effects of distortions arising from Incentive Compatibility Constraints. In principle, MC_t can depend on both the history of shocks in the family and the cost structure for raising new children (e.g., time vs. goods costs). Here (i.e., goods cost independent of type) MC_t is independent of both promised utility and current type and hence, continuation utilities are fully 'reset' for all types. More generally, as long as MC_t is uniformly bounded, it follows that the w_{t+1} that solves this equation will also be

[10] The formal proof of resetting given here relies heavily on the homotheticity properties in the Barro–Becker formulation. The main result, that there is a stationary distribution over continuation utilities can be generalized to non-homothetic settings in some cases. See the Online Appendix, Section 2.1.

[11] Note that $v(w)$ is the difference between present value of income and expenditure so it can be thought of as wealth. In this sense, lack of history dependence in promised utility is equivalent to no history dependence in wealth.

bounded. This implies that w's will not drift to extreme values – there will be no immiseration in per capita variables.[12]

3. Population dynamics

In this section, we study the population dynamics that are implied by the optimal contract discussed in the previous section. First, we show that aggregate population is a random walk with drift. We show that under certain conditions it is possible to determine whether the drift is downward – implying immiseration in population – or upward – implying no immiseration in either population or per capita variables.

We provide two types of results. In the first, we show that in certain cases, $N_t \to 0$ *a.s.*, a property we call 'population immiseration'. In this case, although there is a non-degenerate stationary distribution over per capita consumption, the overall size of the dynasty shrinks over time. Because of this, $N_t^{1-\eta}(u(c_t) + h(y_t, \theta_t))$ converges to the lower bound of utility. Thus, although the flow utility for each individual is bounded below, the flow utility from the dynasty head's perspective – which is adjusted by population size – is shrinking almost surely. We show that this is true whenever R is exogenously set to be less than or equal to the inverse of discount factor, i.e., $\beta R \leqslant 1$.

However, the condition $\beta R \leqslant 1$ is not typically an outcome in Barro–Becker type fertility models when the return on capital, R, is endogenous. Indeed, if model parameters are chosen so as to match empirical (and positive) population growth, it must be true that $\beta R > 1$. Even in this case, this does not translate directly into properties of N_t without further work. However, some sufficient conditions can be given – e.g., if the uncertainty in θ is small enough and $\beta R > 1$, then $N_t \to \infty$ *a.s.*

As we can see from this discussion, the determination of R is key in understanding the long run behavior of N. For this reason, we go on to study the population dynamics of the model in the long run when R is determined endogenously. Under certain conditions the population growth rate can be bounded (sometimes above, sometimes below) by what it would be under full information. We specify these conditions in Proposition 5 below.

3.1. The case with a fixed R

To begin, we extend the results in [13] to our environment with endogenous fertility and derive a version of Inverse Euler Equation. We summarize this as a proposition:

Proposition 2. *If, in the optimal contract, consumption is always interior, the optimal allocation satisfies a version of the Inverse Euler Equation*:

$$E\left[\frac{1}{N_{t+1}(\theta^t)^{\eta-1}u'(\frac{C_{t+1}(\theta^{t+1})}{N_{t+1}(\theta^t)})}|\theta^t\right] = \frac{\beta R}{N_t(\theta^{t-1})^{\eta-1}u'(\frac{C_t(\theta^t)}{N_t(\theta^{t-1})})}.$$

Hence, if $\beta R \leqslant 1$, $\frac{1}{N_{t+1}(\theta^t)^{\eta-1}u'(\frac{C_{t+1}(\theta^{t+1})}{N_{t+1}(\theta^t)})}$ *is a non-negative super-martingale.*

[12] Because of this basic intuition, it can also be shown that the 'no immiseration' result holds in other settings as well (e.g., endowment shock models). See the Online Appendix for more discussion on this point in a two-period example and a partial analysis of an example with persistent shocks.

Please cite this article in press as: R. Hosseini et al., Optimal contracting with dynastic altruism: Family size and per capita consumption, J. Econ. Theory (2013), http://dx.doi.org/10.1016/j.jet.2013.04.022

Proof. The proof of this result follows very closely the one given in [13] and is omitted. □

This result must hold independent of the cost and shock structure as long as the utility function is separable between leisure and consumption. When, $\beta R \leqslant 1$, $X_t = N_{t+1}(\theta^t)^{1-\eta}/u'(c_{t+1}(\theta^{t+1}))$ is a non-negative super-martingale. Thus, the martingale convergence theorem implies that there exists a non-negative random variable with finite mean, X_∞, such that $X_t \to X_\infty$ *a.s.* In order to provide incentives for truthful revelation of types, we must have 'spreading' in $N_{t+1}(\theta^t)^{1-\eta}/u'(c_{t+1}(\theta^{t+1}))$ as long as some incentive constraints are binding. Thomas and Worrall [25] have shown that in an environment where incentive constraints are always binding, spreading leads to immiseration. Intuitively, the planner is relying heavily on overall dynasty size to provide incentives and less on continuation utilities. This is something that sets this model apart from the more standard approach with exogenous fertility.

We present the argument in the context of our model as the following corollary.

Corollary 1. *Assume only local downward constraints are binding. If $\beta R \leqslant 1$, then $N_t(\theta^{t-1}) \to 0$ a.s.*

Proof. It follows from the above that $N_t(\theta^{t-1})^{1-\eta}/u'(c_t(\theta^t))$ is a non-negative super-martingale and therefore it has to converge almost surely to a non-negative random variable X_∞. We will show that X_∞ must be zero.

Consider the first order conditions in problem (P′). First note that

$$n(\theta^t)^{1-\eta} v'(w'(\theta^t)) = \frac{\beta R}{u'(c_t(\theta^t))}.$$

Also, for the highest and lowest types (recall that we assume only local downward constraints are binding)

$$n(\theta^{t-1}, \theta_I)^{1-\eta} v'(w'(\theta^{t-1}, \theta_I)) = \beta R\lambda + \beta R \frac{\mu(\theta_I, \theta_{I-1})}{\pi(\theta^{t-1}, \theta_I)} \geqslant \beta R\lambda,$$

$$n(\theta^{t-1}, \theta_1)^{1-\eta} v'(w'(\theta^{t-1}, \theta_1)) = \beta R\lambda - \beta R \frac{\mu(\theta_2, \theta_1)}{\pi(\theta^{t-1}, \theta_1)} \leqslant \beta R\lambda.$$

Note that if either of the constraints (on θ_I or θ_2) is binding we must have

$$n(\theta^{t-1}, \theta_I)^{1-\eta} v'(w'(\theta^{t-1}, \theta_I)) > n(\theta^{t-1}, \theta_1)^{1-\eta} v'(w'(\theta^{t-1}, \theta_1)).$$

Therefore,

$$\frac{1}{u'(c_t(\theta^{t-1}, \theta_I))} > \frac{1}{u'(c_t(\theta^{t-1}, \theta_1))}.$$

Multiplying both sides by $N_t(\theta^{t-1})^{\eta-1}$, we get

$$\frac{N_t(\theta^{t-1})^{1-\eta}}{u'(c_t(\theta^{t-1}, \theta_I))} > \frac{N_t(\theta^{t-1})^{1-\eta}}{u'(c_t(\theta^{t-1}, \theta_1))}.$$

Recall that $c_t(\theta^t)$ is *i.i.d.* and hence, does not converge to any limit. Therefore, the only possibility for $N_t(\theta^{t-1})^{1-\eta}/u'(c_t(\theta^t))$ to converge to a finite non-negative number is $N_t(\theta^{t-1}) \to 0$ *a.s.* □

The fact that $N_{t+1}(\theta^t) \to 0$ *a.s.* does not mean that fertility converges to zero, rather, it means that it is less than replacement (for example, $n_t(\theta^t) < 1$ most of the time). In fact, $N_{t+1}(\theta^t) \to 0$ *a.s.* does not even imply that $E(N_{t+1}(\theta^t)) \to 0$. In computed examples, it can be shown that for certain parameter configurations (with $\beta R = 1$), $E(N_{t+1}(\theta^t)) \to \infty$ even though $N_{t+1}(\theta^t) \to 0$ *a.s.* The reason for this apparent contradiction is that N_t is not bounded – it converges to zero on some sample paths and to ∞ on others.

When $\beta R > 1$, the situation is more complicated. To see this, note that from Proposition 1, the n_t are *i.i.d.* and thus:

$$\ln N_t = \sum_{s=0}^{t-1} \ln\bigl(n(\theta_s)\bigr),$$

where $n(\theta)$ is the policy function for n from the contracting problem. Thus, N_t is a geometric random walk with drift.

In the version of the model with no uncertainty about θ (i.e., $\theta_t = \bar{\theta}$ *a.s.*), and R fixed, it follows from the Euler Equation that $n^{1-\eta} = \beta R$ with probability one and hence $N_t \to \infty$. By continuity, if the amount of uncertainty about θ is small and $\beta R > 1$ it follows that $N_t \to \infty$ *a.s.* We summarize this as a corollary[13]:

Corollary 2. *Suppose that $\beta R > 1$. If the variance of θ is small, then $N_t(\theta^{t-1}) \to \infty$ a.s.*

Proof. Consider the Inverse Euler Equation

$$E\left[\frac{1}{N_{t+1}(\theta^t)^{\eta-1} u'(c_{t+1}(\theta^{t+1}))} | \theta^t\right] = \frac{\beta R}{N_t(\theta^{t-1})^{\eta-1} u'(c_t(\theta^t))},$$

and recall that when the cost of child rearing is only in terms of consumption goods, the $c_t(\theta^t)$, and hence the $\frac{1}{u'(c_t(\theta^t))}$ as well, only depend on the current shock θ_t, which are *i.i.d.* by assumption. Therefore, taking expectations over θ^{t-1} we have:

$$E\left[\frac{1}{N_{t+1}(\theta^t)^{\eta-1}} | \theta^{t-1}\right] E\left[\frac{1}{u'(c_{t+1}(\theta^{t+1}))} | \theta^{t-1}\right] = \frac{\beta R}{N_t(\theta^{t-1})^{\eta-1}} E\left[\frac{1}{u'(c_t(\theta^t))} | \theta^{t-1}\right].$$

Note that

$$E\left[\frac{1}{u'(c_{t+1}(\theta^{t+1}))} | \theta^{t-1}\right] = E\left[\frac{1}{u'(c_t(\theta^t))} | \theta^{t-1}\right] = E\left[\frac{1}{u'(c_t(\theta^t))}\right],$$

and hence,

$$E\left[N_{t+1}\left(\theta^t\right)^{1-\eta} | \theta^{t-1}\right] = \beta R N_t\left(\theta^{t-1}\right)^{1-\eta}. \tag{9}$$

Eq. (9) implies

$$E\left[n(\theta)^{1-\eta}\right] = \beta R > 1.$$

Hence, by continuity, if the distribution of θ is concentrated around it's mean, $E[\theta]$, it follows that $E[\log(n)] > 0$ and hence $N_t \to \infty$ *a.s.* □

---

[13] Even-though the size of a dynasty may grow, it is possible that almost every dynasty grows more slowly than total population therefore vanishing relative to total population in the limit. See the Online Appendix for more discussion.

The discussion above indicates the importance of both R and the dispersion of the shocks in determining the limiting behavior of overall population. Because of this, we next study the case in which R is endogenously determined.

3.2. Population growth rate with endogenous return on capital

As we can see from the previous discussion, whether or not there is immiseration in population depends, in part, on whether βR is larger or smaller than one. Even in the simple version of the Barro–Becker model (i.e., when the distribution of the θ's is degenerate but R is endogenous) whether βR is larger or smaller than one on the balanced growth path depends on the parameters of preferences and technology (rather than on an assumed, exogenous, growth rate). Moreover, this is coincident with whether population growth is positive – $\beta R > 1$ in this case – or negative – $\beta R < 1$.

In this section, we first show that with shocks to the taste for leisure and full information, a very similar relationship holds – there is population growth (both in mean and almost surely) if and only if the parameters of preferences and technology are such that $\beta R > 1$. Moreover, whether or not this holds is independent of the distribution of the shocks.

Furthermore, we can also show that the equilibrium interest rate is the same with full information and private information. Because of this, it follows that if parameters are such that there is positive population growth with full information and the variance of shocks is low enough, population size will converge to infinity almost surely, $N_t \to \infty$ *a.s.* Finally, even when the variance of the shocks is high, we can say something about overall population growth even when *a.s.* statements are not possible. In this case, whether aggregate population grows or shrinks depends not only on βR, but also on the curvature of the utility function. We conclude by a brief discussion of the model with productivity shocks.

3.2.1. The full information economy with endogenous R

We start by extending the model to allow for the endogenous determination of the interest rate when θ is publicly observed. Throughout, we restrict attention to the case where the shock is to the value of leisure. We assume that capital fully depreciates in each period and that the production function is given by:

$$F(K_t, L_t) = AK_t^{\alpha} L_t^{1-\alpha}$$

where $L_t = \sum_{\theta^t} \pi(\theta^t) Y_t(\theta^t)$ is the aggregate quantity of efficiency units of labor provided by households in period t. In this environment, the sequence planning problem is given by[14]:

$$\max \sum_{t=0}^{\infty} \beta^t \sum_{\theta^t} \pi\left(\theta^t\right) N_t\left(\theta^{t-1}\right)^{\eta} \left[u\left(\frac{C_t(\theta^t)}{N_t(\theta^{t-1})}\right) + \frac{1}{\theta_t} h\left(\frac{Y_t(\theta^t)}{N_t(\theta^{t-1})}\right) \right]$$

subject to

$$\sum_{\theta^t} \pi\left(\theta^t\right)\left[C_t\left(\theta^t\right) + aN_{t+1}\left(\theta^t\right)\right] + K_{t+1} \leqslant F\left(K_t, \sum_{\theta^t} \pi\left(\theta^t\right) Y_t\left(\theta^t\right)\right).$$

To simplify the problem, we make the following assumption:

[14] We are writing this problem as utility maximization problem rather than cost minimization problem for simplicity. It is easy to show that the full information economy has nice properties that make these two problems identical.

Assumption 2. Utility over consumption and leisure are CRRA and have the same curvature:

$$u(c)=\frac{c^{1-\sigma}}{1-\sigma} \quad \text{and} \quad h(y,\theta)=\frac{\phi}{\theta}\frac{(1-y)^{1-\sigma}}{1-\sigma},$$

and $\eta = 1-\sigma$.

These assumptions will simplify the algebra below. In this case, period utility becomes:

$$\sum_{\theta^t}\pi\left(\theta^t\right)\left[\frac{C(\theta^t)_t^{1-\sigma}}{1-\sigma}+\frac{1}{\theta_t}\phi\frac{(N_t(\theta^t)-Y_t(\theta^t))^{1-\sigma}}{1-\sigma}\right].$$

From above, it can be seen that this model can be interpreted as a model with two types of capital, $\sum_{\theta^{t-1}}\pi(\theta^{t-1})N_t(\theta^{t-1})$ and K_t, i.e., aggregate population and physical capital. Since the θ's are publicly observed, the planner must be indifferent between investment in population and investment in physical capital. That is, there should not be any arbitrage opportunity between physical capital and population.[15] The return to physical capital is given by $F_K(K_{t+1},L_{t+1})$. A unit increase in population, $N_{t+1}(\theta^t)$, costs the planner a units of consumption good at date t, and has the benefit of increasing household's utility from leisure in the future. Since households labor supply decisions are not distorted, the marginal utility of leisure in terms of consumption goods is simply given by $F_L(K_{t+1},L_{t+1})$. Hence, the return to investing in children is given by $F_L(K_{t+1},L_{t+1})/a$. Therefore, no arbitrage implies that

$$F_K(K_{t+1},L_{t+1})=\frac{F_L(K_{t+1},L_{t+1})}{a}.$$

Using the Cobb–Douglas formulation for F, we have

$$\frac{1-\alpha}{\alpha}\frac{K_{t+1}}{L_{t+1}}=a.$$

Hence, the return on capital is given by

$$R_{FI}=\alpha A\left(\frac{L_{t+1}}{K_{t+1}}\right)^{1-\alpha}=\alpha A\left(\frac{1-\alpha}{a\alpha}\right)^{1-\alpha}. \tag{10}$$

With full information, the Euler Equation holds and hence, the population growth rate on the balanced growth path (where consumption, total hours worked, as well as capital stock all grow at the rate of population growth), γ_{FI}, satisfies:

$$\beta R_{FI}\gamma_{FI}^{\eta-1}=1.$$

Thus,

$$\gamma_{FI}=(\beta R_{FI})^{\frac{1}{1-\eta}}.$$

Note that from above argument, on the balanced growth path, the capital–labor ratio is constant and independent of time. Hence, we can solve for the population growth rate in terms of primitives:

$$\gamma_{FI}=\left(\beta\left[\alpha A\left(\frac{1-\alpha}{a\alpha}\right)^{1-\alpha}\right]\right)^{\frac{1}{1-\eta}}.$$

[15] See [17] for a more detailed study of this model with a representative agent.

From this, we can see that population will be growing whenever the technology is sufficiently productive (i.e., A is large) or children are cheap enough (i.e., a is small) or the representative parent is sufficiently patient (i.e., β is close to 1). Furthermore, since $\beta R_{FI} = \gamma_{FI}^{1-\eta}$, whenever average population growth rate is positive, i.e., $\gamma_{FI} > 1$, $\beta R_{FI} > 1$ and whenever average population growth rate is negative, i.e., $\gamma_{FI} < 1$, $\beta R_{FI} < 1$. We summarize this discussion in the following proposition:

Proposition 3. *The steady state population growth rate is given by* $\gamma_{FI} - 1 = (\beta[\alpha A(\frac{1-\alpha}{a\alpha})^{1-\alpha}])^{\frac{1}{1-\eta}} - 1$. *When* $\gamma_{FI} \geqslant 1$, $\beta R_{FI} \geqslant 1$ *and vice versa.*

3.2.2. The private information economy with endogenous R

We now turn to the problem of characterizing the optimal contract with private information and an endogenous R. As above, we restrict attention to the case of shocks to the taste for leisure. We show that the interest rate in this economy is the same as in the economy with private information. Because of this, population dynamics depend only on the full information interest rate, R_{FI} (as calculated above) and the dispersion of θ.

To do this, we begin by characterizing the solution to the problem of a representative insurance firm taking the rental rate of capital, $R_{PI} = F_K$, and wage, $p_{PI} = F_L$, as given (the subscript refers to the private information economy). The solution to this problem will give rise to aggregate quantities of capital, K_t and labor supply, L_t. We then impose the balanced growth path equilibrium conditions that the resulting shadow prices for capital and labor are constant over time and satisfy $R_t = F_K(K_t, L_t) = R_{PI}$ and $p_t = F_L(K_t, L_t) = p_{PI}$. Finally, it must also be true that W_0 (promised utility to the head of the dynasty) is such that $V_0(N_0, W_0) = r_0 K_0$ – zero profits for the typical insurance company in equilibrium.

Consider the problem of cost minimization of an insurance firm taking $F_K = R_{PI}$, fixed and $F_L = p_{PI}$ fixed (assuming that $\eta = 1 - \sigma$) with private information:

$$V(N, W) = \min_{C(\theta), Y(\theta), N'(\theta), W'(\theta)} \sum_\theta \pi(\theta)\Big[C(\theta) + aN'(\theta) - p_{PI} Y(\theta) + \frac{1}{R_{PI}} V\big(N'(\theta), W'(\theta)\big)\Big] \tag{11}$$

subject to

$$\sum_\theta \pi(\theta)\left[\frac{C(\theta)^{1-\sigma}}{1-\sigma} + \frac{1}{\theta}\phi\frac{(N - Y(\theta))^{1-\sigma}}{1-\sigma} + \beta W(\theta)\right] = W$$

and

$$\frac{C(\theta)^{1-\sigma}}{1-\sigma} + \frac{1}{\theta}\phi\frac{(N - Y(\theta))^{1-\sigma}}{1-\sigma} + \beta W(\theta) \geqslant \frac{C(\hat\theta)^{1-\sigma}}{1-\sigma} + \frac{1}{\theta}\phi\frac{(N - Y(\hat\theta))^{1-\sigma}}{1-\sigma} + \beta W(\hat\theta) \quad \forall \theta, \hat\theta \in \Theta.$$

As in the full information economy, the solution to this problem will satisfy a no-arbitrage condition. This condition can be characterized by focusing on the first order condition with respect to $N'(\theta)$ in the above problem and the Envelope condition with respect to N. These two conditions are given by

$$\frac{1}{R_{PI}} V_N\big(N'(\theta), W'(\theta)\big) = a, \tag{12}$$

$$V_N(N, W) = \lambda \sum_\theta \pi(\theta) \frac{1}{\theta} \phi\big(N - Y(\theta)\big)^{-\sigma} + \sum_{\theta, \hat{\theta}, \hat{\theta} \neq \theta} \mu(\theta, \hat{\theta}) \frac{1}{\theta} \phi\big[\big(N(\theta) - Y(\theta)\big)^{-\sigma} - \big(N - Y(\hat{\theta})\big)^{-\sigma}\big]. \tag{13}$$

The first equation is the first order condition with respect to $N'(\theta)$. The second equation is the Envelope condition with respect to N, where λ is the Lagrange multiplier on the promise keeping constraint and $\mu(\theta, \hat{\theta})$ is the Lagrange multiplier on the incentive constraint. Next, we show that the right hand side of the Envelope condition must be equal to p_{PI} (the efficiency wage rate in the private information economy). To see this, notice that the first order condition with respect to $Y(\theta)$ is given by

$$\pi(\theta) p = \phi\big(N - Y(\theta)\big)^{-\sigma} \left(\pi(\theta) \lambda \frac{1}{\theta} + \sum_{\hat{\theta} \neq \theta} \mu(\theta, \hat{\theta}) \frac{1}{\theta} - \sum_{\theta \neq \hat{\theta}} \mu(\hat{\theta}, \theta) \frac{1}{\hat{\theta}}\right).$$

A summation of the above equations over all θ implies that the right hand side of (13) must be equal to p. Hence, on a balanced growth path, $V_N(N, W) = p_{PI}$.

It follows from this discussion that $\frac{p_{PI}}{a} = R_{PI}$. We showed in the previous section that a similar equality holds in the full information economy – $\frac{p_{FI}}{a} = R_{FI}$. Thus, it must be the case that $a = \frac{p_{FI}}{R_{FI}} = \frac{p_{PI}}{R_{PI}}$. Thus, the $\frac{F_L(K,L)}{F_K(K,L)}$ are equal in the two economies. It follows that the capital–labor ratios $\frac{K_{FI}}{L_{FI}}$ and $\frac{K_{PI}}{L_{PI}}$ in the two economies must be the same. Therefore, the return on capital in both economies must be equal, i.e., $R_{FI} = R_{PI}$. We summarize this discussion in the following proposition:

Proposition 4. *Suppose the source of heterogeneity is a shock to the taste for leisure and Assumption 2 holds. Then, the return to capital in the full information and private information economy are equal, i.e.,* $R_{FI} = R_{PI}$.

As the above analysis shows, introducing private information in the taste shock model does not affect the balanced growth path capital–labor ratio. To see the intuition for this result, consider an infinitesimal increase in family size at date t. The benefit of this increase comes from the increase in utility of the parents from their children's leisure. The consumption equivalent (at date $t+1$) of this increase in welfare is given by $-p_{PIt+1} E_t[MRS_{Y_{t+1}, N_{t+1}}]$, where $MRS_{Y,N}$ is the marginal rate of substitution between income Y and family size N. The cost of this increase in family size is given by $a R_{PIt+1}$ in terms of consumption goods lost (also at date $t+1$). Hence, we must have

$$-p_{PIt+1} E_t[MRS_{Y_{t+1}, N_{t+1}}] = a R_{PIt+1}.$$

Notice that in the model with a taste shock, the marginal rate of substitution between Y_{t+1} and N_{t+1} is equal to -1. This is true independent of whether there is private information or full information, since in (11), N and $Y(\theta)$ always appear together.[16] When there is no informational friction, the above no arbitrage condition implies that the capital–labor ratio is the same as the

[16] In the model with productivity shocks however, this is not necessarily true.

one with private information (since $MRS_{Y,N} = \theta$). In the presence of private information about productivities, however, this marginal rate of substitution changes and it is affected by how tight the incentive constraints are as well as the ratio of the productivities.

We can now characterize population dynamics in the private information economy when the dispersion in θ is low:

Corollary 3. *Suppose that R_{FI} as defined by (10) is greater than β^{-1}, then if the variance of θ is low enough $N_t \to_{a.s.} \infty$. When $R_{FI} \leqslant \beta^{-1}$ then $N_t \to_{a.s.} 0$.*

In reality, population growth has typically been positive in the past. This suggests that $R_{FI} > \beta^{-1}$ is the more relevant case empirically (but does not prove it).

The above corollary, while useful in showing the long run behavior of population, is limited in that it makes some strong assumption about dispersion of shocks. While we cannot characterize the behavior of the distribution of dynasty size, N_t, in the long run, we are able to characterize the behavior of aggregate population, i.e., average N_t across all dynasties. In particular, we can show that the following proposition holds:

Proposition 5. *Suppose the source of heterogeneity is a shock to the taste for leisure and Assumption 2 holds*:

1. *If $\sigma = 1 - \eta < 1$, then $\gamma_{PI} \geqslant \gamma_{FI}$ with equality only if n is degenerate. Thus, in this case, if $\gamma_{FI} > 1$, there is no 'population immiseration'*;
2. *If $\sigma = 1 - \eta > 1$, then $\gamma_{PI} \leqslant \gamma_{FI}$ with equality only if n is degenerate. In this case, if $\gamma_{FI} < 1$, there is 'population immiseration'*;
3. *If $\sigma = 1 - \eta > 1$, and $\gamma_{FI} > 1$ whether $\gamma_{PI} > (<)\ 1$ depends on the distribution of θ. If the amount of uncertainty is small enough, $\gamma_{PI} > 1$.*

This follows from the fact that allocations are *i.i.d.* through an application of Jensen's inequality to the Inverse Euler Equation.

Remark 1. When the shock is to productivity (as opposed to the taste for leisure), the relationship becomes even more complex because it is no longer true that $R_{FI} = R_{PI}$. It can be shown that with productivity shocks, we still have

$$-p_{PIt+1} E_t[MRS_{Y_{t+1},N_{t+1}}] = a R_{PIt+1}.$$

However, contrary to the case with taste shocks, we have $-MRS_{Y_{t+1},N_{t+1}} > 1$. This implies that the capital–labor ratio has to be lower with private information and therefore, $R_{FI} > R_{PI}$. Thus, in this case, it is more difficult to obtain any concrete results. Of course, if population is growing under full info ($\gamma_{FI} > 1$) and uncertainty is not too large, the continuity argument behind Proposition 5 will still hold, but it is difficult to say much more analytically.

3.3. Numerical example

In this section we present a numerical example. Our goal is to show that for reasonable parameter values it is possible to have population growth in the model (not all dynasties vanish). Thus, there is no immiseration either in per capita terms, or in population. In our example the size of all dynasties grow (regardless of their shock history). To prove the propositions in the

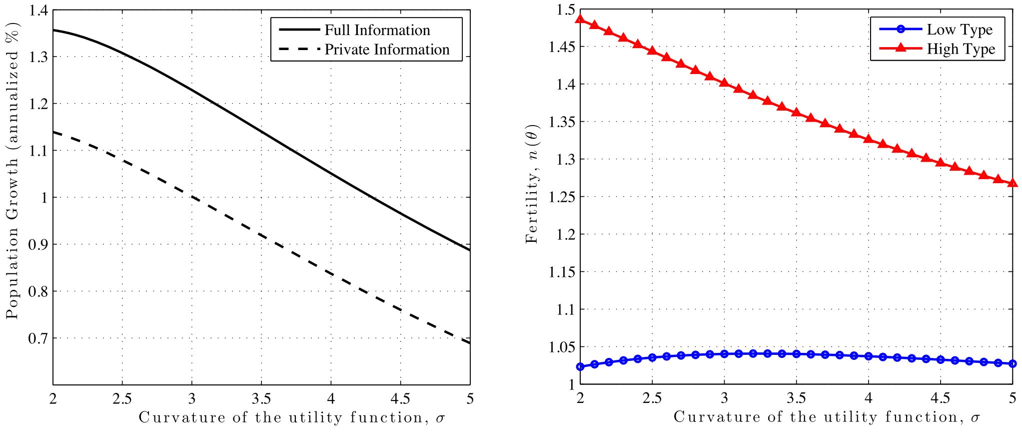

Fig. 1. Panel (a): Population growth under full information (solid line) and private information (dashed line). Panel (b): Fertility for high (red triangle) and low (blue circle) productivity shocks.

previous section we made the restrictive assumption that the private types of parents are taste shock to the value of leisure. In the numerical example we are not constrained by this. Thus, in the example we use the more realistic setting in which the types are productivity shocks. We continue to assume the period utility function is

$$u(c, y; \theta) = \frac{c^{1-\sigma}}{1-\sigma} + \phi \frac{(1-\frac{y}{\theta})^{1-\sigma}}{1-\sigma}$$

where y is efficiency units of labor service and thus, $\frac{y}{\theta}$ is hours worked. We assume production function is of the standard Cobb–Douglas form:

$$F(K, L) = AK^{\alpha}L^{1-\alpha} + (1-\delta)K$$

where δ is the depreciation rate.

Following [17] we choose $\beta = 0.96^{20}$, $\sigma = 3$, $\eta = 1-\sigma$, $\alpha = 0.33$ and $\delta = 1-(1-0.088)^{20}$. We pick $A = 2$ and then choose $\phi = 0.5$, and $a = 1.1$ such that the Frisch Elasticity of labor supply is 0.5 (with parents working on average 40% of their time). Finally, for simplicity, we assume that there are only two shocks, $\theta_L < \theta_H$. We assume that $\pi(\theta_L) = \pi(\theta_H) = 0.5$. We choose the values for θ_L and θ_H so that we match the variance of log-income in the U.S., at 0.6, in the private information version of the model. For this, we use $\theta_L = 5$ and $\theta_H = 15$. For these parameter values, the annual population growth is 1.11% per year in the private information model, close to the actual rate of growth in the U.S. (The growth rate is 1.13% per year under full information.)

The results of this exercise are contained in Fig. 1. In panel (a) we plot the annualized population growth for different values of the curvature parameter (risk aversion, σ). For this range of parameters population always grows under both full and private information. In panel (b), we plot fertility for parents who receive a low shock (blue line) and those who received a high shock (red line). We see that for all values of the utility parameter fertilities are higher than one. Thus, fertility is higher than replacement (i.e., higher than $n = 1.0$), meaning that regardless of history of shocks, all dynasties grow in size.

4. Time costs of kids

In this section, we consider a variant of the model presented in the previous sections by allowing for the possibility that the cost of children is partially (or entirely) in terms of the time of parents. This version of the problem, though more realistic in some ways, does create some technical difficulties. The first of these is that there is no simple transformation of the problem, even in the taste shock case, which guarantees that the contracting problem is convex. Because of this, we will need to make assumptions directly concerning which incentive constraints are binding (and which are slack), and the convexity and differentiability of the value function.

The second difficulty comes about because the version of the 'resetting equation' that we derive only holds in certain circumstances – always for the highest (undistorted) type, and for low enough continuation values for other types (i.e., in the limit). The resetting property at the top has important implications about intergenerational social mobility. In fact, it makes sure that any productive parent will have children with a high level of wealth (if we consider promised utility as a proxy for wealth).

The fact that only a limited version of resetting holds here has two other important implications. First, it implies that the stationary distribution over continuation values is no longer degenerate. Second, per capita allocations are no longer *i.i.d.* Thus, what we take away is that in more complicated and realistic versions of the model, there is still a stationary distribution over the key variables, but their dynamics are more complicated than suggested by the case in which cost of having kids is only in terms of consumption goods.

4.1. The contracting problem

Consider the model studied in Section 2 (that is with a fixed, exogenous, return on capital). We focus on the version of the problem in which θ's are productivity shocks and we add the assumption that each child that is born requires b units of time from the parent as well as a units of the consumption good. It follows that the relevant amount of leisure per parent is given by $1 - \frac{y_t(\theta^t)}{\theta_t} - bn_t(\theta^t)$. Thus, the contracting problem in its recursive form is given by[17]:

$$v(w) = \min_{\{c(\theta), y(\theta), n(\theta), w'(\theta)\}} \sum_{\theta} \pi(\theta)\left[c(\theta) + an(\theta) - y(\theta) + \frac{1}{R} n(\theta) v\big(w'(\theta)\big)\right] \tag{Pt}$$

s.t.

$$\sum_{\theta} \pi(\theta)\left[u\big(c(\theta)\big) + h\left(1 - \frac{y(\theta)}{\theta} - bn(\theta)\right) + \beta n(\theta)^{\eta} w'(\theta)\right] = w,$$

$$u\big(c(\theta)\big) + h\left(1 - \frac{y(\theta)}{\theta} - bn(\theta)\right) + \beta n(\theta)^{\eta} w'(\theta)$$
$$\geqslant u\big(c(\hat\theta)\big) + h\left(1 - \frac{y(\hat\theta)}{\theta} - bn(\hat\theta)\right) + \beta n(\hat\theta)^{\eta} w'(\hat\theta) \quad \forall \hat\theta \leqslant \theta.$$

For technical reasons, we will assume that only downward incentive constraints are binding.

[17] The reduction of the problem from aggregate to per capita and its recursive formulation is identical to the case in which cost of having kids is only in terms of consumption goods. We omit that step.

Please cite this article in press as: R. Hosseini et al., Optimal contracting with dynastic altruism: Family size and per capita consumption, J. Econ. Theory (2013), http://dx.doi.org/10.1016/j.jet.2013.04.022

Including a time cost in the model, introduces further complications. In particular, it is possible for some types never to work. This happens when it is more efficient for an individual to produce goods through the indirect method of having children now with the children working in the future. This will hold when $\theta < \frac{E[\theta]}{bR}$. We will rule out this situation by making the following assumption:

Assumption 3. Assume that for all $\theta \in \Theta$, $\theta > \frac{E[\theta]}{bR}$.

Note that this assumption does not guarantee that hours are always positive. The fact that fertility enters leisure implies that Inada condition on hours will never be satisfied and therefore, it might be optimal for an agent with a very low marginal utility from consumption not to work. However, when marginal utility from consumption is high enough, it is optimal for an agent to work.

Finally, in order to simplify the analysis, we assume that utility functions $u(\cdot)$ and $h(\cdot)$ are both negative and that utility from consumption is unbounded below[18]:

Assumption 4. The utility functions $u(\cdot)$ and $h(\cdot)$ are both negative and $\eta < 0$. Furthermore, $\lim_{c\to 0} u(c) = -\infty$ and $\lim_{m\to 0} h(m) = -\infty$.

4.2. The resetting property

We have shown that children's utility is independent of parent's promised utility when the cost of children is in terms of goods. In this section, we show that this result extends to the time cost version of the model, but only for the highest type, i.e., $\theta = \theta_I$. This can be derived from the first order conditions of the recursive formulation.

The analogue of the resetting equation for this version of the model is:

$$v\big(w'(\theta_I, w)\big) + bR\theta_I + aR = \eta w'(\theta_I, w) v'\big(w'(\theta_I, w)\big). \tag{14}$$

Note that this equation is very similar to Eq. (8). The only added term is $bR\theta_I$, which is the value of time used by the highest type to raise a child in terms of the future consumption good.

In Eq. (14), $w'(\theta_I, w)$ is the per capita promised utility to each child of a parent whose current promised utility is w and whose productivity shock is the highest possible, θ_I. Using the same argument as before, we see that $w'(\theta_I, w)$ is independent of promised utility to the parent, w. That is, $w'(\theta_I, w) = w'(\theta_I, \hat{w})$ for all w, $\hat{w}$. We denote this level of promised utility by $w_I^* = w'(\theta_I, w)$.

The resetting property, in this environment, means that once a parent receives the highest productivity shock, the per capita allocation for her descendants is independent of the parents level of wealth (and all other history) – an extreme version of social mobility holds. We summarize this in a proposition.

Proposition 6. *Assume that v is continuously differentiable and that there is a unique solution to (14). Then, there exists a w_I^* such that the continuation utility for the parent who receives the highest shock has the 'resetting' property, $w'(\theta_I, w) = w_I^*$ for all w.*

[18] An alternative utility configuration would be to assume that u and h are both non-negative and that $\eta > 0$. With sufficient Inada conditions on u and h, the optimal allocation will still be interior and the results below will still hold.

Please cite this article in press as: R. Hosseini et al., Optimal contracting with dynastic altruism: Family size and per capita consumption, J. Econ. Theory (2013), http://dx.doi.org/10.1016/j.jet.2013.04.022

Intuitively, the reason that the resetting property holds here mirrors the argument given above in the case in which cost of having kids is only in terms of consumption goods. That is, since no 'type' wants to pretend to have $\theta = \theta_I$, the allocation for this type is undistorted. Due to the homogeneity properties of the problem, it follows that per capita variables (i.e., continuation utility) are independent of promised utility.

Because of this, it follows that there is no immiseration in this model, under very mild assumptions, in the sense that per capita utility does not converge to its lower bound. To see this, first consider the situation when $n(\theta, w)$ is independent of (θ, w). In this case, from any initial position, the fraction of the population that will be assigned to w_I^* next period is at least $\pi(\theta_I)$. This, by itself, implies that there is no *a.s.* convergence to the lower bound of continuation utilities. When $n(\theta, w)$ is not constant, the argument involves more steps. Assume that $n(\theta, w)$ is bounded above and below – $0 < \underline{n} \leqslant n(w, \theta) \leqslant \bar{n}$. Then, the fraction of descendants being assigned to w_I^* next period is at least $\frac{\pi(\theta_I)\underline{n}}{(1-\pi(\theta_I))\bar{n}}$. Again then, we see that there will not be *a.s.* immiseration. Thus, the key property that must be shown is that $0 < \underline{n} \leqslant n(w, \theta) \leqslant \bar{n}$ for some $\underline{n}$ and $\bar{n}$.

To show this, we show that a version of the resetting property holds when continuation utility is low enough for any type. That is, even as promised utility, w, gets lower and lower, continuation utility, $w'(\theta, w)$, is bounded below. From this, it follows that $n(\theta, w)$ is bounded below in the relevant range of continuation utilities. (That this bound can't be zero follows from promise keeping.) To this end, we show that as $w \to -\infty$, the optimal allocation converges to $c(\theta, w) = 0$, $l(\theta, w) = 1$, $n(\theta, w) = 0$. None of the incentive constraints are binding at this allocation and hence, the optimal allocation has properties similar to those that hold for the highest type. Formally:

Proposition 7. *Suppose that V is continuously differentiable and strictly convex and Assumptions 3 and 4. Then there exists a $\underline{w}_i \in \mathbb{R}$, such that*

$$\lim_{w \to -\infty} w'(w, \theta_i) = \underline{w}_i.$$

Proof. See Appendix B. □

A key step in the proof is to show that when promised utility is sufficiently low, distortions, as measured by the values of the multipliers on the incentive constraints, converge to zero. An important part of the proof uses the fact that the utility from consumption and leisure, $u(\cdot)$ and $h(\cdot)$, is unbounded below. Then, from Proposition 7, it follows that as long as $w'(w, \theta)$ is continuous, w' will be bounded below on any closed set bounded away from 0. We have the following corollary:

Corollary 4. *Suppose that v is continuously differentiable and that $\eta w v'(w) - v(w)$ is strictly increasing. Then for all $\hat{w} < 0$, $w'(w, \theta)$ is bounded below on $(-\infty, \hat{w}]$ – there is a $\underline{w}(\hat{w})$ such that $w'(w, \theta) \geqslant \underline{w}(\hat{w})$ for all $w \leqslant \hat{w}$ and all θ.*

4.3. Stationary distributions

The results from the previous section effectively rule out *a.s.* immiseration as long as fertility, n, is bounded away from 0. This, however, is not quite enough to show that a stationary distribution exists. This is the topic of this section. There are two issues here. First, is there a stationary distribution for continuation utilities and is it non-trivial? Second, because the size of population is endogenous here, it could be growing, or shrinking, in a non-stationary way. Hence, we must also show that the growth rate of population is constant. We deal with this problem in general here.

Consider a measure on continuation utilities in $\mathbb{R}$, Ψ. Then, applying the policy functions to the measure Ψ, gives rise to a new measure on continuation utilities, $T\Psi$:

$$T(\Psi)(A) = \int\limits_{w} \sum_{\theta} \pi(\theta) \mathbf{1}_{\{(\theta,w);\, w'(\theta,w)\in A\}}(w,\theta) n(\theta,w)\, d\Psi(w)$$
$$\forall A\text{: Borel set in } \mathbb{R}. \tag{15}$$

For a given measure over promised values today, Ψ, $T(\Psi)(A)$ is the measure of agents with continuation utility in the set A tomorrow. The overall population growth rate generated by Ψ is given by:

$$\gamma(\Psi) = \frac{\int_{\mathbb{R}} \sum_{\theta} \pi(\theta) n(\theta,w)\, d\Psi(w)}{\Psi(\mathbb{R})} = \frac{T(\Psi)(\mathbb{R})}{\Psi(\mathbb{R})}.$$

Ψ is said to be a stationary distribution if:

$$T(\Psi) = \gamma(\Psi)\cdot\Psi.$$

Thus, in a stationary distribution, the distribution of per capita continuation utility is constant while population grows at rate $(\gamma(\Psi) - 1) \times 100$ percent per period.

To show that there is a stationary distribution, we will show that the mapping $\Psi \to \frac{T(\Psi)}{\gamma(\Psi)}$ is a well-defined and continuous function on the set of probability measures on a compact set of possible continuation utilities. To do this, we need to construct a compact set of continuation utilities, $[\underline{w}, \overline{w}]$, such that:

1. For all $w \in [\underline{w}, \overline{w}]$, there is a solution to problem (Pt);
2. For all $w \in [\underline{w}, \overline{w}]$, $w'(w,\theta) \in [\underline{w}, \overline{w}]$;
3. $n(w,\theta)$ and $w'(w,\theta)$ are continuous functions of w on $[\underline{w}, \overline{w}]$;
4. $\gamma(\Psi)$ is bounded away from zero for the distributions concentrated on $[\underline{w}, \overline{w}]$.

The construction of the interval $[\underline{w}, \overline{w}]$, is relegated to Appendix C.

Now, we are ready to prove our main result about the existence of a stationary distribution. Let $M([\underline{w}, \overline{w}])$ be the set of regular probability measures on $[\underline{w}, \overline{w}]$.

Theorem 1. *Assume that for all $w \in [\underline{w}, \overline{w}]$, there is a solution to the problem* (Pt) *and that it is unique. Then there exists a measure $\Psi^* \in M([\underline{w}, \overline{w}])$ such that $T(\Psi^*) = \gamma(\Psi^*)\cdot\Psi^*$.*

Proof. See Appendix C. □

This theorem immediately implies that there is a stationary distribution for per capita consumption, labor supply and fertility. Moreover, since promised utility is fluctuating in a bounded set, per capita consumption has the same property. This shows that there is no immiseration in per capita variables. However, since full resetting does not hold, it follows that per capita allocations will not be *i.i.d.* in general.

Please cite this article in press as: R. Hosseini et al., Optimal contracting with dynastic altruism: Family size and per capita consumption, J. Econ. Theory (2013), http://dx.doi.org/10.1016/j.jet.2013.04.022

5. Conclusion

In this paper, we have provided results concerning the form of the optimal contract in a dynamic setting with dynastic altruism and fertility choice. We show that in this case the optimal contract no longer features in immiseration in consumption under a wide variety of assumptions about the costs of children and the nature of private information. Indeed, with Barro–Becker style preferences, goods costs for children and *i.i.d.* shocks, we have shown that there is a unique continuation value for children. Because of this feature, it follows that consumption, labor supply and fertility are *i.i.d.* in this case.

We have gone on to show that there may or may not be immiseration in dynasty size, depending on the specifics of the problem. Examples are given where each is shown to hold.

The model can also be used to study the form of the optimal subsidies to family size, the efficient provision of incentives for bequests and how these two interact with labor income taxes. Some preliminary work has been done along these lines (see the Online Appendix for details). Our preliminary work, suggests that whether children should be taxed or subsidized depends on the details of how incentives are provided over time as well as the nature of the cost of children. For example, one can show that with only goods cost of child rearing it may be optimal to subsidize children while introducing a time cost moves the optimal tax system towards taxing children. These are interesting issues and questions to explore and are left for future research.

Appendix A. The special case of shocks to the taste for leisure

There is a technical issue with the statement of Proposition 1. This is that in the statement of the proposition, we are making assumptions about properties of the value function (that it is convex and differentiable), an endogenous object. Here, we show that there is at least one case in which we can dispense with these two extra assumptions. This is true when θ is the taste shock to the value leisure:

$$u(c) + h(y, \theta) = u(c) + \frac{1}{\theta} v(1 - y).$$

In this case, we will show that there is a simple transformation so that the constraint set in the minimization problem (P) is convex and the objective function is strictly convex. Let's define:

$$U(\theta) = N^{\eta} u\left(\frac{C(\theta)}{N}\right), \qquad H(\theta) = N^{\eta} h\left(1 - \frac{Y(\theta)}{N}\right).$$

For a given $U(\theta)$ and $H(\theta)$, we have

$$C(\theta) = N u^{-1}\bigl(N^{-\eta} U(\theta)\bigr).$$

Also, $h^{-1}(N^{-\eta} H(\theta)) = 1 - Y(\theta)/N$, therefore

$$Y(\theta) = N - N h^{-1}\bigl(N^{-\eta} H(\theta)\bigr).$$

Under these transformations, the promise keeping constraint (4) becomes:

$$\sum_{\theta} \pi(\theta)\left[U(\theta) + \frac{1}{\theta} H(\theta) + \beta W(\theta)\right] = W,$$

and, the incentive constraints (5) become:

$$U(\theta)+\frac{1}{\theta}H(\theta)+\beta W(\theta)\geqslant U(\hat{\theta})+\frac{1}{\theta}H(\hat{\theta})+\beta W(\hat{\theta})\quad \forall\theta,\hat{\theta}\in\Theta.$$

Thus, the promise keeping and incentive constraints are linear in the transformed variables.

Next, we will show that the value function in problem (P) is strictly convex. If we rewrite the objective in terms of transformed variables we obtain:

$$\begin{aligned}&\sum_{\theta}\pi(\theta)\left[C(\theta)+aN'(\theta)-Y(\theta)+\frac{1}{R}V\big(N'(\theta),W'(\theta)\big)\right]\\&\quad=\sum_{\theta}\pi(\theta)\left[Nu^{-1}\big(N^{-\eta}U(\theta)\big)+aN'(\theta)+Nh^{-1}\big(N^{-\eta}H(\theta)\big)\right.\\&\qquad\left.-N+\frac{1}{R}V\big(N'(\theta),W'(\theta)\big)\right].\end{aligned}$$

For this objective function to be strictly convex, it is necessary that the two functions $Nh^{-1}(N^{-\eta}H)$ and $Nu^{-1}(N^{-\eta}U)$ be strictly convex functions. That is what we assume from now on.

We can rewrite the transformed version of the problem (P) as:

$$\begin{aligned}V(N,W)=\min_{U(\theta),H(\theta),N'(\theta),W'(\theta)}&\sum_{\theta}\pi(\theta)\left[Nu^{-1}\big(N^{-\eta}U(\theta)\big)+aN'(\theta)\right.\\&\left.-N\big(1-h^{-1}\big(N^{-\eta}H(\theta)\big)\big)+\frac{1}{R}V\big(N'(\theta),W'(\theta)\big)\right]\end{aligned}\tag{16}$$

subject to

$$\sum_{\theta}\pi(\theta)\left[U(\theta)+\frac{1}{\theta}H(\theta)+\beta W(\theta)\right]=W,$$

and,

$$U(\theta)+\frac{1}{\theta}H(\theta)+\beta W(\theta)\geqslant U(\hat{\theta})+\frac{1}{\theta}H(\hat{\theta})+\beta W(\hat{\theta})\quad \forall\theta,\hat{\theta}\in\Theta.$$

We state and prove the main result of this section in the following proposition:

Proposition 8. *Suppose that $Nh^{-1}(N^{-\eta}H)$ and $Nu^{-1}(N^{-\eta}U)$ are strictly convex functions of (N,H) and (N,U) and that for some convex set A, and all $(N,W)\in A$, there is a solution to the sequence problem (1). Then, the solution to the functional equation (16) is strictly convex on A and differentiable on the interior of A. Moreover, there is a function, $v:\mathbb{R}\to\mathbb{R}$, such that $V(N,W)=Nv(N^{-\eta}W)$.*

Proof. To show that $V(\cdot,\cdot)$ is strictly convex, we use the sequence problem version of the recursive problem (16) and use the principle of optimality to show that $V(\cdot,\cdot)$ is convex. Consider the following sequence problem

$$\begin{aligned}\min_{U_t,H_t,N_{t+1},W_{t+1}}&\sum_{t=0}^{\infty}\left(\frac{1}{R}\right)^t\sum_{\theta^t}\pi(\theta^t)\big[N_t(\theta^{t-1})u^{-1}\big(N_t(\theta^{t-1})^{-\eta}U_t(\theta^t)\big)+aN_{t+1}(\theta^t)\\&-N_t(\theta^{t-1})\big(1-h^{-1}\big(N_t(\theta^{t-1})^{-\eta}H_t(\theta^t)\big)\big)\big]\end{aligned}$$

Please cite this article in press as: R. Hosseini et al., Optimal contracting with dynastic altruism: Family size and per capita consumption, J. Econ. Theory (2013), http://dx.doi.org/10.1016/j.jet.2013.04.022

subject to

$$\sum_{\theta^t \geqslant \theta^{t-1}} \pi\left(\theta^t|\theta^{t-1}\right)\left[U_t\left(\theta^t\right)+\frac{1}{\theta_t}H_t\left(\theta^t\right)+\beta W_{t+1}\left(\theta^t\right)\right]=W_t\left(\theta^{t-1}\right) \quad \forall\theta^{t-1},$$

$$U_t\left(\theta^t\right)+\frac{1}{\theta_t}H_t\left(\theta^t\right)+\beta W_{t+1}\left(\theta^t\right)$$

$$\geqslant U_t\left(\theta^{t-1},\hat{\theta}_t\right)+\frac{1}{\theta_t}H_t\left(\theta^{t-1},\hat{\theta}_t\right)+\beta W_{t+1}\left(\theta^{t-1},\hat{\theta}_t\right) \quad \forall\theta^t,\hat{\theta}_t,$$

given $W_0 = W,\ N_0 = N$.

By the principle of optimality, the value associated with the above problem is given by $V(N, W)$. Now, consider two possible initial conditions $(N^1, W^1) \neq (N^2, W^2)$. Associated with each of these is a set of allocations that solve the above maximization problem given by $\{U_t^1(\theta^t), H_t^1(\theta^t), N_{t+1}^1(\theta^t), W_{t+1}^1(\theta^t)\}$ and $\{U_t^2(\theta^t), H_t^2(\theta^t), N_{t+1}^2(\theta^t), W_{t+1}^2(\theta^t)\}$, respectively. Let $(\tilde{N}, \tilde{W}) = \lambda(N^1, W^1) + (1-\lambda)(N^1, W^2)$ and define the following allocations

$$\tilde{U}_t\left(\theta^t\right)=\lambda U_t^1\left(\theta^t\right)+(1-\lambda)U_t^2\left(\theta^t\right),$$
$$\tilde{H}_t\left(\theta^t\right)=\lambda H_t^1\left(\theta^t\right)+(1-\lambda)H_t^2\left(\theta^t\right),$$
$$\tilde{N}_{t+1}\left(\theta^t\right)=\lambda N_{t+1}^1\left(\theta^t\right)+(1-\lambda)N_{t+1}^2\left(\theta^t\right),$$
$$\tilde{W}_{t+1}\left(\theta^t\right)=\lambda W_{t+1}^1\left(\theta^t\right)+(1-\lambda)W_{t+1}^2\left(\theta^t\right).$$

Note that, since the constraint set is linear, the above allocation satisfies promise keeping and incentive compatibility when the initial state is given by $(\tilde{N}, \tilde{W})$. Let $\tilde{V}$ denote the value of objective function at the allocation constructed above. Then, we must have $V(\tilde{N}, \tilde{W}) \leqslant \tilde{V}$. Note also that, since the objective function is convex

$$\tilde{V} \leqslant \lambda V\left(N^1, W^1\right)+(1-\lambda)V\left(N^2, W^2\right).$$

Moreover, since the period cost functions, $Nh^{-1}(N^{-\eta}H)$ and $Nu^{-1}(N^{-\eta}U)$, are strictly convex, the above inequality is strict (provided that for some θ^t, either $U_t^1(\theta^t) \neq U_t^2(\theta^t)$, or $H_t^1(\theta^t) \neq H_t^2(\theta^t)$, or $N_{t+1}^1(\theta^t) \neq N_{t+1}^2(\theta^t)$). Hence, the function $V(\cdot,\cdot)$ must be strictly convex. The same proof can be used to show that the optimal allocation is unique.

In order to show that the value function is differentiable, we use a technique similar to [7]. To do so, we also need to use results developed by Milgrom and Segal [21]. From the argument above, we know that the value function is strictly convex and the policy functions are unique. Now consider an interior point (N, W) and let the optimal policy at this point be given by $\{U^*(\theta), H^*(\theta), N'^*(\theta), W'^*(\theta)\}$. For any $(\tilde{N}, \tilde{W})$ close to (N, W), define the following function

$$G(\tilde{N}, \tilde{W}) = \min_{H(\theta),U(\theta)} \sum_{\theta}\pi(\theta)\Big[\tilde{N}u^{-1}\left(\tilde{N}^{-\eta}U(\theta)\right)+aN'^*(\theta)-\tilde{N}\left(1-h^{-1}\left(\tilde{N}^{-\eta}H(\theta)\right)\right)$$
$$+\frac{1}{R}V\left(N'^*(\theta), W'^*(\theta)\right)\Big]$$

subject to

$$\sum_{\theta}\pi(\theta)\left[U(\theta)+\frac{1}{\theta}H(\theta)+\beta W'^*(\theta)\right]=\tilde{W}$$

Please cite this article in press as: R. Hosseini et al., Optimal contracting with dynastic altruism: Family size and per capita consumption, J. Econ. Theory (2013), http://dx.doi.org/10.1016/j.jet.2013.04.022

and

$$U(\theta)+\frac{1}{\theta}H(\theta)+\beta W'^{*}(\theta)\geqslant U(\hat{\theta})+\frac{1}{\theta}H(\hat{\theta})+\beta W'^{*}(\hat{\theta})\quad\forall\theta,\hat{\theta}\in\Theta.$$

Note that since we have fixed future policies, we must have that $G(\tilde{N},\tilde{W})\leqslant V(\tilde{N},\tilde{W})$ with equality at $(N,W)=(\tilde{N},\tilde{W})$. Since the cost functions are strictly convex and the set of constraints is linear, the solution to the above minimization problem is unique. Moreover, by the theorem of maximum the solution is continuous in $(\tilde{N},\tilde{W})$. Hence by applying Theorem 5 and Corollary 4 in [21], $G(\cdot,\cdot)$ is differentiable. Furthermore, since the cost functions are strictly convex, we must have that $G(\cdot,\cdot)$ is strictly convex. Hence, by Lemma 1 in [7], $V(\cdot,\cdot)$ must be differentiable at (N,W).

In order to show the final property of the value function, we show that the transformation associated with the functional equation (16) preserves that property. That is, if $T(\cdot)$ is the transformation associated with the functional equation (16) and the set S is defined as the following:

$$S=\{V(N,W);\ \exists v:\mathbb{R}\to\mathbb{R},\ V(N,W)=Nv(N^{-\eta}W)\},$$

then $T(S)\subseteq S$. Suppose that $V\in S$, then $\exists v$ such that $V(N,W)=Nv(N^{-\eta}W)$, and

$$TV(N,W)=\min_{H(\theta),U(\theta),N'(\theta),W'(\theta)}\sum_{\theta}\pi(\theta)\Big[Nu^{-1}\big(N^{-\eta}U(\theta)\big)+aN'(\theta)\\ -N\big(1-h^{-1}\big(N^{-\eta}H(\theta)\big)\big)+\frac{1}{R}V\big(N'(\theta),W'(\theta)\big)\Big]$$

subject to

$$\sum_{\theta}\pi(\theta)\left[U(\theta)+\frac{1}{\theta}H(\theta)+\beta W(\theta)\right]=W,$$

and,

$$U(\theta)+\frac{1}{\theta}H(\theta)+\beta W(\theta)\geqslant U(\hat{\theta})+\frac{1}{\theta}H(\hat{\theta})+\beta W(\hat{\theta})\quad\forall\theta,\hat{\theta}\in\Theta.$$

Now, let $\hat{V}(N,N^{-\eta}W)=N^{-1}TV(N,W)$ and define $h(\theta)=N^{-\eta}H(\theta)$, $u(\theta)=N^{-\eta}U(\theta)$, $n(\theta)=N^{-1}N'(\theta)$ and $w'(\theta)=N(\theta)^{-\eta}W'(\theta)$. Then the above maximization can be written as

$$T\hat{V}(N,W)=\min_{H(\theta),U(\theta),N'(\theta),W'(\theta)}N\sum_{\theta}\pi(\theta)\Big[u^{-1}\big(u(\theta)\big)+an(\theta)-\big(1-h^{-1}\big(h(\theta)\big)\big)\\ +\frac{1}{R}n(\theta)v\big(w'(\theta)\big)\Big]$$

subject to

$$\sum_{\theta}\pi(\theta)\left[u(\theta)+\frac{1}{\theta}h(\theta)+\beta n(\theta)^{\eta}w'(\theta)\right]=N^{-\eta}W,$$

and,

$$u(\theta)+\frac{1}{\theta}h(\theta)+\beta n(\theta)^{\eta}w'(\theta)\geqslant u(\hat{\theta})+\frac{1}{\theta}h(\hat{\theta})+\beta n(\hat{\theta})^{\eta}w'(\hat{\theta})\quad\forall\theta,\hat{\theta}\in\Theta.$$

Clearly, $\hat{V}(N,N^{-\eta}W)$ is independent of N and solely depends on $N^{-\eta}W$. Hence $TV\in S$. □

From this proposition, it follows that the result of Proposition 1 will hold as long as the solution to the contracting problem is interior. In order to guarantee that this holds we will make the following standard assumptions[19]:

Assumption 5. Period utility function satisfies the following:

1. u is unbounded below and negative and $\eta < 0$.
2. h is unbounded below and negative.

Under these conditions, the solution will be interior and thus, we have:

Corollary 5. *Under Assumption 5*:

1. *The solution to problem* (P′) *has the property that* $w'(\theta, w) = w^*$ *for all* θ *and* w. *Hence, there is a stationary distribution over continuation values* (*a point mass at* w^*).
2. *The allocations* $n(\theta^t)$, $c(\theta^t)$, *and* $y(\theta^t)$ *are i.i.d. and monotone increasing in current shock* θ.[20]

Appendix B. Proof of Proposition 7

We show this claim in two steps:

- Step 1: We show that leisure, consumption and fertility converge to zero as promised utility converges to $-\infty$ – this is the content of Lemma 2 below;
- Step 2: Using step 1, we show that a limiting version of the resetting property, with type specific resetting values, must hold as $w \to -\infty$.

We first prove the following lemma:

Lemma 2. *Suppose that the value function in* (Pt) *is differentiable and that Assumptions 3 and 4 hold. Then, the value function and the policy functions satisfy the following properties*:

$$\lim_{w \to -\infty} v(w) = -\sum_{\theta_i} \pi(\theta_i)\theta_i,$$

$$\lim_{w \to -\infty} c(\theta_i, w) = 0,$$

$$\lim_{w \to -\infty} n(\theta_i, w) = 0,$$

$$\lim_{w \to -\infty} l(\theta_i, w) = 1.$$

[19] In addition to this assumption, we have made two extra assumptions on u and h in the statement of Proposition 8. It can be checked that all of these assumptions are satisfied for the following specification: $u = \frac{c^{1-\sigma_c}}{1-\sigma_c}$, $h(\ell) = \psi \frac{\ell^{1-\sigma_l}}{1-\sigma_l}$ with $1 < \sigma_l, \sigma_c < 1 - \eta$.

[20] Assumption 1 is only used to prove, from first principles, that the optimal allocation is interior. Alternative methods are possible. For example, if $u(0) = 0$, $u'(0) = \infty$ and $\eta > 0$, this will also hold.

Proof. Consider the following set of functions:

$$S = \left\{ \hat{v} \colon \ \hat{v} \in C(\mathbb{R}_-), \ \hat{v} \text{ is weakly increasing}, \ \lim_{w \to -\infty} \hat{v}(w) = -\sum_{\theta_i} \pi(\theta_i)\theta_i \right\}.$$

Define the following mapping on S as

$$T\hat{v}(w) = \min \sum_{\theta_i} \pi(\theta_i) \left[c(\theta_i) + an(\theta_i) - \theta l(\theta_i) + \frac{1}{R} n(\theta_i) \hat{v}\big(w'(\theta_i)\big) \right]$$

s.t.

$$\sum_{\theta_i} \pi(\theta_i) \big[u\big(c(\theta_i)\big) + h\big(1 - l(\theta_i) - bn(\theta_i)\big) + \beta n(\theta_i)^{\eta} w'(\theta_i) \big] \geqslant w,$$

$$u\big(c(\theta_i)\big) + h\big(1 - l(\theta_i) - bn(\theta_i)\big) + \beta n(\theta_i)^{\eta} w'(\theta_i)$$
$$\geqslant u\big(c(\theta_j)\big) + h\left(1 - \frac{\theta_j l(\theta_j)}{\theta_i} - bn(\theta_j)\right) + \beta n(\theta_j)^{\eta} w'(\theta_j) \quad \forall i > j,$$
$$l(\theta_i) + bn(\theta_i) \leqslant 1,$$
$$c(\theta_i), l(\theta_i), n(\theta_i) \geqslant 0.$$

We first show that the solution to the above program has the claimed property for the policy function and that $T\hat{v}$ satisfies the claimed property. Then, since S is closed and T preserves S, by the Contraction Mapping Theorem we have that the fixed point of T belongs to S.

Now, suppose the claim about the policy function for fertility, does not hold. Then, there exists a sequence $w_k \to -\infty$ such that for some θ_i, $n(w_k, \theta_i) \to \bar{n}_i > 0$. For each $j \neq i$, define $\bar{n}_j = \limsup_{k \to \infty} n(w_k, \theta_j)$. Then, we must have

$$\liminf_{n \to \infty} T\hat{v}(w_k) \geqslant \sum_{\theta_j} \pi(\theta_j) \left[a\bar{n}_j - \theta_j(1 - b\bar{n}_j) + \frac{1}{R} \bar{n}_j \left[-\sum_{\theta_k} \pi(\theta_k)\theta_k \right] \right].$$

Note that by Assumption 3, we have

$$b\theta_j > \frac{1}{R} \sum_{\theta_k} \pi(\theta_k)\theta_k \quad \forall j$$

and therefore, if $\bar{n}_j \geqslant 0$, we must have

$$a\bar{n}_j - \theta_j + b\bar{n}_j\theta_j - \frac{1}{R}\bar{n}_j \sum_{\theta_k} \pi(\theta_k)\theta_k \geqslant -\theta_j$$

with equality only if $\bar{n}_j = 0$. This implies that

$$\liminf_{n \to \infty} T\hat{v}(w_k) > -\sum_{\theta_j} \pi(\theta_j)\theta_j$$

since $\bar{n}_i > 0$. Now, we construct a sequence of allocations and show that the above allocation cannot be an optimal. Consider a sequence of numbers ϵ_m that converges to zero. Define

$$c_m(\theta_i) = u^{-1}\left(-\epsilon_m^\eta\right),$$
$$n_m(\theta_i) = (I-i)^{\frac{1}{\eta}}\epsilon_m,$$
$$w'_m(\theta_i) = \tilde{w} < 0,$$

for some $\tilde{w} < 0$.

Since, h is unbounded below and that the utility of deviation is bounded away from $-\infty$, it is possible to construct a sequence for l_m that converges to 1. To see this, let $l_m(\theta_i)$ be defined such that

$$h\left(1 - l_m(\theta_i) - bn_m(\theta_i)\right) = \frac{1}{2}\tilde{w}\epsilon_m^\eta.$$

Hence, we have

$$u\left(c_m(\theta_j)\right) + h\left(1 - l_m(\theta_j) - bn_m(\theta_j)\right) + \beta n_m(\theta_j)^\eta w'_m(\theta_j) \\ - u\left(c_m(\theta_i)\right) - \beta n_m(\theta_i)^\eta w'_m(\theta_i) = \tilde{w}\epsilon_m^\eta\left(i - j + \frac{1}{2}\right)$$

which converges to ∞ as m converges to ∞. Moreover, by definition, $l_m(\theta_i)$ converges to 1 and $n_m(\theta_i)$ converges to 0. Therefore, the deviation value for leisure, $h(1 - \frac{\theta_i l_m(\theta_i)}{\theta_j} - bn_m(\theta_i))$, converges to $h(1 - \frac{\theta_i}{\theta_j})$. This implies that for m large enough

$$u\left(c_m(\theta_j)\right) + h\left(1 - l_m(\theta_j) - bn_m(\theta_j)\right) + \beta n_m(\theta_j)^\eta w'_m(\theta_j) \\ - u\left(c_m(\theta_i)\right) - \beta n_m(\theta_i)^\eta w'_m(\theta_i) \geqslant h\left(1 - \frac{\theta_i l_m(\theta_i)}{\theta_j} - bn_m(\theta_i)\right)$$

and for m large enough the allocation is incentive compatible. Therefore, the utility from the constructed allocation is:

$$\hat{w}_m = \epsilon_m^\eta\left[-1 + \beta\sum_j (I-j)^\eta\tilde{w}\right] + \sum_{\theta_j}\pi(\theta_j)\left[h\left(1 - l_m(\theta_j) - bn_m(\theta_j)\right)\right].$$

It is clear that the $\hat{w}_m$'s converge to $-\infty$ and the allocation's cost converges to $-\sum_j \pi(\theta_j)\theta_j$. Now since $\hat{w}_m$ and w_k converge to $-\infty$, there exist subsequences $\hat{w}_{m_k}$ and w_{n_k} such that $\hat{w}_{m_k} \geqslant w_{n_k}$ and therefore by optimality:

$$\sum_{\theta_j}\pi(\theta_j)\left[c_{m_k}(\theta_j) + an_{m_k}(\theta_j) - \theta_j l_{m_k}(\theta_j) + \frac{1}{R}n_{m_k}(\theta_j)\hat{v}(\tilde{w})\right] \geqslant T\hat{v}(w_{n_k}).$$

Hence,

$$-\sum_{\theta_k}\pi(\theta_k)\theta_k \geqslant \liminf_{n\to\infty} T\hat{v}(w_k) > -\sum_{\theta_k}\pi(\theta_k)\theta_k$$

and we have a contradiction. This completes the proof. □

Since $h(\cdot)$ is unbounded below, given the above lemma, for low enough $w \in \mathbb{R}_-$, the allocations will be interior. Since v is differentiable by assumption, positive Lagrange multipliers $\lambda, \mu(\theta_i, \theta_j)|_{i>j}$ must exist such that

$$u'\big(c(\theta_i,w)\big)\Big[\pi(\theta_i)\lambda(w)+\sum_{j<i}\mu(\theta_i,\theta_j;w)-\sum_{j>i}\mu(\theta_j,\theta_i;w)\Big]=\pi(\theta_i),$$

$$\beta n(\theta_i,w)^{\eta-1}\Big[\pi(\theta_i)\lambda(w)+\sum_{j<i}\mu(\theta_i,\theta_j;w)-\sum_{j>i}\mu(\theta_j,\theta_i;w)\Big]=\pi(\theta_i)\frac{1}{R}v'\big(w'(\theta_i,w)\big),$$

$$\begin{aligned}&h'\big(1-l(\theta_i,w)-bn(\theta_i,w)\big)\Big[\pi(\theta_i)\lambda(w)+\sum_{j<i}\mu(\theta_i,\theta_j;w)\Big]\\&\quad-\sum_{j>i}\mu(\theta_j,\theta_i;w)\frac{\theta_i}{\theta_j}h'\Big(1-\frac{\theta_i l(\theta_i,w)}{\theta_j}-bn(\theta_i,w)\Big)=\pi(\theta_i)\theta_i,\end{aligned}$$

$$\begin{aligned}&\big\{-bh'\big(1-l(\theta_i,w)-bn(\theta_i,w)\big)\\&\quad+\beta\eta n(\theta_i,w)^{\eta-1}w'(\theta_i,w)\big\}\Big[\pi(\theta_i)\lambda(w)+\sum_{j<i}\mu(\theta_i,\theta_j;w)\Big]\\&\quad-\sum_{j>i}\mu(\theta_j,\theta_i;w)\Big\{-bh'\Big(1-\frac{\theta_i l(\theta_i,w)}{\theta_j}-bn(\theta_i,w)\Big)+\beta\eta n(\theta_i,w)^{\eta-1}w'(\theta_i,w)\Big\}\\&=\pi(\theta_i)\frac{1}{R}\big(v\big(w'(\theta_i,w)\big)+a\big).\end{aligned}$$

By Lemma 2, we must have:

$$\lim_{w\to-\infty}c(\theta_j,w)=0,$$
$$\lim_{w\to-\infty}n(\theta_j,w)=0,$$
$$\lim_{w\to-\infty}l(\theta_j,w)=1.$$

Then for every $\epsilon>0$, there exists W such that for all $w<W$, we have $u'(c(\theta_j,w))>\frac{M}{\epsilon}$, $h'(1-l(\theta_j,w)-bn(\theta_j,w))>\frac{M}{\epsilon}$, for some large number M. This implies that

$$\lambda(w)=\sum_j\frac{\pi_j}{u'(c(\theta_j,w))}<\frac{\epsilon}{M},$$

$$\pi(\theta_I)\lambda(w)+\sum_{j<I}\mu(\theta_n,\theta_j;w)=\frac{\pi(\theta_I)}{u'(c(\theta_I,w))}<\frac{\epsilon}{M}$$

$$\Rightarrow\quad\mu(\theta_I,\theta_j;w)<\frac{\epsilon}{M}.$$

In addition,

$$\pi(\theta_{I-1})\lambda(w)+\sum_{j<I-1}\mu(\theta_{I-1},\theta_j;w)-\mu(\theta_I,\theta_{I-1};w)=\frac{\pi(\theta_I)}{u'(c(\theta_{I-1},w))}<\frac{\epsilon}{M}$$

$$\Rightarrow\quad\mu(\theta_{I-1},\theta_j;w)<\frac{2\epsilon}{M}.$$

By an induction argument, we have

$$\mu(\theta_j,\theta_j;w)<\frac{z_i\epsilon}{M}$$

where $z_{I-1} = 1$, $z_{I-2} = 2$, $z_{I-i} = z_{I-1} + \cdots + z_{I-i+1} + 1$. If we pick M so that $z_1 < M$, we have that

$$\mu(\theta_j, \theta_j; w) < \epsilon \quad \forall w < W.$$

Next, we define the type specific resetting values, $\underline{w}_i$, as the values of w that solve the following equations:

$$\eta w v'(w) - v(w) = aR + bR\theta_i.$$

Under our convexity assumptions, the left hand side – $\eta v'(w)w - v(w)$ – is strictly increasing in w, so that if a solution exists, it is unique.

From Proposition 6 we know that there is a w_I^* such that:

$$\eta v'(w)w - v(w) = aR + bR\theta_I.$$

Moreover, from the first order conditions, we know that

$$\eta v'\big(w'(\theta_i, w)\big)w'(\theta_i, w) - v\big(w'(\theta_i, w)\big) \leqslant aR + bR\theta_i.$$

Therefore, by the Intermediate Value Theorem, there exists a unique $\underline{w}_i > -\infty$ which satisfies

$$\eta v'(\underline{w}_i)\underline{w}_i - v(\underline{w}_i) = aR + bR\theta_i.$$

Moreover, by substituting first order conditions, we get:

$$\begin{aligned}
\pi\big((\theta_i)b\theta_i + a\big) &\geqslant \pi(\theta_i)\frac{1}{R}\eta v'\big(w'(\theta_i, w)\big)w'(\theta_i, w) - \pi(\theta_i)\frac{1}{R}v\big(w'(\theta_i, w)\big) \\
&= \pi\big((\theta_i)b\theta_i + a\big) \\
&\quad - b\sum_{j>i}\left(1 - \frac{\theta_i}{\theta_j}\right)\mu(\theta_i, \theta_j)h'\left(1 - \frac{\theta_i l(\theta_i, w)}{\theta_j} - bn(\theta_i, w)\right) \\
&> \pi\big((\theta_i)b\theta_i + a\big) - b\epsilon\sum_{j>i}\left(1 - \frac{\theta_i}{\theta_j}\right)h'\left(1 - \frac{\theta_i l(\theta_i, w)}{\theta_j} - bn(\theta_i, w)\right).
\end{aligned}$$

Since hours worked converges to 1, the term multiplied by ϵ in the above expression is bounded away from ∞ as $w \to -\infty$. From this it follows that

$$\lim_{w \to -\infty} \eta v'\big(w'(\theta_i, w)\big)w'(\theta_i, w) - v\big(w'(\theta_i, w)\big) = aR + bR\theta_i.$$

Continuity of v' implies that

$$\lim_{w \to -\infty} w'(\theta_i, w) = \underline{w}_i.$$

This completes the proof.

Appendix C. Proof of Theorem 1

Before we prove the theorem, we need to construct a compact set of continuation utilities, $[\underline{w}, \overline{w}]$ such that:

1. For all $w \in [\underline{w}, \overline{w}]$, there is a solution to problem (Pt);
2. For all $w \in [\underline{w}, \overline{w}]$, $w'(w, \theta) \in [\underline{w}, \overline{w}]$;
3. $n(w, \theta)$ and $w'(w, \theta)$ are continuous functions of w on $[\underline{w}, \overline{w}]$;
4. $\gamma(\Psi)$ is bounded away from zero for the probability distributions on $[\underline{w}, \overline{w}]$.

We proceed by defining $\underline{w}$ and $\overline{w}$. For any fixed $w < 0$, consider the problem:

$$\max_{n \in [0, 1/b]} h(1 - bn) + \beta n^\eta w.$$

Note that there is a unique solution to this problem for every $w < 0$. Moreover, this solution is continuous in w. Let $g(w)$ denote the maximized value in this problem and note that it is strictly increasing in w. Because of this, $\lim_{w \to 0} g(w)$ exists. In a slight abuse of notation, let $g(0) = \lim_{w \to 0} g(w)$. Further, since $w < 0$, it follows that $g(w) < h(1)$ and hence, $g(0) \leqslant h(1)$. In fact, $g(0) = h(1)$. To see this, consider the sequences $w_k = -1/k$, $n_k = k^{1/(2\eta)}$. Then, for k large enough, n_k is feasible and therefore, $g(w_k) \geqslant h(1 - bn_k) + \beta n_k^\eta w_k$. Hence,

$$\begin{aligned} h(1) &= \lim_{k \to \infty} h(1 - bn_k) - \beta k^{-1/2} \\ &= \lim_{k \to \infty} h(1 - bn_k) + \beta n_k^\eta w_k \leqslant \lim_k g(w_k) = g(0) \leqslant h(1). \end{aligned}$$

Thus, in a neighborhood of $w = 0$, $g(w) < w$.

Assume that $b < 1$ (thus, it is physically possible for the population to reproduce itself). Then, it also follows that for w small enough, $g(w) > w$. Hence, there is at least one fixed point for g. Since g is continuous, the set of fixed points is closed. Given this there is a largest fixed point for g. Let $\overline{w}$ be this fixed point. Since $g(w) < w$ in a neighborhood of 0, it follows that $\overline{w} < 0$. Following Corollary 4, choose $\underline{w} = \underline{w}(\overline{w})$.

With these definitions, it follows that, as long as a solution to the functional equation exists for all $w \in [\underline{w}, \overline{w}]$, the condition $w'(w, \theta) \in [\underline{w}, \overline{w}]$ will hold. I.e., requirement 2 above is satisfied.

As noted above, we have no way to guarantee from first principles that the requisite convexity assumptions are satisfied to guarantee that a unique solution to the functional equation exists and is unique (i.e., 1 and 3 above). Thus, we will simply assume that this holds. Given this assumption, 4 can be shown to hold since n must be bounded away from zero on $[\underline{w}, \overline{w}]$ for the promise keeping constraint to be satisfied (i.e., $n = 0$ would imply expected discounted utility is $-\infty$).

Now we can proceed with the proof of Theorem 1.

Proof of Theorem 1. Since $[\underline{w}, \overline{w}]$ is compact in $\mathbb{R}$, by the Riesz Representation Theorem [10, IV.6.3], the space of regular measures is isomorphic to the space $C^*([\underline{w}, \overline{w}])$, the dual of the space of bounded continuous functions over $[\underline{w}, \overline{w}]$. Moreover, by Banach–Alaoglu Theorem [24, Theorem 3.15], the set $\{\Psi \in C^*([\underline{w}, \overline{w}]);\ \|\Psi\| \leqslant k\}$ is a compact set in the weak-$*$ topology for any $k > 0$. Equivalently the set of regular measures, Ψ, with $\|\Psi\| \leqslant 1$, is compact. Since non-negativity and full measure on $[\underline{w}, \overline{w}]$ are closed restrictions, we must have that the set

$$\bigl\{\Psi\colon\ \Psi \text{ a regular measure on } [\underline{w}, \overline{w}],\ \Psi\bigl([\underline{w}, \overline{w}]\bigr) = 1,\ \Psi \geqslant 0\bigr\}$$

is compact in weak-$*$ topology.

By definition,

$$T(\Psi)(A) = \int_{[\underline{w},\overline{w}]} \sum_{i=1}^{n} \pi_i \mathbf{1}\big\{w'(w,\theta_i) \in A\big\} n(w,\theta_i)\, d\Psi(w).$$

The assumption that the policy function is unique implies that it is continuous by the theorem of the maximum. It also follows from this that n is bounded away from 0 on $[\underline{w}, \overline{w}]$ (since otherwise utility would be $-\infty$). From this, it follows that T is continuous in Ψ. Moreover,

$$\gamma(\Psi) = \int_{[\underline{w},\overline{w}]} \sum_{i=1}^{n} \pi_i n(w,\theta_i)\, d\Psi(w) \geqslant \underline{n} > 0$$

is a continuous function of Ψ and is bounded away from zero.

Therefore, the function

$$\hat{T}(\Psi) = \frac{T(\Psi)}{\gamma(\Psi)} : \mathcal{M}\big([\underline{w}, \overline{w}]\big) \to \mathcal{M}\big([\underline{w}, \overline{w}]\big)$$

is continuous. Thus, by the Schauder–Tychonoff Theorem [10, V.10.5], $\hat{T}$ has a fixed point $\Psi^* \in M([\underline{w}, \overline{w}])$. □

Appendix D. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jet.2013.04.022.

References

[1] S. Albanesi, C. Sleet, Dynamic optimal taxation with private information, Rev. Econ. Stud. 73 (2006) 1–30.
[2] F. Alvarez, Social mobility: The Barro–Becker children meet the Laitner–Loury dynasties, Rev. Econ. Dynam. 2 (1999) 65–103.
[3] A. Atkeson, R. Lucas, On efficient distribution with private information, Rev. Econ. Stud. 59 (1992) 427–453.
[4] A. Atkeson, R. Lucas, Efficiency and equality in a simple model of efficient unemployment insurance, J. Econ. Theory 66 (1995) 64–88.
[5] R.J. Barro, G.S. Becker, Fertility choice in an endogenous growth model, Econometrica 57 (1989) 481–501.
[6] G.S. Becker, R.J. Barro, A reformulation of the economic theory of fertility, Quart. J. Econ. 103 (1988) 1–25.
[7] L.M. Benveniste, J.A. Scheinkman, On the differentiability of the value function in dynamic models of economics, Econometrica 47 (1979) 727–732.
[8] G. Clark, G. Hamilton, Survival of the richest: The Malthusian mechanism in pre-industrial England, J. Econ. Hist. 66 (2006) 707.
[9] D. de la Croix, M. Doepke, Inequality and growth: Why differential fertility matters, Amer. Econ. Rev. 93 (2003) 1091–1113.
[10] N. Dunford, J. Schwartz, Linear Operators, Interscience Publishers, 1958.
[11] E. Farhi, I. Werning, Inequality and social discounting, J. Polit. Economy 115 (2007) 365–402.
[12] E. Farhi, I. Werning, Progressive estate taxation, Quart. J. Econ. 125 (2010) 635–673.
[13] M. Golosov, N. Kocherlakota, A. Tsyvinski, Optimal indirect and capital taxation, Rev. Econ. Stud. 70 (2003) 569–587.
[14] M. Golosov, A. Tsyvinski, Designing optimal disability insurance: A case for asset testing, J. Polit. Economy 114 (2006) 257–279.
[15] M. Golosov, A. Tsyvinski, Optimal taxation with endogenous insurance markets, Quart. J. Econ. 122 (2007) 487–534.

[16] E.J. Green, Lending and the smoothing of uninsurable income, in: Contractual Arrangements for Intertemporal Trade, 1987, pp. 3–25.

[17] L. Jones, A. Schoonbroodt, Complements versus substitutes and trends in fertility choice in dynastic models, Int. Econ. Rev. 51 (2010) 671–699.

[18] A. Khan, B. Ravikumar, Growth and risk-sharing with private information, J. Monet. Econ. 47 (2001) 499–521.

[19] N. Kocherlakota, Zero expected wealth taxes: A Mirrlees approach to dynamic optimal taxation, Econometrica 73 (2005) 1587–1621.

[20] N.R. Kocherlakota, The New Dynamic Public Finance, Princeton University Press, 2010.

[21] P. Milgrom, I. Segal, Envelope theorems for arbitrary choice sets, Econometrica 70 (2002) 583–601.

[22] C. Phelan, On the long run implications of repeated moral hazard, J. Econ. Theory 79 (1998) 174–191.

[23] C. Phelan, Opportunity and social mobility, Rev. Econ. Stud. 73 (2006) 487–504.

[24] W. Rudin, Functional Analysis, McGraw–Hill, 1991.

[25] J. Thomas, T. Worrall, Income fluctuation and asymmetric information: An example of a repeated principal–agent problem, J. Econ. Theory 51 (1990) 367–390.

Please cite this article in press as: R. Hosseini et al., Optimal contracting with dynastic altruism: Family size and per capita consumption, J. Econ. Theory (2013), http://dx.doi.org/10.1016/j.jet.2013.04.022