

Optimal Contracting with Dynastic Altruism: Family Size and Per Capita Consumption*

Roozbeh Hosseini

Arizona State University

rhosseini@asu.edu

Larry E. Jones

University of Minnesota

lej@umn.edu

Ali Shourideh

Wharton and NYU

shourideh@nyu.edu

March 19, 2012

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, in addition to the usual variables, 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.

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 [Green \(1987\)](#), [Thomas and Worrall \(1990\)](#), [Atkeson and Lucas \(1992\)](#) and [Phelan \(1998\)](#) as examples). This arises due to the desire, by both parties, to use future

*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 thank NSF for financial support, NSF Grant No. SES-0962432.

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 current period.

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 [Phelan \(2006\)](#) and [Farhi and Werning \(2007, 2010\)](#) (and implicitly earlier in [Atkeson and Lucas \(1992\)](#)) 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 [\(de la Croix and Doepke, 2003\)](#) 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 \(1989\)](#) (and [Becker and Barro \(1988\)](#)) style dynastic preferences 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 (Farhi and Werning, 2007) 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 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.

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. If, for whatever reason, the discount factor is less than the inverse of the (endogenous) rate of return on capital, we show that family size shrinks over time (i.e., fertility is below replacement) in the optimal contract. On the other hand, if, as seems to be true empirically, the discount rate exceeds the inverse of the rate of return on capital, two additional key factors determine the long run properties of population size: the curvature of utility and the amount of uncertainty about 'type.' 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.

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 [Green \(1987\)](#), [Thomas and Worrall \(1990\)](#), [Atkeson and Lucas \(1992\)](#), [Phelan \(2006\)](#), [Farhi and Werning \(2007, 2010\)](#) (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 \(1998\)](#) shows that this result is robust to many variations in the assumptions of the model. Moreover, [Khan and Ravikumar \(2001\)](#) establish numerically that in a production economy, the same result holds and although the economy grows, the detrended distribution of consumption has a negative trend. 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 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 basic technical results concerning the efficient allocation. Our model is an extension of the intergenerational interpretation of [Atkeson and Lucas \(1992\)](#) [Farhi and Werning \(2007, 2010\)](#) 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 consumption good. We assume that it takes a units of consumption good to raise each child. We will report the results with 'time cost' included in [section 4](#).¹

¹Goods 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.

Preferences. Following [Barro and Becker \(1989\)](#) and [Becker and Barro \(1988\)](#), we assume that the family head has preferences over own consumption, c_0 , own leisure, ℓ_0 , own fertility, n_0 , and per child utility, u_1 given by:

$$U_0 = u(c_0, \ell_0) + \beta g(n_0)u_1,$$

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:

$$U_0 = u(c_0, \ell_0) + \beta g(n_0)u(c_1, \ell_1) + \beta^2 g(n_0)g(n_1)u(c_2, \ell_2) + \dots$$

As in [Barro and Becker \(1989\)](#) and [Becker and Barro \(1988\)](#), this simplifies when it is assumed that $g(\cdot)$ is iso-elastic, $g(n) = n^n$. In this case, it follows that $g(n_0)g(n_1) = g(n_0n_1)$, etc. Under this assumption, the utility of the family head is given by:

$$U_0 = u(c_0, \ell_0) + \beta n_0^n u(c_1, \ell_1) + \beta^2 (n_0n_1)^n u(c_2, \ell_2) + \dots$$

We will assume that utility is additively separable in consumption and leisure², and we will write flow utility in terms of amount of efficiency units of labor, y 's, in place of levels of leisure:³

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

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 .

Written in this way, θ can be interpreted as a productivity shock, as well 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,

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

in which $\hat{h}(\cdot)$ is the utility function over leisure and it is strictly increasing and strictly

²This is standard assumption in dynamic Mirrleesian models. see for example [Albanesi and Sleet \(2006\)](#); [Farhi and Werning \(2010\)](#); [Goloso et al. \(2003\)](#); [Goloso and Tsyvinski \(2006, 2007\)](#); [Kocherlakota \(2005, 2010\)](#) among many others.

³Alternatively, one can think of $y = \theta \cdot l$ as output that is required from worker of type θ .

concave. Similarly, if θ is a taste shock, assume that

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

Information. The set of possible types is given by $\Theta = \{\theta_1 < \dots < \theta_I\}$. 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 $\theta^t = (\theta_0, \dots, \theta_t)$. We assume that the planner can observe the amount of effective labor units 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 are of type θ .

Technology. There is a constant returns to scale production function $F(K_t, Y_t)$ where K_t is aggregate capital and $Y_t = \sum_{\theta^t} \pi(\theta^t) y_t(\theta^t)$ 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

$$F(K, Y) = RK + Y.$$

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

Allocation. An allocation is a vector $\{c_t, y_t, n_t\}_{t=0}^{\infty}$ where $c_t : \Theta^t \rightarrow \mathbb{R}_+$, $y_t : \Theta^t \rightarrow \mathbb{R}_+$ and $n_t : \Theta^t \rightarrow \mathbb{R}_+$. Here, $c_t(\theta^t)$ is per capita consumption of a generation with history θ^t ; $y_t(\theta^t)$ is the amount of per capita effective labor units this generation supplies; $n_t(\theta^t)$ is number of children that each individual in this generation has.

2.1 The Contracting Problem

Consider a social planner who offers a contract $\{c_t(\theta^t), y_t(\theta^t), n_t(\theta^t)\}_{t=0}^{\infty}$ to each dynasty in period $t = 0$. Each dynasty chooses a reporting strategy $\sigma = \{\sigma_t\}$, which is a sequence of functions $\sigma_t : \Theta^{t+1} \rightarrow \Theta$ that maps a history of shocks θ^t into a current report $\hat{\theta}_t \in \Theta$. Any strategy σ induces history of reports $\sigma^t = (\sigma_0(\theta_0), \dots, \sigma_t(\theta^t))$ with $\sigma^t : \Theta^{t+1} \rightarrow \Theta^{t+1}$. 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 $c_t(\sigma_t(\theta^t))$ units of consumption, will supply $y_t(\sigma_t(\theta^t))$ units of efficiency labor units and will have $n_t(\sigma_t(\theta^t))$ kids.

We study 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.

Define $N_t(\theta^{t-1}) = N_{-1} \times n_0(\theta_0) \times n_1(\theta_0, \theta_1) \times \dots \times n_{t-1}(\theta^{t-1})$ as the current population of a cohort (in date t) whose ancestors received the history of shocks of θ^{t-1} . We write the contracting problem as

$$\min_{\{c_t(\theta^t), y_t(\theta^t), n_t(\theta^t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \sum_{\theta^t} \frac{1}{R^t} \pi(\theta^t) N_t(\theta^{t-1}) [c_t(\theta^t) + a n_t(\theta^t) - y_t(\theta^t)]$$

subject to

$$\sum_{t=0}^{\infty} \sum_{\theta^t} \beta^t \pi(\theta^t) N_t(\theta^{t-1})^\eta [u(c_t(\theta^t)) + h(y_t(\theta^t), \theta_t)] \geq W_0 \quad (1)$$

and

$$\begin{aligned} \sum_{t=0}^{\infty} \sum_{\theta^t} \beta^t \pi(\theta^t) N_t(\theta^{t-1})^\eta [u(c_t(\theta^t)) + h(y_t(\theta^t), \theta_t)] \geq \\ \sum_{t=0}^{\infty} \sum_{\theta^t} \beta^t \pi(\theta^t) N_t(\sigma_{t-1}(\theta^{t-1}))^\eta [u(c_t(\sigma_t(\theta^t))) + h(y_t(\sigma_t(\theta^t)), \theta_t)] \end{aligned} \quad \text{for all } \sigma \in \Sigma_0 \quad (2)$$

in which equation (1) and (2) are promise keeping and incentive compatibility constraints.⁴

Note that the objective and constraint set in the above problem may not be convex due to terms that involve multiplication of fertility (or population) -which is a choice variable- by other allocations. Therefore, we follow Alvarez (1999) and reformulate the problem in terms of aggregates rather than per capita variables. For any history θ^t , define $C_t(\theta^t) = N_t(\theta^{t-1}) c_t(\theta^t)$ and $Y_t(\theta^t) = N_t(\theta^{t-1}) y_t(\theta^t)$ to be aggregate consumption and aggregate effective labor unit for a generation with history θ^t whose ancestors received history of

⁴Note that if W_0 is such that the profit earned at the solution to this problem is 0, this problem is the dual (assuming convexity) of a planning problem where agent utility is maximized.

shocks of θ^{t-1} . Using this notation, we can rewrite the contracting problem as:⁵

$$\min_{\{C_t(\theta^t), Y_t(\theta^t), N_{t+1}(\theta^t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \sum_{\theta^t} \frac{1}{R^t} \pi(\theta^t) [Y_{t+1}(\theta^t) - C_{t+1}(\theta^t) - aN_{t+1}(\theta^t)] \quad (3)$$

subject to

$$\sum_{t=0}^{\infty} \sum_{\theta^t} \beta^t \pi(\theta^t) N_t(\theta^{t-1})^\eta \left[u\left(\frac{C_t(\theta^t)}{N_t(\theta^{t-1})}\right) + h\left(\frac{Y_t(\theta^t)}{N_t(\theta^{t-1})}, \theta_t\right) \right] \geq W_0 \quad (4)$$

and

$$\begin{aligned} \sum_{t=0}^{\infty} \sum_{\theta^t} \beta^t \pi(\theta^t) N_t(\theta^{t-1})^\eta \left[u\left(\frac{C_t(\theta^t)}{N_t(\theta^{t-1})}\right) + h\left(\frac{Y_t(\theta^t)}{N_t(\theta^{t-1})}, \theta_t\right) \right] \geq \\ \sum_{t=0}^{\infty} \sum_{\theta^t} \beta^t \pi(\theta^t) N_t(\sigma_{t-1}(\theta^{t-1}))^\eta \left[u\left(\frac{C_t(\sigma_t(\theta^t))}{N_t(\sigma_{t-1}(\theta^{t-1}))}\right) + h\left(\frac{Y_t(\sigma_t(\theta^t))}{N_t(\sigma_{t-1}(\theta^{t-1}))}, \theta_t\right) \right] \end{aligned} \quad \text{for all } \sigma \in \Sigma_0 \quad (5)$$

Using standard arguments for the i.i.d. shock, this problem can be reformulated as a dynamic programming problem with state variables W , 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:

$$V(N, W) = \min_{C(\theta), Y(\theta), N'(\theta), W'(\theta)} \sum_{\theta} \pi(\theta) \left[C(\theta) + aN'(\theta) - Y(\theta) + \frac{1}{R} V(N'(\theta), W'(\theta)) \right] \quad (\text{P})$$

subject to

$$\sum_{\theta} \pi(\theta) \left[N^\eta \left\{ u\left(\frac{C(\theta)}{N}\right) + h\left(\frac{Y(\theta)}{N}, \theta\right) \right\} + \beta W'(\theta) \right] = W \quad (6)$$

⁵The resulting 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 the next section we state sufficient conditions on preferences that guarantee the convexity of the problem.

and

$$N^\eta \left\{ u \left(\frac{C(\theta)}{N} \right) + h \left(\frac{Y(\theta)}{N}, \theta \right) \right\} + \beta W'(\theta) \geq N^\eta \left\{ u \left(\frac{C(\hat{\theta})}{N} \right) + h \left(\frac{Y(\hat{\theta})}{N}, \theta \right) \right\} + \beta W'(\hat{\theta})$$

$$\forall \theta, \hat{\theta} \in \Theta,$$
(7)

in which $W'(\theta)$ 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 homotheticity property and so, it can be equivalently be formulated in per capita terms⁶. 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):

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

s.t.

$$\sum_{\theta} \pi(\theta) c(\theta) (u(c(\theta)) + h(y(\theta), \theta) + \beta n(\theta)^\eta w'(\theta)) = w \quad (8)$$

$$u(c(\theta)) + h(y(\theta), \theta) + \beta n(\theta)^\eta w'(\theta) \geq u(c(\hat{\theta})) + h(y(\hat{\theta}), \theta) + \beta n(\hat{\theta})^\eta w'(\hat{\theta})$$

$$\forall \theta, \hat{\theta} \in \Theta. \quad (9)$$

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 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 [Green \(1987\)](#), [Thomas and Worrall \(1990\)](#) and [Atkeson and Lucas \(1992\)](#)) – immiseration. More-

⁶See [Hosseini et al. \(2009\)](#) for a formal proof of this result

over, from any initial condition (N_0, W_0) we jump to the stationary distribution in one period. This follows because of 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 the consumption, hours and efficiency units of labor are i.i.d. Thus, when the cost of raising children is only in terms of consumption goods and shocks are i.i.d., the optimal contract has an extreme lack of history dependence once fertility is introduced.⁷

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 value function $V(\cdot, \cdot)$ is both strictly convex and differentiable. In the next section, we show that in some cases, when θ is a taste shock to the utility value of leisure, these assumptions on V can be derived from first principles.⁸

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 concave and differentiable and that the solution to problem (P) (and (P’)) is interior. Then,*

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 .

As a first step in proving part 1, we will give a partial characterization of the solution in terms of continuation utilities, w . Let $\mu(\theta, \hat{\theta})$ and λ be multipliers on constraints (8) and (9). Taking first order conditions with respect to $n(\theta)$ and $w'(\theta)$ in problem (P’) gives us the following two equations:

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

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

⁷The situation is more complex when the costs of children is in terms of parental time. Although there is still a stationary distribution over per capita variables, they are not i.i.d., so that there is more history dependence in the optimal contract. See the more detailed discussion of the time cost case in section 4 below.

⁸In appendix C we show that this result holds even in a sequence formulation of the problem without relying on convexity and differentiability. We focus on the recursive formulation here for better exposition.

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:

$$Ra + v(w'(\theta)) = \eta w'(\theta) v'(w'(\theta)). \quad (10)$$

For intuition, imagine the planner who wants to deliver a certain level of utility to the parent in the future. Once the planner decides about the level of future utility to the parent, it needs to find the mix of the number of children and promised utility to each children that delivers that level of utility. Let's keep the future utility to the parent – that is $\beta n(\theta)^\eta w'(\theta)$ – fixed. Increasing the number children by one has the cost $Ra + v(w'(\theta))$ in terms of the future consumption good. On the other hand, to keep the future utility of parents constant implies that promised utility to each child has to go down by $\eta \frac{w'(\theta)}{n(\theta)}$. This reduction in the promised utility to each child has a cost saving effect of $n(\theta) v(w'(\theta)) \times \eta \frac{w'(\theta)}{n(\theta)}$ in the units of the future consumption good. The marginal cost of increasing the number of children and the marginal cost saved from reducing each child's promised utility must be equal at the optimal contract. Otherwise, the planner can change the combination of $n(\theta)$ and $w'(\theta)$ and reduce the cost of delivering the same future utility to the parent. This implies that equation (10) must hold at the optimal contract.

From our assumption that value function in problem (P), V , is strictly convex and differentiable, it follows that $\eta w v'(w) - v(w)$ is strictly decreasing.⁹ Thus, there is a unique solution to equation (10). 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 properties of this equation: First, promised utility to the parent, w , does not appear in the equation. This implies that promised utility for each children,

⁹See appendix A

$w'(\theta)$, is independent of promised utility of the parent, 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 either. 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:

Lemma 1 *Assume that value function in problem (P), V , is strictly concave 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 (10).*

It 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 and all t and any initial promised utility w_0 . Therefore, the per capita allocations are:

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

and hence they depend only on 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 allocations are the solution of the following static problem

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

s.t.

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

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

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.

This results of proposition 1 do not hold in general, they hold because of the special assumption that the cost of children (i.e., a) is completely in terms of goods – children do not take any parental time to raise. When there is a time cost for children, the ‘resetting’ property (lemma 1) only holds for the highest type (undistorted types) and at the very bottom of the promised utility distribution (i.e., low w). We will discuss this in more details in section 4.

The results of proposition 1 and lemma 1 imply that the long run distribution of future promised utilities is concentrated on w^* . i.e., in contrast to the case where fertility is not included (e.g. Green (1987), Thomas and Worrall (1990) and Atkeson and Lucas (1992)), 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 rate (as in Farhi and Werning (2007, 2010)).

Unlike standard models of dynamic contracting, where incentives to parent are provided only by changing the promised utility to each child, in our environment, incentives to parents are provided using a combination of promised utility to each child and the number of children a parent has. Given the homotheticity of the Barro-Becker formulation, the planner will use fertility more heavily to provide incentives. Hence, promised utility is not used at all for incentive provision.

This is an extreme result in that it says that there is no history dependence of ‘wealth’ at all¹⁰ – promised utility of children is independent of the shock received by the parent. This is something that is unique to this particular example – in which there is no time cost for raising children. As an intermediate case, it can be shown that if the goods cost of children depends on the type of the parent – $a(\theta)$ – continuation utility depends on the type of the parent, but not on the promised utility to the parent. Thus, in this case, w' (as well as other per capita allocations) is $MA(1)$.

The intuition for our result can best be understood in two steps. First, since the shock

¹⁰Note that $v(w)$ is the different between present value of income and expenditure so it can be thought of as wealth. Lack of history dependence in promised utility is equivalent to no history dependence in wealth

is only to the utility of leisure and the costs of children are entirely in terms of goods, the optimal contract features an ‘undistorted’ choice between current consumption and future children/future consumption for all types of parents (i.e., for all θ). Second, the Barro–Becker model of dynastic altruism has a natural homotheticity property in aggregate allocations, i.e. let $U(C, Y, N, \theta) = N^\eta (u(C/N) + h(Y/N, \theta))$, then

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

Now consider the problem 3. 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 in problem 3 are homogenous 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 after one period. Therefore, the per capita allocations in the first period are independent of W_0 . The same intuition applies to future periods.

If part of the costs of children are in terms of parental time a more limited version of this characterization result still holds. In that case there is no history dependence in continuation utility only for types whose allocations are undistorted – typically for the highest (or lowest) shock. We will overview this result in section 4.¹¹

2.3 A Special Case: Taste Shocks

As can be seen from the previous section, the technical difficulty that arises is to show that the value function is strictly concave and differentiable. Once we have these properties, our main result – that continuation utility is reset to the same value for the children of all types – follows directly from the first order conditions of the contracting problem.

In this section, 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:

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

¹¹For more details we refer the reader to [Hosseini et al. \(2009\)](#).

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:

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

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

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

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

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

Under these transformations, the promise keeping constraint **6** becomes:

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

and, the incentive constraints **7** become:

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

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:

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

For this objective function to be strictly convex, it is necessary that the two function $Nh^{-1}(N^{-\eta}H)$ 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:

$$V(N, W) = \max_{U(\theta), H(\theta), N'(\theta), W'(\theta)} \sum_{\theta} \pi(\theta) \left[Nu^{-1}(N^{-\eta}U(\theta)) + aN'(\theta) - N(1 - h^{-1}(N^{-\eta}H(\theta))) + \frac{1}{R}V(N'(\theta), W'(\theta)) \right] \quad (11)$$

subject to

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

and,

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

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

Proposition 2 *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 (3). Then, the solution to the functional equation (11) is strictly convex on A and differentiable on the interior of A . Moreover, there is a function, $v : \mathbb{R} \rightarrow \mathbb{R}$, 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 (11) and use the principle of optimality to show that $V(\cdot, \cdot)$ is convex. Consider the following sequence problem

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

subject to

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

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

$$\text{given } W_0 = W, N_0 = N$$

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 $(\tilde{N}, \tilde{W}) = \lambda(N^1, W^1) + (1 - \lambda)(N^2, W^2)$ and define the following allocations

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

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 $V(\tilde{N}, \tilde{W}) \leq \tilde{V}$. Note also that, since the objective function is convex

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

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 the value function is differentiable, we use a technique similar to [Benveniste and Scheinkman \(1979\)](#). To do so, we also need to use results developed by [Milgrom and Segal \(2002\)](#). 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

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

subject to

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

and

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

Note that since we have fixed future policies, we must have that $G(\tilde{N}, \tilde{W}) \leq V(\tilde{N}, \tilde{W})$ with equality at $(N, W) = (\tilde{N}, \tilde{W})$. Since the cost functions are strictly convex and the set of constraints are linear, the solution to the above maximization 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 [Milgrom and Segal \(2002\)](#), $G(\cdot, \cdot)$ 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 [Benveniste and Scheinkman \(1979\)](#), $V(\cdot, \cdot)$ 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 (11) preserves that property. That is, if $T(\cdot)$ is the transformation associated with the functional equation (11) and the set S is defined as the following:

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

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

$$TV(N, W) = \max_{H(\theta), U(\theta), N'(\theta), W'(\theta)} \sum_{\theta} \pi(\theta) \left[N - Nh^{-1}(N^{-\eta}H(\theta)) - Nu^{-1}(N^{-\eta}U(\theta)) \right. \\ \left. - aN'(\theta) + \frac{1}{R}V(N'(\theta), W'(\theta)) \right]$$

subject to

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

and,

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

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

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

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,

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

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 the following standard assumptions¹²:

Assumption 1 *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 1 *Under assumptions (1):*

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 .

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.

¹²In addition to this assumption, we have made two extra assumptions on u and h in the statement of Proposition 2. 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 $1 < \sigma_l, \sigma_c < 1 - \eta$

We provide two types of results. In the first, we show that in certain cases, $N_t \rightarrow 0$ *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 and hence, $N_t^{1-\eta} (u(c_t) + h(y_t, \theta_t)) \rightarrow -\infty$, *a.s.* Therefore, 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., $\beta R \leq 1$.

However, the condition $\beta R \leq 1$ 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 $\beta R > 1$. For this reason, we go on to study the population dynamics of the model in the long run when R is determined endogenously.

Our goal is to present sufficient conditions for no ‘population immiseration’ when R is endogenous. To this end, we first study population dynamics when there is full information. In this case we can fully characterize population dynamics in the long run. In particular, we can determine the population growth rate. Under certain sufficient conditions the population growth rate under full information can provide an appropriate upper (or lower) bound for population growth rate in the private information model. We specify these conditions in proposition 6 below.

3.1 An Inverse Euler Equation

To begin, we extend the results in Golosov et al. (2003) to our environment with endogenous fertility and derive a version of Inverse Euler Equation. We summarize this as a Proposition:

Proposition 3 *If, in the optimal contract, consumption is always interior, the optimal allocation satisfies a version of the Inverse Euler Equation:*

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

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

Proof. Let $C_t(\theta^t)$ be $N_t(\theta^{t-1})$ part of the efficient plan and let $c_t(\theta^t) \equiv \frac{C_t(\theta^t)}{N_t(\theta^{t-1})}$ be per capita consumption at history θ^t . Consider decreasing per capita consumption of each parent parent with history θ^t and saving that unit. There will be R units available for next generation which can be distributed among the descendants. We increase consumption of each descendant of type θ^{t+1} by $\epsilon(\theta_{t+1})$ such that:

$$\begin{aligned} N_{t+1}(\theta^t) \sum_{\theta^{t+1}} \pi(\theta^{t+1}) \epsilon(\theta_{t+1}) &= N_t(\theta^{t-1}) R \\ N_{t+1}(\theta^t)^\eta u'(c_{t+1}(\theta^t, \theta)) \epsilon(\theta) &= N_{t+1}(\theta^t)^\eta u'(c_{t+1}(\theta^t, \theta')) \epsilon(\theta') = \Delta \text{ for all } \theta, \theta'. \end{aligned}$$

The first equation is the resource constraint implied by redistributing the available resources. The second equation makes sure that the incentives are aligned by insuring that utility of all descendants increase by the same increment regardless of their ability shock θ_{t+1} . This implies that the change in the utility of all types is the same – there is no incentive to lie. The above equations imply that

$$\sum_{\theta^{t+1}} \pi(\theta_{t+1}) \frac{\Delta}{N_{t+1}(\theta^t)^\eta u'(c_{t+1}(\theta^t, \theta_{t+1}))} = N_t(\theta^{t-1}) R$$

Since the change in utility from this perturbation must be zero, we must have $\beta \Delta = N_t(\theta^{t-1})^\eta u'(c_t(\theta^t))$. Replacing Δ in the above equation implies the claim of the proposition. ■

This result must hold independent of the cost and shock structure as long as utility function is separable between leisure and consumption (See [Hosseini et al. \(2009\)](#) for the derivation in the case where cost of child rearing is in terms of parents time). When, $\beta R \leq 1$, we see from above that $X_t = N_{t+1}(\theta^t)^{1-\eta} / u'(c_{t+1}(\theta^{t+1}))$ 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 $X_t \rightarrow X_\infty$ *a.s.* As is standard in this literature, 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 \(1990\)](#) 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 2 *Assume only local downward constraints are binding. If $\beta R \leq 1$, then $N_t(\theta^{t-1}) \rightarrow 0$ 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 supermartingale 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

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

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

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

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

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

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

Note that when the cost of child rearing is only in terms of consumption goods $c_t(\theta^t)$ 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 $N_t(\theta^{t-1}) \rightarrow 0$ a.s. ■

The fact that $N_{t+1}(\theta^t) \rightarrow 0$ a.s. does not mean that fertility converges to zero almost surely, rather, it means that it is less than replacement (for example, $n_t(\theta^t) < 1$ most of

the time). In fact, $N_{t+1}(\theta^t) \rightarrow 0$ *a.s.* does not even imply that $E(N_{t+1}(\theta^t)) \rightarrow 0$. In computed examples, it can be shown that for certain parameter configurations (with $\beta R = 1$), $E(N_{t+1}(\theta^t)) \rightarrow \infty$ even though $N_{t+1}(\theta^t) \rightarrow 0$ *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. Even though a stationary distribution over per capita variables still exists, it need not follow that $N_{t+1}(\theta^t) \rightarrow 0$ *a.s.* In fact, numerical examples can be constructed in which $N_{t+1}(\theta^t) \rightarrow \infty$ *a.s.* (see [Hosseini et al. \(2010b\)](#) and the discussion below where R is endogenous). However, it can be shown that almost every path converges to zero relative to a trend. In some sense, there is always population immiseration ‘relative to a trend’. The next proposition is a formal statement and proof of this. To see this note that $X_t = \frac{N_t^{1-\eta}}{(\beta R)^t} v'(w_t)$ is always a non-negative martingale.

Proposition 4 *If $\beta R > 1$ and cost of child rearing is only in terms of consumption goods, then $\frac{N_t(\theta^{t+1})}{(\beta R)^{\frac{t}{1-\eta}}} \rightarrow 0$ *a.s.**

Proof. Consider the inverse Euler equation

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

and recall that when the cost of child rearing is only in terms of consumption goods, $c_t(\theta^t)$ is i.i.d. Therefore taking expectations over θ^t we will have

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

and hence

$$E [N_{t+1}(\theta^t)^{1-\eta} | \theta^{t-1}] = \beta R N_t(\theta^{t-1})^{1-\eta}. \quad (12)$$

Dividing both sides by $(\beta R)^t$ we get

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

Therefore, $X_t = \frac{N_{t+1}(\theta^t)^{1-\eta}}{(\beta R)^t}$ is non-negative martingale and by the Martingale Convergence Theorem there exists a non-negative random variable with finite mean, X_∞ , such that $X_t \rightarrow X_\infty$ *a.s.* To establish the claim of the proposition we need to show that $X_\infty = 0$. Note that equation (12) implies

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

By Jensen's inequality

$$E [\log (n (\theta)^{1-\eta})] < \log (E [n (\theta)^{1-\eta}]) = \log (\beta R).$$

Recall that $n_t (\theta^t)$ is i.i.d and we can apply the Law of Large Numbers to show that

$$\begin{aligned} \lim_{t \rightarrow \infty} \log (X_t) &= \lim_{t \rightarrow \infty} \left(\log \left(N_{t+1} (\theta^t)^{1-\eta} \right) - t \log (\beta R) \right) \\ &= \lim_{t \rightarrow \infty} t \left(\frac{\sum_{i=0}^t \log \left(n_i (\theta^i)^{1-\eta} \right)}{t} - \log (\beta R) \right) \\ &= \lim_{t \rightarrow \infty} t \left(E [\log (n (\theta)^{1-\eta})] - \log (\beta R) \right) \\ &= -\infty. \end{aligned}$$

This implies that X_∞ has to be zero. ■

The discussion above indicates the importance of R in determining the limiting behavior of overall population size as well as population of each dynasty. Next, we study the case in which R is endogenously determined in the model.

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 unity. In this section, we consider a more general version of the model in which R is determined endogenously through general equilibrium effects. To simplify the analysis, we focus on the case where the cost of children is in terms of the consumption good and individual types are taste shocks to parent's leisure. We also make a specific simplifying assumption about the utility function over consumption and leisure and its curvature.

We show that whether or not there is population immiseration in depends on two factors. The first has nothing to do with private information and can be summarized by the return on capital in full information economy, which we refer to as R_{FI} . In the economy with full information, fertility, $n (\theta)$, is the same across all realizations of θ and therefore, the growth rate of population is directly connected to the value of βR_{FI} . For different parameter configurations, any of the three possibilities: i) $\beta R_{FI} < 1$; ii) $\beta R_{FI} = 1$; or iii) $\beta R_{FI} > 1$ may arise under full information. In the first case, $N_t \rightarrow 0$, in the second case N_t is constant and in the third case, $N_t \rightarrow \infty$.

In the case where uncertainty is in the form of a taste shock to leisure we can show that return on capital in the private information economy, R_{PI} is equal the return on capital in the full information economy. However, in this case fertility depends on the realization of θ for each parent. The inverse Euler equation establishes a connection between average of $n(\theta)^{1-\eta}$ over realizations of θ and the value of βR_{PI} . The second factor affecting the growth rate of population shows up here. This is the curvature of the utility function. Using Jensen's inequality in the Inverse Euler equation, it follows that increases in uncertainty (or increase in curvature) causes mean growth rates to fall when $\eta < 0$ and rise when $\eta > 0$.¹³

Whether population grows or shrinks in the economy with private information depends on the interplay of these two factors: return on capital and the degree of uncertainty in θ . Our goal in this section is to identify conditions under which we can determine the overall effect. We start by studying the problem with full information and develop, in special cases, expressions for R_{FI} based on the parameters of the model. We go on to re-introduce private information and study the extent to which the interplay of return on capital and the degree of uncertainty in θ can be sorted out.

3.2.1 The full information economy with endogenous R

Here, we extend the model discussed above to allow for the endogenous determination of the interest rate. 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

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

subject to

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

¹³This is the opposite of the usual effect of uncertainty on growth rates because it is coming from the Inverse Euler Equation rather than the Euler Equation (See Jones et al. (2005)).

¹⁴We 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.

To simplify the problem, we make the following assumption:

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

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

and $\eta = 1 - \sigma$

These assumptions will simplify the algebra below. Hence, period utility is simplified to:

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

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 there are no frictions, the planner will 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 simply given by $F_K(K_{t+1}, L_{t+1})$. At the optimum, the planner should be indifferent between investment in physical capital and investment in any dynasty's population. 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

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

¹⁵See Jones and Schoonbroodt (2010) for a more detailed study of this model with a representative agent.

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 population growth rate in terms of primitives:

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

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).

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 . Ideally, we would like to characterize the limiting properties of N_t when R is endogenous to match up with the discussion in section 3.1. This turns out to be difficult however. Because of this, we will focus on mean population growth rates instead. Accordingly, in this section, we discuss conditions when mean population is growing or shrinking.

To do this, we begin by characterizing the solution to the problem of a representative insurance firm taking rental rate of capital, $R_{PI} = F_K$, and wage, $p_{PI} = F_L$, as given (the subscript refers private information economy). The solution to this problem will give rise to aggregate quantities of capital, K_t and labor supply, L_t . At the optimum, on a balanced growth path, it must be true 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.

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

$$V(N, W) = \max_{C(\theta), Y(\theta), N'(\theta), W'(\theta)} \sum_{\theta} \pi(\theta) \left[C(\theta) + aN'(\theta) - p_{PI}Y(\theta) + \frac{1}{R_{PI}}V(N'(\theta), W'(\theta)) \right] \quad (13)$$

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

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

As in the case 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

$$\frac{1}{R_{PI}} V_N(N'(\theta), W'(\theta)) = a \quad (14)$$

$$V_N(N, W) = \lambda \sum_{\theta} \pi(\theta) \frac{1}{\theta} \phi (N - Y(\theta))^{-\sigma} \quad (15)$$

$$+ \sum_{\theta, \hat{\theta}, \hat{\theta} \neq \theta} \mu(\theta, \hat{\theta}) \frac{1}{\theta} \phi \left[(N(\theta) - Y(\theta))^{-\sigma} - (N - Y(\hat{\theta}))^{-\sigma} \right] \quad (16)$$

The first equation is 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

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

A summation of the above equations over all θ implies that the right hand side of (15) 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 full information economy – $\frac{p_{FI}}{a} = R_{FI}$. Thus, it must be the case that $\frac{p_{FI}}{R_{FI}} = \frac{p_{PI}}{R_{PI}} = \frac{F_L(K, L)}{F_K(K, L)}$. It follows that the capital–labor ratios $\frac{K_{FI}}{L_{FI}}$ and $\frac{K_{PI}}{L_{PI}}$ in both economy must be equal. 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 5 *Suppose the source of heterogeneity is a shock to the taste for leisure and assumption 2 holds. Then, the return to capital in both 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 this increase in family size is given by aR_{PIt+1} in terms of consumption goods lost (also at date $t+1$). Hence, we must have

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

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 (13), N and $Y(\theta)$ always appear together. In the model with productivity shocks however, this is not necessarily true. When there is no friction, the above no arbitrage condition implies that the capital–labor ratio is the same as the one with the taste shock model (since $MRS_{Y, N} = \theta$). In the presence of private information, 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 turn to studying the properties of population dynamics when private information is present. First, recall our version of the Inverse Euler Equation:

$$E_t \left[\frac{1}{u'(c_{t+1})} \right] = \frac{\beta R_{PI}}{n_t^{1-\eta} u'(c_t)}.$$

Since the θ_t are i.i.d., and we have ‘resetting property’ (i.e., per capita promised utility to each chile is w^* and allocations are i.i.d), it follows that $\frac{1}{u'(c_{t+1})}$ are also i.i.d.. Thus, $E_t \left[\frac{1}{u'(c_{t+1})} \right] = E \left[\frac{1}{u'(c_{t+1})} \right]$ – the left hand side of this equation doesn’t depend on t . Hence, the same must be true of the right hand side. It follows that $n_t^{1-\eta} u'(c_t) \equiv B$ is independent of t . Substituting $u'(c_{t+1}) = B n_{t+1}^{\eta-1}$ we have:

$$E \left[n_{t+1}^{1-\eta} \right] = \beta R_{PI}.$$

Next, using our assumption that $\eta = 1 - \sigma$, it follows that when $\sigma > 1$, $\eta < 0$ (i.e., $n^{1-\eta}$ is convex). Hence, in this case, Jensen’s Inequality implies:

$$\beta R_{PI} = E \left[n_{t+1}^{1-\eta} \right] > [E(n_{t+1})]^{1-\eta} = \gamma_{PI}^{1-\eta}.$$

Similarly, when $\sigma < 1$, $0 < \eta < 1$ and so $n^{1-\eta}$ is concave. In this case:

$$\beta R_{PI} = E \left[n_{t+1}^{1-\eta} \right] < [E(n_{t+1})]^{1-\eta} = \gamma_{PI}^{1-\eta}.$$

As shown above, with full information, we have the property $\beta R_{FI} = \gamma_{FI}^{1-\eta}$ in all cases. Also, as we have shown in proposition 5 $R_{PI} = R_{FI}$. Using these facts, we obtain:

Proposition 6 *In the case with a goods cost and taste shocks:*

1. *If $\sigma = 1 - \eta < 1$, then $\gamma_{PI} \geq \gamma_{FI}$ 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 $\gamma_{PI} \leq \gamma_{FI}$ 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, $\gamma_{PI} > 1$.*

Remark 1: the no ‘population immiseration’ results in this proposition refer to immiseration in an absolute sense. As we saw in proposition 4 whenever $\beta R > 1$ there is immiseration in population relative to the trend $(\beta R)^{t/(1-\eta)}$.

Remark 2: As noted above, although $\gamma_{PI} >< 1$ is related to the asymptotic properties of N_t (i.e., $\rightarrow \infty, 0$ a.s.), they are not identical. (since N_t is a geometric random walk with drift, the asymptotic behavior of N_t is governed by $E(\log(n))$ not $E(n)$.)

From this, we can see that population grows on average as long as $\gamma_{FI} > 1$, i.e., as long as there population growth under full information, provided that either $\sigma < 1$, or uncertainty is not too large (this is essentially a continuity argument).

Remark 3: When the shock is to productivity (as opposed to 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

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

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 even 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 6 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 previous section we made the restrictive assumption that the private types of parents are taste shock to 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

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

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.

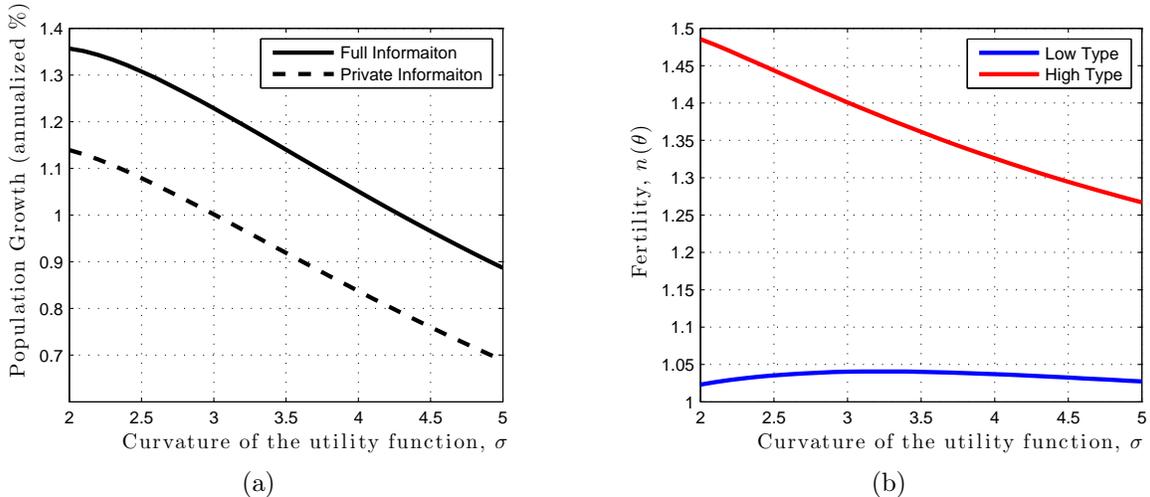


Figure 1: Panel (a) Population growth under full information (sold line) and private information (dashed line). Panel (b) Fertility for high (red) and low (blue) productivity shocks.

Following Jones and Schoonbroodt (2010) 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 Figure 1. In panel (a) we plot the annualized population growth for different values of the curvature parameter (risk aversion, σ). As we see, 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 assuming that the cost of children is primarily 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 makes the contracting problem convex. Because of this, we will need to make assumptions directly on binding incentive constraints (which are binding and which are slack at the solution) as well as convexity and differentiability of value functions.

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. 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, we also lose a result from the previous sections: the per capita allocations are 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 goods cost case.

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:¹⁶

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

s.t.

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

¹⁶The reduction of the problem from aggregate to per capita and its recursive formulation is identical to the goods cost case. So we omit that step.

$$\begin{aligned}
& u(c(\theta)) + h \left(1 - \frac{y(\theta)}{\theta} - bn(\theta) \right) + \beta n(\theta)^\eta w'(\theta) \geq \\
& u(c(\hat{\theta})) + h \left(1 - \frac{y(\hat{\theta})}{\theta} - bn(\hat{\theta}) \right) + \beta n(\hat{\theta})^\eta w'(\hat{\theta}) \quad \forall \hat{\theta} \leq \theta
\end{aligned}$$

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

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 an indirect method of having children 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 be 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.

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:

$$v(w'(\theta_I, w)) + bR\theta_I + aR = \eta w'(\theta_I, w) v'(w'(\theta_I, w)). \quad (17)$$

Note that that this equation is very similar to equation (10). 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 future consumption good.

In equation (17), $w'(\theta_I, w)$ is the per capita promised utility to child of a parent who has current promised utility of w and has received the highest productivity shock, θ_I . Using the same argument as before, we can 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 – an extreme version of social mobility holds. We summarize this in a proposition.¹⁷

Proposition 7 *Assume that v is continuously differentiable and that there is a unique solution to (17). 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 .*

Intuitively, the reason that the resetting property holds here mirrors the argument given above in the goods cost case. That is, since no ‘type’ wants to pretend to have $\theta = \theta_I$, the allocation for this type is marginally undistorted. Again, due to the homogeneity properties of the problem, 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_0 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 – $0 < \underline{n} \leq n(w, \theta) \leq \bar{n}$. Then, the fraction of descendants being assigned to w_0 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 $0 < \underline{n} \leq n(w, \theta) \leq \bar{n}$.

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

¹⁷For a detailed derivation and proof of this proposition see [Hosseini et al. \(2009\)](#)

zero follows from promise keeping.)

To this end, we show that as $w \rightarrow -\infty$, 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 8 *Suppose that V is continuously differentiable and strictly convex. Then there exists a $\underline{w}_i \in \mathbb{R}$, such that*

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

Proof. See Appendix B.1. ■

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 utility of leisure, $h(\cdot)$, is unbounded below. From this, it follows that as long as $w'(\theta, w)$ is continuous, w' will be bounded below on any closed set bounded away from 0.

From Proposition 8, it follows that as long as $w'(w, \theta)$ is continuous, w' will be bounded below on any closed set bounded away from 0 or we have the following corollary:

Corollary 3 *Suppose that v is continuously differentiable and that $\eta w v'(w) - v(w)$ is strictly decreasing. 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 $w'(w, \theta) \geq \underline{w}(\hat{w})$ for all $w \leq \hat{w}$ 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 of continuation utilities over \mathbb{R} , Ψ . Then, applying the policy functions to the measure Ψ , gives rise to a new measure over continuation utilities, $T\Psi$:

$$T(\Psi)(A) = \int_{\mathbb{R}} \sum_{\theta} \pi(\theta) \mathbf{1}_{\{(\theta, w); w'(\theta, w) \in A\}}(w, \theta) n(\theta, w) d\Psi(w) \quad (18)$$

$\forall A : \text{Borel Set in } \mathbb{R}$

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:

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

Now, suppose Ψ is a probability measure over continuation utilities. Ψ is said to be a stationary distribution if:

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

This is equivalent to having a constant distribution of per capita continuation utility along a balanced growth path in which 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 $\Psi \rightarrow \frac{T(\Psi)}{\gamma(\Psi)}$ 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, $[\underline{w}, \bar{w}]$, such that:

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

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

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

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

Proof. See Appendix [B.2](#). ■

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.

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 dynamic 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.*

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 and 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 ([Hosseini et al., 2010a](#))), but more work needs to be done.

References

- ALBANESI, S. AND C. SLEET (2006): “Dynamic Optimal Taxation with Private Information,” *Review of Economic Studies*, 73, 1–30.
- ALVAREZ, F. (1999): “Social Mobility: The Barro–Becker Children Meet the Laitner–Loury Dynasties,” *Review of Economic Dynamics*, 2, 65–103.
- ATKESON, A. AND R. LUCAS (1992): “On Efficient Distribution with Private Information.” *Review of Economic Studies*, 59, 427–53.

- BARRO, R. J. AND G. S. BECKER (1989): “Fertility Choice in an Endogenous Growth Model,” *Econometrica*, 57, 481–501.
- BECKER, G. S. AND R. J. BARRO (1988): “A Reformulation of the Economic Theory of Fertility,” *The Quarterly Journal of Economics*, 103, 1–25.
- BENVENISTE, L. M. AND J. A. SCHEINKMAN (1979): “On the Differentiability of the Value Function in Dynamic Models of Economics,” *Econometrica*, 47, 727–32.
- DE LA CROIX, D. AND M. DOEPKE (2003): “Inequality and Growth: Why Differential Fertility Matters,” *American Economic Review*, 93, 1091–1113.
- DUNFORD, N. AND J. SCHWARTZ (1958): *Linear Operators*, Interscience publishers.
- FARHI, E. AND I. WERNING (2007): “Inequality and Social Discounting,” *Journal of Political Economy*, 115, 365–402.
- (2010): “Progressive Estate Taxation,” *Quarterly Journal of Economics*, 125, 635–673.
- GOLOSOV, M., N. KOCHERLAKOTA, AND A. TSYVINSKI (2003): “Optimal Indirect and Capital Taxation,” *Review of Economic Studies*, 70, 569–587.
- GOLOSOV, M. AND A. TSYVINSKI (2006): “Designing Optimal Disability Insurance: A Case for Asset Testing,” *Journal of Political Economy*, 114, 257–279.
- (2007): “Optimal Taxation With Endogenous Insurance Markets,” *The Quarterly Journal of Economics*, 122, 487–534.
- GREEN, E. J. (1987): “Lending and the Smoothing of Uninsurable Income,” *Contractual Arrangements for Intertemporal Trade*, 3–25.
- HOSSEINI, R., L. E. JONES, AND A. SHOURIDEH (2009): “Risk Sharing, Inequality and Fertility,” *NBER Working Paper*, 15111.
- (2010a): “Risk Sharing, Inequality and Fertility,” Available at http://www.public.asu.edu/~rhossein/Roozbeh_Hosseini/Research_files/fertimpr_29.pdf.
- (2010b): “Risk Sharing, Inequality and Fertility, Supplemental Appendix,” Available at http://www.public.asu.edu/~rhossein/Roozbeh_Hosseini/Research_files/appendix2.pdf.

- JONES, L., R. MANUELLI, H. SIU, AND E. STACCHETTI (2005): “Fluctuations in convex models of endogenous growth, I: Growth effects,” *Review of Economic Dynamics*, 8, 780–804.
- JONES, L. AND A. SCHOONBROODT (2010): “Complements Versus Substitutes And Trends In Fertility Choice In Dynastic Models,” *International Economic Review*, 51, 671–699.
- KHAN, A. AND B. RAVIKUMAR (2001): “Growth and Risk-Sharing with Private Information,” *Journal of Monetary Economics*, 47, 499–521.
- KOCHERLAKOTA, N. (2005): “Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation,” *Econometrica*, 73, 1587–1621.
- KOCHERLAKOTA, N. R. (2010): *The New Dynamic Public Finance*, Princeton University Press.
- MILGROM, P. AND I. SEGAL (2002): “Envelope Theorems for Arbitrary Choice Sets,” *Econometrica*, 70, 583–601.
- PHELAN, C. (1998): “On the Long Run Implications of Repeated Moral Hazard,” *Journal of Economic Theory*, 79, 174–191.
- (2006): “Opportunity and Social Mobility,” *Review of Economic Studies*, 73, 487–504.
- RUDIN, W. (1991): *Functional Analysis*, McGraw-Hill.
- THOMAS, J. AND T. WORRALL (1990): “Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-agent Problem,” *Journal of Economic Theory*, 51, 367–390.

Appendix

A Relationship between $v(w)$ and $V(N, W)$

Lemma 2 *Assume that $V(\cdot, \cdot)$ is twice differentiable. Then $v(w)$ is convex and strictly increasing and $v(w) - \eta w v'(w)$ is increasing.*

Proof. We know from above that $V(N, W) = Nv(N^{-\eta}W)$. Strict concavity of $V(N, W)$ implies that $V_{WW} > 0, V_{NN} > 0, V_{WW}V_{NN} > V_{WN}^2$. We have

$$\begin{aligned} V_{WW} &= N^{1-2\eta}v''(N^{-\eta}W) = N^{1-2\eta}v''(w) \\ V_{WN} &= (1-\eta)N^{-\eta}v'(N^{-\eta}W) - \eta N^{-2\eta}Wv''(N^{-\eta}W) = N^{-\eta}((1-\eta)v'(w) - \eta w v''(w)) \\ V_{NN} &= \eta(\eta-1)N^{-\eta-1}Wv'(N^{-\eta}W) + \eta^2N^{-2\eta-1}W^2v''(N^{-\eta}W) \\ &= \eta N^{-1}w((\eta-1)v'(w) + \eta w v''(w)) \end{aligned}$$

After some algebra, we have

$$V_{WW}V_{NN} - V_{WN}^2 = N^{-2\eta}((\eta-1)v'(w) + \eta w v''(w))(1-\eta)v'(w)$$

Therefore, the strict convexity of $V(\cdot, \cdot)$ implies that:

$$\begin{aligned} v''(w) &> 0 \\ (\eta-1)v'(w) + \eta w v''(w) &= \frac{d}{dw}(\eta w v'(w) - v(w)) > 0 \\ v'(w) &> 0 \end{aligned}$$

■

B Main Proofs

B.1 Proof of Proposition 8

We first prove the following lemma:

Lemma 3 *Suppose that the value function in Pt is differentiable and 3 hold, then the value function and the policy functions satisfy the following properties:*

$$\begin{aligned}
\lim_{w \rightarrow -\infty} v(w) &= - \sum_{\theta_i} \pi(\theta_i) \theta_i \\
\lim_{w \rightarrow -\infty} c(\theta_i, w) &= 0 \\
\lim_{w \rightarrow -\infty} n(\theta_i, w) &= 0 \\
\lim_{w \rightarrow -\infty} l(\theta_i, w) &= 1
\end{aligned}$$

Proof. Consider the following set of function:

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

Moreover define the following mapping on S as

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

s.t.

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

$$\begin{aligned}
u(c(\theta_i)) + h(1 - l(\theta_i) - bn(\theta_i)) + \beta n(\theta_i)^\eta w'(\theta_i) &\geq u(c(\theta_j)) + h \left(1 - \frac{\theta_j l(\theta_j)}{\theta_i} - bn(\theta_j) \right) \\
&\quad + \beta n(\theta_j)^\eta w'(\theta_j), \forall i > j
\end{aligned}$$

$$\begin{aligned}
l(\theta_i) + bn(\theta_i) &\leq 1 \\
c(\theta_i), l(\theta_i), n(\theta_i) &\geq 0
\end{aligned}$$

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 Contraction Mapping Theorem we have that the fixed point of T belongs to S .

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

$$\liminf_{n \rightarrow \infty} T\hat{v}(w_n) \geq \sum_{\theta_j} \pi(\theta_j) [-\theta_j(1 - b\bar{n}_j) + \frac{1}{R}\bar{n}_j \left[-\sum_{\theta_k} \pi(\theta_k) \theta_k \right]]$$

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 $\bar{n}_j \geq 0$, we must have

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

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

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

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

$$\begin{aligned} c_m(\theta_i) &= u^{-1}(-\epsilon_m^\eta) \\ n_m(\theta_i) &= (I - i)^{\frac{1}{\eta}} \epsilon_m \\ w'_m(\theta_i) &= \tilde{w} < 0 \end{aligned}$$

If $h(\cdot)$ is bounded above and below, define

$$l_m(\theta_i) = 1 - \epsilon_m - bn_m(\theta_i)$$

By construction,

$$\begin{aligned} c_m(\theta_i) &\rightarrow 0 \\ n_m(\theta_i) &\rightarrow 0 \\ l_m(\theta_i) &\rightarrow 1 \end{aligned}$$

Moreover,

$$u(c_m(\theta_j)) + \beta n_m(\theta_j)^\eta w'_m(\theta_j) - u(c_m(\theta_i)) - \beta n_m(\theta_i)^\eta w'_m(\theta_i) = \beta \tilde{w} \epsilon_m^\eta (i - j), : \forall j > i$$

This expression converges to ∞ and therefore, since $h(\cdot)$ is bounded above and below for m large enough, the allocations are incentive compatible.

When, h is unbounded below, since the utility of deviation is bounded away from $-\infty$, it is possible to construct a sequence for l_m that converges to 1. Find $l_m(\theta_i)$ such that

$$h(1 - l_m(\theta_i) - b n_m(\theta_i)) = \frac{1}{2} \tilde{w} \epsilon_m^\eta$$

Hence, we have that

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

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

$$\begin{aligned} u(c_m(\theta_j)) + h(1 - l_m(\theta_j) - b n_m(\theta_j)) + \beta n_m(\theta_j)^\eta w'_m(\theta_j) \\ - u(c_m(\theta_i)) - \beta n_m(\theta_i)^\eta w'_m(\theta_i) \geq h(1 - \frac{\theta_i l_m(\theta_i)}{\theta_j} - b n_m(\theta_i)) \end{aligned}$$

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

$$\hat{w}_m = \epsilon_n^\eta \left[-1 + \beta \sum A_j^\eta \tilde{w} \right] + \sum_k \pi_j [h(1 - l_m(\theta_j) - b n_m(\theta_j))]$$

It is clear that \hat{w}_m 's converge to $-\infty$ and the allocation's cost converges to $-\sum_j \pi_j \theta_j$. Now since \hat{w}_m and w_n converge to $-\infty$, there exists subsequences \hat{w}_{m_k} and w_{n_k} such that $\hat{w}_{m_k} \geq w_{n_k}$ and therefore by optimality:

$$\sum_j \pi_j \left[c_{m_k}(\theta_j) - \theta_j l_{m_k}(\theta_j) + \frac{1}{R} n_{m_k}(\theta_j) \hat{v}(\tilde{w}) \right] \geq T \hat{v}(w_{n_k})$$

and therefore,

$$-\sum_k \pi_k \theta_k \geq \liminf_{n \rightarrow \infty} T \hat{v}(w_n) > -\sum_k \pi_k \theta_k$$

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

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

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

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

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

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

By Lemma 3, we must have:

$$\begin{aligned} \lim_{w \rightarrow -\infty} c(\theta_j, w) &= 0 \\ \lim_{w \rightarrow -\infty} n(\theta_j, w) &= 0 \\ \lim_{w \rightarrow -\infty} l(\theta_j, w) &= 1 \end{aligned}$$

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‘

$$\begin{aligned}\lambda(w) &= \sum_j \frac{\pi_j}{u'(c(\theta_j, w))} < \frac{\epsilon}{M} \\ \pi(\theta_I)\lambda(w) + \sum_{j < I} \mu(\theta_n, \theta_j; w) &= \frac{\pi(\theta_I)}{u'(c(\theta_I, w))} < \frac{\epsilon}{M} \\ \Rightarrow \mu(\theta_I, \theta_j; w) &< \frac{\epsilon}{M}\end{aligned}$$

In addition,

$$\begin{aligned}\pi(\theta_{I-1})\lambda(w) + \sum_{j < I-1} \mu(\theta_{I-1}, \theta_j; w) - \mu(\theta_I, \theta_{I-1}; w) &= \frac{\pi(\theta_{I-1})}{u'(c(\theta_{I-1}, w))} < \frac{\epsilon}{M} \\ \Rightarrow \mu(\theta_{I-1}, \theta_j; w) &< \frac{2\epsilon}{M}\end{aligned}$$

By an induction argument, we have

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

where $z_{I-1} = 1$, $z_{I-2} = 2$, $z_{I-i} = z_{I-1} + \dots + z_{I-i+1} + 1$. 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 7 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

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

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

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

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

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

Continuity of v' implies that

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

This completes the proof.

B.2 Proof of Theorem 1

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

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

We proceed by defining \underline{w} and \bar{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, $\lim_{w \rightarrow 0} g(w)$ exists. In a slight abuse of notation, let $g(0) = \lim_{w \rightarrow 0} g(w)$. Further, since $w < 0$, it follows that $g(w) < h(1)$ and hence, $g(0) \leq h(1)$. In fact, $g(0) = h(1)$. To see this, consider the sequences $w_k = -1/k$, $n_k = k^{1/(2n)}$. Then, for k large enough, n_k is feasible and therefore, $g(w_k) \geq h(1 - bn_k) + \beta n_k^\eta w_k$. Hence,

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

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 \bar{w} be this fixed point. Since $g(w) < w$ in a neighborhood of 0, it follows that $\bar{w} < 0$. Following Corollary 3, choose $\underline{w} = \underline{w}(\bar{w})$.

With these definitions, it follows that, as long as a solution to the functional equation exists for all $w \in [\underline{w}, \bar{w}]$, $w'(w, \theta) \in [\underline{w}, \bar{w}]$. 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 $[\underline{w}, \bar{w}]$ 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. Since $[\underline{w}, \bar{w}]$ is compact in \mathbb{R} , by Riesz Representation Theorem (Dunford and Schwartz (1958), IV.6.3), the space of regular measures is isomorphic to the space $C^*([\underline{w}, \bar{w}])$, the dual of the space of bounded continuous functions over $[\underline{w}, \bar{w}]$. Moreover, by Banach-Alaoglu Theorem (Rudin (1991), Theorem 3.15), the set $\{\Psi \in C^*([\underline{w}, \bar{w}]); \|\Psi\| \leq k\}$ is a compact set in the weak-* topology for any $k > 0$. Equivalently the set of regular measures, Ψ , with $\|\Psi\| \leq 1$, is compact. Since non-negativity and full measure on $[\underline{w}, \bar{w}]$ are closed

restrictions, we must have that the set

$$\{\Psi : \Psi \text{ a regular measure on } [\underline{w}, \bar{w}], \Psi([\underline{w}, \bar{w}]) = 1, \Psi \geq 0\}$$

is compact in weak-* topology.

By definition,

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

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 $[\underline{w}, \bar{w}]$ (since otherwise utility would be $-\infty$). From this, it follows that T is continuous in Ψ . Moreover,

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

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

Therefore, the function

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

is continuous. Therefore, by Schauder-Tychonoff Theorem (Dunford and Schwartz (1958), V.10.5), \hat{T} has a fixed point $\Psi^* \in M([\underline{w}, \bar{w}])$. ■

C An Alternative Proof for Proposition 1

In this section we state and prove the main result of section 2 (proposition 1) without relying on convexity and differentiability.

Suppose the size of the initial dynasty is N_{-1} and the promised utility to the head of dynasty is W_0 . Consider the minimization problem (P) in section

$$\min_{\{C_t(\theta^t), Y_t(\theta^t), N_{t+1}(\theta^t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \sum_{\theta^t} \frac{1}{R^t} \pi(\theta^t) [Y_t(\theta^t) - C_t(\theta^t) - aN_{t+1}(\theta^t)] \quad (19)$$

subject to

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

and

$$\begin{aligned} \sum_{t=0}^{\infty} \sum_{\theta^t} \beta^t \pi(\theta^t) N_t(\theta^{t-1})^\eta \left[u \left(\frac{C_t(\theta^t)}{N_t(\theta^{t-1})} \right) + h \left(\frac{Y_t(\theta^t)}{N_t(\theta^{t-1})}, \theta_t \right) \right] \geq \\ \sum_{t=0}^{\infty} \sum_{\theta^t} \beta^t \pi(\theta^t) N_t(\sigma_{t-1}(\theta^{t-1}))^\eta \left[u \left(\frac{C_t(\sigma_t(\theta^t))}{N_t(\sigma_{t-1}(\theta^{t-1}))} \right) + h \left(\frac{Y_t(\sigma_t(\theta^t))}{N_t(\sigma_{t-1}(\theta^{t-1}))}, \theta_t \right) \right] \end{aligned}$$

for all $\sigma \in \Sigma_0$

Let $C_t^*(\theta^t, N_{-1}, W_0)$, $N_{t+1}^*(\theta^t, N_{-1}, W_0)$ and $Y_t^*(\theta^t, N_{-1}, W_0)$ be the solution to this problem. We call this solution the *constrained efficient allocations*. Define

$$c_t^*(\theta^t, N_{-1}, W_0) \equiv \frac{C_t^*(\theta^t, N_{-1}, W_0)}{N_t^*(\theta^{t-1}, N_{-1}, W_0)},$$

$$y_t^*(\theta^t, N_{-1}, W_0) \equiv \frac{Y_t^*(\theta^t, N_{-1}, W_0)}{N_t^*(\theta^{t-1}, N_{-1}, W_0)}$$

and

$$n_t^*(\theta^t, N_{-1}, W_0) \equiv \frac{N_{t+1}^*(\theta^t, N_{-1}, W_0)}{N_t^*(\theta^{t-1}, N_{-1}, W_0)}$$

to be per capita consumption, effective unit of labor supply and per capita fertility for generation after history θ^t . We first prove that after the initial generation, the per capita allocations of consumption and effective labor supply do not depend on W_0 (the promised utility promised to the initial generation).

Proposition 9 *Let $c_t^*(\theta^t, N_{-1}, W_0)$, $y_t^*(\theta^t, N_{-1}, W_0)$ and $n_t^*(\theta^t, N_{-1}, W_0)$ be constrained efficient per capita consumption, effective unit of labor supply and per capita fertility. Then, $c_t^*(\theta^t, N_{-1}, W_0)$, $y_t^*(\theta^t, N_{-1}, W_0)$ and $n_t^*(\theta^t, N_{-1}, W_0)$ are i.i.d. and are independent of W_0 for any generation $t > 0$.*

Proof. Let $C_t^*(\theta^t, N_{-1}, W_0)$, $Y_t^*(\theta^t, N_{-1}, W_0)$ and $N_{t+1}^*(\theta^t, N_{-1}, W_0)$ be solution to the prob-

lem (19). For any history θ^{t-1} define

$$W_t^*(\theta^{t-1}, W_0) = \sum_{s=t}^{\infty} \sum_{\theta^s | \theta^{t-1}} \beta^s \pi(\theta^s) N_s^*(\theta^{s-1}, N_{-1}, W_0)^\eta \left[u \left(\frac{C_s^*(\theta^s, N_{-1}, W_0)}{N_s^*(\theta^{s-1}, N_{-1}, W_0)} \right) + h \left(\frac{Y_s^*(\theta^s, N_{-1}, W_0)}{N_s^*(\theta^{s-1}, N_{-1}, W_0)}, \theta_s \right) \right].$$

The allocation $\{C_s^*(\theta^s, N_{-1}, W_0), Y_s^*(\theta^s, N_{-1}, W_0), N_{s+1}^*(\theta^s, N_{-1}, W_0)\}_{\theta^s | \theta^{t-1}}$ must be a solution to the following cost minimization problem for any history θ^{t-1} :

$$\min_{\{C_s(\theta^s), Y_s(\theta^s), N_{s+1}(\theta^s)\}_{\theta^s | \theta^{t-1}}} -aN_1(\theta_0) + \sum_{s=t}^{\infty} \sum_{\theta^s | \theta^{t-1}} \frac{1}{R^s} \pi(\theta^s) [Y_s(\theta^s) - C_s(\theta^s) - aN_{s+1}(\theta^s)] \quad (20)$$

subject to

$$\sum_{s=t}^{\infty} \sum_{\theta^s | \theta^{t-1}} \beta^s \pi(\theta^s) N_s(\theta^{s-1})^\eta \left[u \left(\frac{C_s(\theta^s)}{N_s(\theta^{s-1})} \right) + h \left(\frac{Y_s(\theta^s)}{N_s(\theta^{s-1})}, \theta_s \right) \right] = W_t^*(\theta^{t-1}, W_0)$$

and

$$\sum_{s=t}^{\infty} \sum_{\theta^s | \theta^{t-1}} \beta^s \pi(\theta^s) N_s(\theta^{s-1})^\eta \left[u \left(\frac{C_s(\theta^s)}{N_s(\theta^{s-1})} \right) + h \left(\frac{Y_s(\theta^s)}{N_s(\theta^{s-1})}, \theta_s \right) \right] \geq \sum_{s=t}^{\infty} \sum_{\sigma^s | \theta^{t-1}} \beta^s \pi(\sigma^s) N_s(\sigma_{s-1}(\theta^{s-1} | \theta^{t-1}))^\eta \left[u \left(\frac{C_s(\sigma_s(\theta^s | \theta^{t-1}))}{N_s(\sigma_{s-1}(\theta^{s-1} | \theta^{t-1}))} \right) + h \left(\frac{Y_s(\sigma_s(\theta^s | \theta^{t-1}))}{N_s(\sigma_{s-1}(\theta^{s-1} | \theta^{t-1}))}, \theta_s \right) \right] \quad \text{for all } \sigma \in \Sigma_0.$$

Suppose not. Then we can replace the solution to (19) (after each history θ^{t-1}) by the solution to this problem (for each history θ^{t-1}) and reduce the cost to the planner in problem (19). Note that this allocation satisfies promise keeping in problem (19) by construction. It is also incentive compatible. This is a contradiction.

The objective in (20) is homogenous of degree one and the constraint set is homogenous of degree $1/\eta$. It is straight-forward to show that after any history θ^{t-1} :

$$\begin{aligned} C_s^*(\theta^s, N_{-1}, W_0) &= \bar{C}_s(\theta^s, N_{-1}) W_t^*(\theta^{t-1}, W_0)^{1/\eta} \\ Y_s^*(\theta^s, N_{-1}, W_0) &= \bar{Y}_s(\theta^s, N_{-1}) W_t^*(\theta^{t-1}, W_0)^{1/\eta} \\ N_{s+1}^*(\theta^s, N_{-1}, W_0) &= \bar{N}_{s+1}(\theta^s, N_{-1}) W_t^*(\theta^{t-1}, W_0)^{1/\eta}. \end{aligned}$$

Therefore after any history θ^{t-1} :

$$c_s^*(\theta^s, N_{-1}, W_0) = \frac{\bar{C}_s(\theta^s, N_{-1})}{\bar{N}_s(\theta^{s-1}, N_{-1})}$$

$$y_s^*(\theta^s, N_{-1}, W_0) = \frac{\bar{Y}_s(\theta^s, N_{-1})}{\bar{N}_s(\theta^{s-1}, N_{-1})},$$

and

$$n_s^*(\theta^s, N_{-1}, W_0) = \frac{\bar{N}_{s+1}(\theta^s, N_{-1})}{\bar{N}_s(\theta^{s-1}, N_{-1})}.$$

Note that the problem (20) depends on history θ^{t-1} and W_0 only through $W_t^*(\theta^{t-1}, W_0)$. Therefore, its solution depends on history θ^{t-1} and W_0 only through $W_t^*(\theta^{t-1}, W_0)$. This implies that $c_s^*(\theta^s, N_{-1}, W_0)$, $y_s^*(\theta^s, N_{-1}, W_0)$ and $n_s^*(\theta^s, N_{-1}, W_0)$ do not depend on θ^{t-1} and W_0 for $s \geq t$. In particular $c_t^*(\theta^t, N_{-1}, W_0)$, $y_t^*(\theta^t, N_{-1}, W_0)$ and $n_t^*(\theta^t, N_{-1}, W_0)$ depend only on θ_t (and N_{-1}), i.e., they are i.i.d. ■