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

Online Appendix

Roozbeh Hosseini
Arizona State University
rhossein@asu.edu

Larry E. Jones
University of Minnesota
lej@umn.edu

Ali Shourideh
Wharton School
shouride@wharton.upenn.edu

April 15, 2013

Contents

1 Introduction **2**

2 A Two Period Model **2**

2.1 Resetting Property in a Two Period Example 2

2.2 A Non-Homothetic Example 3

3 Persistent Shocks **5**

4 Alternative Proof of Stationarity Based on Sequence Problem **8**

5 Aggregate and Per Capita Recursive Formulation **10**

6 Dynasty Size vs. Total Population **12**

7 Implementation **14**

References **17**

1 Introduction

This is the Online (or Supplementary) Appendix for the paper, "Optimal Contracting with Dynastic Altruism: Family Size and Per Capita Consumption," by Roozbeh Hosseini, Larry E. Jones and Ali Shourideh. It contains various related examples and results not included in either the paper or it's Appendix. Included are:

1. A detailed examination of a two period example including a discussion of the role of homothetic preferences;
2. A short discussion of what is known when the shocks in the model are persistent;
3. An alternative proof of the existence of a stationary distribution in the goods cost case from primitives on the sequence problem;
4. A discussion of the homotheticity properties of the aggregate version of the cost minimization problem showing that it can be reduced to the per capita problem;
5. A short, preliminary discussion on what is known about implementation of the optimal contract using an income tax system.

2 A Two Period Model

In this section, we study a two period version of the model with a more generalized specification of preferences.

2.1 The Resetting Property in a Two Period Example

In this section, we study a two period example of the model outlined in the paper. To simplify, we assume that there is no labor supply in the second period, i.e., of the children. Because of this, the utility of a parent is given by $U(c_0, l, n, \theta) + \beta n^\eta u(c_1)$,where l is labor supply of the parent and θ is parent's productivity. For example, when there is only a time cost of children, $U(c_0, y, n, \theta) = u(c_0) + h(1 - l - bn)$. Suppose that having children has an additional cost $k(n, \theta)$ in terms of parent's consumption good. The problem of the parent is given by

$$V(A_0) = \max_{c_0, C_1, n, l \geq 0} U(c_0, l, n, \theta) + n^\eta u\left(\frac{C_1}{n}\right)$$

s.t.

$$c_0 + C_1 + k(n, \theta) \leq A_0 + \theta l \tag{1}$$

The solution of the problem (1) has the following property:

Lemma 1 *Suppose that the solution to problem (1) is interior. Then*

$$\eta \frac{u\left(\frac{C_1}{n}\right)}{u'\left(\frac{C_1}{n}\right)} - \frac{C_1}{n} = k_n(n,\theta) + \theta \frac{U_n(c_0,l,n,\theta)}{U_l(c_0,l,n,\theta)}. \tag{2}$$

Proof. The first order conditions for the above problem are given by

$$\begin{aligned} U_c &= \lambda \\ -U_l &= \theta\lambda \quad (3) \\ U_n + \beta\eta n^{\eta-1} u\left(\frac{C_1}{n}\right) - \beta n^{\eta-1}\frac{C_1}{n} u'\left(\frac{C_1}{n}\right) &= \lambda k_n(n,\theta) \quad (4) \\ \beta n^{\eta-1} u'\left(\frac{C_1}{n}\right) &= \lambda \quad (5) \end{aligned}$$

where λ is the multiplier on budget constraint. By combining (3), (4) and (5), we get the following:

$$-\theta\frac{U_n}{U_l} + \eta \frac{u\left(\frac{C_1}{n}\right)}{u'\left(\frac{C_1}{n}\right)} - \frac{C_1}{n} = k_n(n,\theta)$$

which implies the claim. ■

Note that if, $k_n = a$ and $U_n/U_l = b$, (2) becomes

$$\eta \frac{u\left(\frac{C_1}{n}\right)}{u'\left(\frac{C_1}{n}\right)} - \frac{C_1}{n} = a + b\theta.$$

Given the above characterization for $c_1 = C_1/n$, one can state the following result:

Remark 1 *Consumption of each child, $c_1 = C_1/n$, is independent of parent's wealth, A_0, if and only if the function*

$$k_n(n,\theta) + \theta\frac{U_n(c,l,n,\theta)}{U_l(c,l,n,\theta)}$$

is only a function of θ and does not depend on allocations (c_0, y, n).

2.2 A Non-Homothetic Example

As it is argued in Alvarez (1999) and we demonstrated above, Barro-Becker style dynastic altruism implicitly has a homotheticity property which, if coupled with linear cost of raising children, delivers a very stark result: The consumption of children is independent of parent's wealth (or as we discussed in the paper, independent of promised utility to the parent).

Here we use the two period example presented above to discuss what will happen if we drop the homotheticity assumption. We will show that each child's consumption is no longer constant and depends on parent's wealth. However, for very general assumption there is a lower bound on each child's consumption. In other words, we show that for very general assumptions on preferences the consumption allocation of each child does not converge to zero as the parent's asset becomes smaller and smaller.

Suppose parents have the following, non-homothetic preferences:

$$U\left(c_0, l, n, \theta\right) + \beta g(n) u\left(C_1/n\right)$$

and assume that $g(n)u(c_1)$ is strictly increasing, strictly concave, differentiable and

$$\frac{ng'(n)/g(n)}{c_1 u'\left(c_1\right)/u\left(c_1\right)} < D < \infty \quad \forall c_1, n,$$

i.e., the (negative of) elasticity of substitution between c_1 and n is uniformly bounded above.

Parents solve the following decision problem:

$$V\left(A_0\right) = \max_{c_0, C_1, n, l \geq 0} U\left(c_0, l, n, \theta\right) + g(n) u\left(\frac{C_1}{n}\right)$$

s.t.

$$c_0 + C_1 + k\left(n, \theta\right) \leq K_0 + \theta l \tag{6}$$

Then, the analog of equation (2) is:

$$n \frac{g'\left(n\right)}{g\left(n\right)} \frac{u\left(\frac{C_1}{n}\right)}{u'\left(\frac{C_1}{n}\right)} - \frac{C_1}{n} = k_n\left(n, \theta\right) + \theta \frac{U_n(c, l, n, \theta)}{U_l(c, l, n, \theta)} \tag{7}$$

Replace $c_1 = C_1/n$ and divide both sides by c_1

$$\frac{ng'(n)/g(n)}{c_1 u'\left(c_1\right)/u\left(c_1\right)} = 1 + \frac{k_n + \theta U_n/U_l}{c_1}$$

The term $k_n + \theta U_n/U_l$ is the marginal cost of having a child in terms of the parent's consumption good. The first term is the marginal good's cost and the second term is the marginal time cost (or utility cost).[1] Suppose this term is bounded below i.e., assume that

$$k_n\left(n, \theta\right) + \theta \frac{U_n(c_0, l, n, \theta)}{U_l(c_0, l, n, \theta)} > d > 0 \quad \forall c_0, l, n.$$

[1] Note that $\theta U_n/U_l = -U_n/U_c$.

Now consider moving parents' wealth towards $-\theta$. As parent's become poorer, they choose lower consumption, c_0, less leisure (more l), fewer kids,n, and less consumption for kids, C_1. However, if the assumptions above hold the consumption of each child will be bounded below, away from zero as parent's wealth moves towards $-\theta$.

Homotheticity in the utility function and linear cost is required to get stark result that children's consumption is independent of parent's wealth (or promised utility). But they are not required for keeping each child's consumption away from misery. Rather what is required is that income expansion paths in (C_2, n) space should have a slope that is bounded away from zero. The example given below illustrates this point.

Example. Suppose $g(n) = n^{\eta_1} + An^{\eta_2}$ with $0 > \eta_1 > \eta_2$, $k(n, \theta) = 0$ and $U_n/U_l = b$. In this case, equation (7) becomes:

$$\frac{\eta_1 n^{\eta_1 - 1} + \eta_2 A n^{\eta_2 - 1}}{n^{\eta_1 - 1} + A n^{\eta_2 - 1}} \frac{u(c_2)}{u'(c_2)} - c_2 = \theta b.$$

Now suppose that A_0 converges to $-\theta$. In this case, one can argue that n has to converge to zero. To see this, c_1 has to converge to 0 because of the budget constraint. This violates the equation above. Note that

$$\lim_{n \to 0} \frac{\eta_1 n^{\eta_1 - 1} + \eta_2 A n^{\eta_2 - 1}}{n^{\eta_1 - 1} + A n^{\eta_2 - 1}} = \eta_2.$$

This means that as A_0 converges to $-\theta$, the above equation becomes:

$$\eta_2 \frac{u(c_1)}{u'(c_1)} - c_1 = \theta b$$

which implies that c_1 is bounded away from 0. Income expansion paths for this example are given in Figure 1. Note that at $A_0 = -\theta$, $c_2(W_0, \theta_H)$ is the slope of the income expansion path at the origin which is positive. Moreover, $\frac{C_1}{n}$ is bounded away from zero for all points on the curve.

3 Persistent Shocks

Here, we briefly discuss the long-run implications of the model when shocks are persistent. Unfortunately, due to the known complications of dynamic contracting with persistent pri-

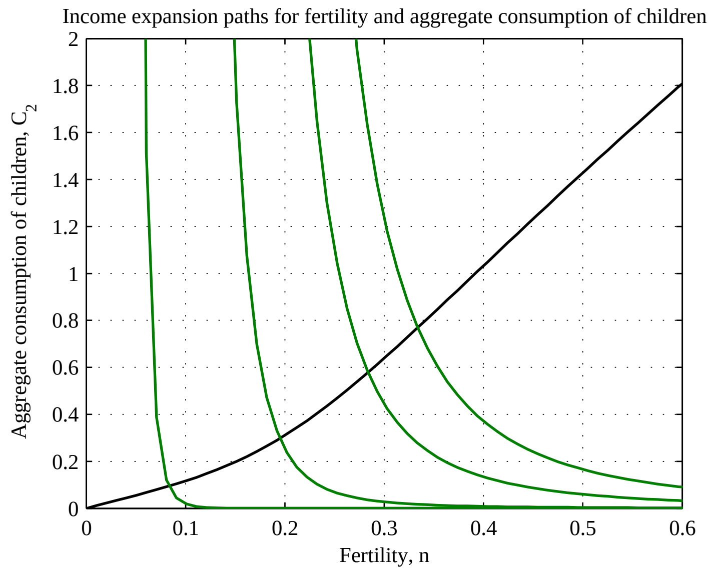

Figure 1: Income expansion path in an example with non-homothetic formulation. The slope of the income expansion path is per capita consumption. Example for $g(n) = n^{\eta_1} + An^{\eta_2}$.

vate information, we cannot fully extend our results to persistent shocks. [2] However, we are able to provide a preliminary analysis that suggests that our result on long run properties of per capita allocations hold.

We consider an example in which private shocks affect the utility of leisure and there is a goods cost of rearing children. We assume that the shock takes on two values, $\{\theta_L < \theta_H\}$, and it follows a markov process of order 1 with transition probability $\pi(\theta|\theta_-)$. Given, this we can use Fernandes and Phelan (2000)'s approach to formulate this problem recursively:

$$v(w, w_-; \theta_-) = \min \sum_{i=H,L} \pi(\theta_i|\theta_-) \left[c_i + an_i - y_i + \frac{1}{R} n_i v\left(w'_i, w'_{-,i}; \theta_i\right) \right]$$

[2] See Fernandes and Phelan (2000) for a recursive representation of a dynamic contracting problem with persistent shocks. Pavan et al. (2009), Farhi and Werning (2010), as well as Golosov et al. (2010), extend this approach. We are unaware of a general proof of immiseration with persistent private information in environments without endogenous fertility decision. There are some examples solved in Williams (2009) and Zhang (2009).

subject to

$$\begin{aligned}
\sum_{i=L,H} \pi\left(\theta_i|\theta_-\right)\left[u\left(c_i\right)+\frac{1}{\theta_i}h\left(y_i\right)+\beta n_i^{\eta} w_i'\right] &= w \\
\sum_{i=L,H} \pi\left(\theta_i|\theta_-^c\right)\left[u\left(c_i\right)+\frac{1}{\theta_i}h\left(y_i\right)+\beta n_i^{\eta} w_i'\right] &= w_-,\ \theta_-^c \neq \theta_- \\
u\left(c_i\right)+\frac{1}{\theta_i}h\left(y_i\right)+\beta n_i^{\eta} w_i' \geq & \\
u\left(c_j\right)+\frac{1}{\theta_i}h\left(y_j\right)+\beta n_j^{\eta} w_{-,i}' &
\end{aligned}$$

where w_-, as defined by Fernandes and Phelan (2000), is the threat-keeping utility. It is the continuation utility that type θ_-^c receives if he pretends to be of type θ and tells the truth from then onward. In what follows, we make the following assumptions:

1. Only donward incentive constraints are binding, i.e., we only need to consider θ_H pretending to be θ_L;
2. The value function is differentiable;
3. Relative fertility: for some $\varepsilon > 0$, $\frac{n_H(w,w_-;\theta_-)}{n_L(w,w_-;\theta_-)} > \varepsilon$, for all $(w, w_-; \theta_-)$.

The above assumptions, although, on endogenous variables, help us illustrate our main point. Furthermore, they all hold with i.i.d. shocks. The first assumption implies that threat-keeping utility when the previously announced type was been θ_H, is irrelevant. In other words, in $v(w, w_-; \theta_H)$, we can drop w_-. In each state, the following necessary conditions hold:

$$\begin{aligned}
v\left(w_H';\theta_H\right)-\eta w_H' v'\left(w_H';\theta_H\right) &= aR \\
v\left(w_L', w_{-,L}';\theta_L\right)-\eta w_L' v_1\left(w_L', w_{-,L}';\theta_L\right)-\eta w_{-,L}' v_2\left(w_L', w_{-,L}';\theta_L\right) &= aR
\end{aligned}$$

where we have suppressed the dependence on the state (w, w_-, θ_-). These equations are equivalent to resetting in the i.i.d. case. The first equation is identical to equation (10) in the paper and states that promised utility following a high value of the shock is independent of history. The second equation, however, defines a locus of points (w, w_-) that promise and threat keeping utilities should belong to following a low value of the shock. Together with the third assumption above, they imply that a stationary distribution exists for per capita allocation in the long-run.

As we see, when shocks are persistent, we no longer have complete independence from history in the optimal per capita allocations that holds in the *i.i.d.* case. However, by using

$w'_{-,L}$, the planner can provide incentives for the truthful revelation of types. The above analysis, although not fully derived from first principles, suggests that our main result on stationarity of per capita allocations holds with persistence.

4 Alternative Proof of Stationarity Based on the Sequence Problem

In this section we state and prove the main result of Section 2.1 (Proposition 1) without relying on convexity and differentiability.

Suppose the size of the initial dynasty is N_{-1} and that promised utility to the head of the dynasty is W_0. Consider the minimization problem (1) in Section 2.1 of the paper:

$$\min_{\{C_t(\theta^t), Y_t(\theta^t), N_{t+1}(\theta^t)\}_{t=0}^{\infty}} \sum_{t=0}^{\infty} \sum_{\theta^t} \frac{1}{R^t} \pi\left(\theta^t\right) \left[C_t\left(\theta^t\right) + aN_{t+1}\left(\theta^t\right) - Y_t\left(\theta^t\right)\right] \tag{8}$$

subject to

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

and

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

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 allocation*. 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

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

These are the per capita consumption, effective units of labor supply and per capita fertility after history θ^t. We 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 of the initial generation). Hence the per capita allocations have stationary distribution.

Proposition 1 *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 the constrained efficient per capita consumption, effective units 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 the solution to the problem (8). For any history θ^{t-1} define

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

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

$$\min_{\{C_s(\theta^s), Y_s(\theta^s), N_{s+1}(\theta^s)\}_{\theta^s|\theta^{t-1}}} aN_1\left(\theta_0\right) + \sum_{s=t}^{\infty} \sum_{\theta^s|\theta^{t-1}} \frac{1}{R^s} \pi\left(\theta^s\right)\left[C_s\left(\theta^s\right) + aN_{s+1}\left(\theta^s\right) - Y_s\left(\theta^s\right)\right] \quad (9)$$

subject to

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

and

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

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

for all $\sigma \in \Sigma_0$.

To see this, suppose that this doesn't hold. Then, we can replace the solution to (8) (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 (8). Note that this allocation satisfies promise keeping in problem (8) by construction. It is also incentive compatible. This is a contradiction.

The objective in (9) 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}:

$$C_s^*(\theta^s, N_{-1}, W_0) = \bar{C}_s(\theta^s, N_{-1}) W_t^* \left(\theta^{t-1}, W_0\right)^{1/\eta}$$

$$Y_s^*(\theta^s, N_{-1}, W_0) = \bar{Y}_s(\theta^s, N_{-1}) W_t^* \left(\theta^{t-1}, W_0\right)^{1/\eta}$$

$$N_{s+1}^*(\theta^s, N_{-1}, W_0) = \bar{N}_{s+1}(\theta^s, N_{-1}) W_t^* \left(\theta^{t-1}, W_0\right)^{1/\eta}.$$

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 (9) 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. ■

5 Aggregate and Per Capita Recursive Formulations

In this section, we describe how convexity of the aggregate value function $V(\cdot, \cdot)$ translates into the properties of the per capita value function $v(\cdot)$:

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

Proof. It can be shown that $V(N,W) = Nv(N^{-\eta}W)$. Strict convexity of $V(N,W)$ implies that $V_{WW} > 0, V_{NN} > 0, V_{WW}V_{NN} > V_{WN}^2$. We have:

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

After some algebra, we have

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

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

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

■

Furthermore, we show that the solution to the functional equation (P) in the main text of the paper satisfies:

$$V(N,W) = Nv(N^{\eta}W)$$

To see this, let $\hat{v}(N,w) = N^{-1}V(N,N^{\eta}W)$. Then, $\hat{v}$ must satisfy the following functional equation:

$$N\hat{v}(N,w) = \min_{C(\theta),Y(\theta),N'(\theta),w'(\theta)} \sum_{\theta} \pi(\theta)\left[C(\theta) + aN'(\theta) - Y(\theta) + \frac{1}{R}N'(\theta)\,\hat{v}\left(N'(\theta), w'(\theta)\right)\right]$$

subject to

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

and

$$N^{\eta}\left\{u\left(\frac{C(\theta)}{N}\right)+h\left(\frac{Y(\theta)}{N},\theta\right)\right\}+\beta N'(\theta)^{\eta}w'(\theta)\geq N^{\eta}\left\{u\left(\frac{C(\hat{\theta})}{N}\right)+h\left(\frac{Y(\hat{\theta})}{N},\theta\right)\right\}$$
$$+\beta N'\left(\hat{\theta}\right)^{\eta}w'(\hat{\theta})\qquad\forall\theta,\hat{\theta}\in\Theta.$$

We can rewrite the above in terms of per capita variables and arrive at the following functional equation – the objective has been divided by N and the constrains by N^{η}:

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

subject to

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

and

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

The above manipulations show that the transformation associated with the functional equation (P) in the paper maps the set of the functions of the form $N\hat{v}(N^{-\eta}w)$ to itself. Since this set is closed, the solution to (P) belongs to this set.

6 Dynasty Size vs. Total Population

In Corollary 2 of Section 3.1 we show that if $\beta R>1$, and the variance of the shocks is small, the size of all dynasties grow without bound, almost surely. Here, we provide an upper bound on the growth rate of dynasty size when there is only goods cost for having kids. We also show that if $\eta>0$, then almost all dynasties grow slower than total population. Hence, almost all dynasties vanish relative to total population (even though the number of people in each dynasty may grow without bound).

Lemma 3 $\frac{N_{t+1}\left(\theta^{t}\right)^{1-\eta}}{(\beta R)^{t+1}}\overset{a.s.}{\to}0.$

Proof. Consider the Inverse Euler Equation

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

Divide both sides by $(\beta R)^{t+1}$ to obtain:

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

Therefore $X_t = \frac{N_{t+1}\left(\theta^{t}\right)^{1-\eta}}{(\beta R)^{t+1}u'\left(c_{t+1}\left(\theta^{t+1}\right)\right)}$ is a non-negative martingale. Hence, there must exist a nonnegative random variable X_∞ with finite mean such that $X_t \overset{a.s.}{\to} X_\infty$. We show that X_∞ must be equal to zero.

Suppose not. Recall, from Corollary 2, that when the cost of child rearing is only in terms of consumption goods

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

In other words

$$n(\theta)^{1-\eta}=\beta R\epsilon(\theta),$$

for some $\epsilon(\theta)$ with $E[\epsilon(\theta)]=1$. Therefore

$$(1-\eta)\log(n(\theta))=\log(\beta R)+\log(\epsilon(\theta)),$$

with $E[\log(\epsilon(\theta))]<\log(E[\epsilon(\theta)])=0$. Therefore,

$$(1-\eta)\log\left(N_{t+1}\left(\theta^{t}\right)\right)=(1-\eta)\sum_{s=0}^{t}\log(n(\theta_s))=(1+t)\log(\beta R)+\sum_{s=0}^{t}\log(\epsilon(\theta_s)).$$

This implies that $(1-\eta)\log(N_{t+1}(\theta^t))-(1+t)\log(\beta R)$ is a random walk with negative drift.

Now, if $X_\infty>0$, then $\log(X_\infty)>-\infty$, but

$$\log(X_t)=(1-\eta)\log\left(N_{t+1}\left(\theta^{t}\right)\right)-(1+t)\log(\beta R)-\log\left(u'\left(c_{t+1}\left(\theta^{t+1}\right)\right)\right)$$

cannot converge to a finite number. Contradiction.

Note that we have already established that $c_t(\theta_t)$ is i.i.d and bounded. Therefore, it must be true that $\frac{N_{t+1}\left(\theta^{t}\right)^{1-\eta}}{(\beta R)^{t+1}} \overset{a.s.}{\to} 0$. ■

Note that growth rate of total population is $E\left[n\left(\theta\right)\right]$. To see this, let $\bar{N}_{t+1}=\sum_{\theta^t}\pi\left(\theta^t\right)N_{t+1}\left(\theta^t\right)$ be total population. Then,

$$\begin{aligned}\bar{N}_{t+1} &= \sum_{\theta^t}\pi\left(\theta^t\right)N_t\left(\theta^{t-1}\right)n\left(\theta_t\right)=\sum_{\theta^{t-1}}\pi\left(\theta^{t-1}\right)N_t\left(\theta^{t-1}\right)\sum_{\theta}\pi\left(\theta\right)n\left(\theta\right)\\ &= \bar{N}_t\sum_{\theta}\pi\left(\theta\right)n\left(\theta\right)\end{aligned}$$

where the last equality follows because the $n\left(\theta\right)$ are i.i.d.

Note that if $\eta>0$, then $\left(E\left[n\left(\theta\right)\right]\right)^{1-\eta}>E\left[n\left(\theta\right)^{1-\eta}\right]=\beta R$. Therefore, $E\left[n\left(\theta\right)\right]>\left(\beta R\right)^{1/(1-\eta)}$, and population growth is higher than $\left(\beta R\right)^{1/(1-\eta)}$. In this case, almost all dynasties grow more slowly than total population.

If, however, $\eta<0$, then $\left(E\left[n\left(\theta\right)\right]\right)^{1-\eta}<E\left[n\left(\theta\right)^{1-\eta}\right]=\beta R$. In this case, $E\left[n\left(\theta\right)\right]<\left(\beta R\right)^{1/(1-\eta)}$, and we cannot say whether all dynasties grow faster or slower than total population.

7 Implementation

Here, we discuss the implementation of efficient allocations via decentralized decision making with taxes. To simplify the presentation we restrict attention to a two period example and explicitly characterize how tax implementations are used to alter private fertility choices.

We assume that there is a one time shock, realized in the first period.

The constrained efficient allocation $c_{1i}^*,l_i^*,n_i^*,c_{2i}^*$ solves the following problem:

$$\sum_{i=H,L}\pi_i\left[u(c_{1i})+h(1-l_i-bn_i)+\beta n_i^{\eta}u(c_{2i})\right]$$

s.t.

$$\sum_{i=H,L}\pi_i\left[c_{1i}+\frac{1}{R}n_ic_{2i}\right]\leq\sum_{i=H,L}\pi_i\theta_il_i+RK_0$$

$$u(c_{1H})+h(1-l_H-bn_H)+\beta n_H^{\eta}u(c_{2H})\geq u(c_{1L})+h(1-\frac{\theta_Ll_L}{\theta_H}-bn_L)+\beta n_L^{\eta}u(c_{2L}).$$

Now suppose that we want to implement the above allocation with a tax function of the form $T(y,n,c_2)$. Then the consumer's problem is the following:

$$\max_{c_1,n,l,c_2}u(c_1)+h(1-l-bn)+\beta n^{\eta}u(c_2)$$

s.t.

$$\begin{aligned} c_1 + k_1 &\leq Rk_0 + \theta l - T(\theta l, n, c_2) \\ nc_2 &\leq Rk_1 \end{aligned}$$

It can be shown that if T is differentiable and if y is interior for both types $T_n(\theta_H l_H^*, n_H^*, c_{2H}^*) = T_y(\theta_H l_H^*, n_H^*, c_{2H}^*) = T_{c_2}(\theta_H l_H^*, n_H^*, c_{2H}^*) = 0$ – there are no (marginal) distortions on the decisions of the agent with the high shock. Thus, what we need to do is to characterize the types of distortions that are used to get the low type to choose the correct allocation.

It is well known that when the type space is discrete, the constrained efficient allocation cannot be implemented by a continuously differentiable tax function. (This is also true in our environment.) However, there exists continuous and piecewise differentiable tax functions which implement the constrained efficient allocation. Next, we construct the analog of this for our environment.

Let $\bar{u}_L$ (resp. $\bar{u}_H$) be the level of utility received at the socially efficient allocation by the low (resp. high) type, and define two versions of the tax function:

$$\begin{aligned} \bar{u}_L &= u(y - T_L(y, n, c_2) - \frac{1}{R} nc_2) + h(1 - \frac{y}{\theta_L} - bn) + \beta n^\eta u(c_2), \\ \bar{u}_H &= u(y - T_H(y, n, c_2) - \frac{1}{R} nc_2) + h(1 - \frac{y}{\theta_H} - bn) + \beta n^\eta u(c_2). \end{aligned}$$

T_L, is designed to make sure that the low type always gets utility $\bar{u}_L$ if they satisfy their budget constraint with equality while T_H, is defined similarly. It can be shown that such T_L and T_H always exist, and from the Theorem of the Maximum, they are continuous functions of (y, n, c_2). Moreover, since $c_1 > 0$ (i.e., $y - T - \frac{1}{R} nc_2 > 0$) they are each differentiable.

We will build the overall tax code, $T(y, n, c_2)$, by using T_L as the effective tax code for the low type and T_H as the one for the high type. Given this, it follows that the distortions, at the margin, faced by the two types are described by the derivatives of T_L (T_H) with respect to y and n.

Remark 2 ***Remark 3*** *If the allocation is interior,*

1. *The tax function*

$$T(y, n, c_2) = \max\{T_L(y, n, c_2), T_H(y, n, c_2)\}$$

implements the efficient allocation.

2. *If the incentive constraint for the low type is slack, there are no distortions in the decisions of the high type –* $\frac{\partial T}{\partial y}(y_H^*, n_H^*, c_{2H}^*) = \frac{\partial T_H}{\partial y}(y_H^*, n_H^*, c_{2H}^*) = 0$ *and* $\frac{\partial T}{\partial n}(y_H^*, n_H^*, c_{2H}^*) =$

$\frac{\partial T_H}{\partial n}(y_H^*, n_H^*, c_{2H}^*) = 0.$

3. *At the choice of the low type, (y_L^*, n_L^*, c_{2L}^*), $T = T_L$ and (i) $\frac{\partial T_L}{\partial y}(y_L^*, n_L^*, c_{2L}^*) > 0$; (ii) $\frac{\partial T_L}{\partial n}(y_L^*, n_L^*, c_{2L}^*) > 0$; (iii) $\frac{\partial T_L}{\partial c_2}(y_L^*, n_L^*, c_{2L}^*) = 0$.*

Proof. First we show, using incentive compatibility, that $T_L(y_L^*, n_L^*, c_{2L}^*) = T_H(y_L^*, n_L^*, c_{2L}^*)$. We know that at the constrained efficient allocation, type θ_H is indifferent between the allocations $(c_{1H}^*, y_H^*, n_H^*, c_{2H}^*)$ and $(c_{1L}^*, y_L^*, n_L^*, c_{2L}^*)$. Hence we have the following equality:

$$\bar{u}_H = u(c_{1H}^*) + h(1 - \frac{y_H^*}{\theta_H} - bn_H^*) + \beta n_H^{*\eta} u(c_{2H}^*) = u(c_{1L}^*) + h(1 - \frac{y_L^*}{\theta_H} - bn_L^*) + \beta n_L^{*\eta} u(c_{2L}^*)$$

Replace for c_{1H}^* and c_{1L}^* from budget constraints to get

$$\begin{aligned} \bar{u}_H &= u(y_H^* - T_H(y_H^*, n_H^*, c_{2H}^*) - \frac{1}{R} n_H^* c_{2H}^*) + h(1 - \frac{y_H^*}{\theta_H} - bn_H^*) + \beta n_H^{*\eta} u(c_{2H}^*) \\ &= u(y_L^* - T_L(y_L^*, n_L^*, c_{2L}^*) - \frac{1}{R} n_L^* c_{2L}^*) + h(1 - \frac{y_L^*}{\theta_H} - bn_L^*) + \beta n_L^{*\eta} u(c_{2L}^*) \end{aligned}$$

Moreover, from the definition of T_H we know that

$$\bar{u}_H = u(y_L^* - T_H(y_L^*, n_L^*) - \frac{1}{R} n_L^* c_{2L}^*) + h(1 - \frac{y_L^*}{\theta_H} - bn_L^*) + \beta n_L^{*\eta} u(c_{2L}^*)$$

Hence, the last two equalities imply that $T_L(y_L^*, n_L^*, c_{2L}^*) = T_H(y_L^*, n_L^*, c_{2L}^*)$. We can also show that $T_H(y_H^*, n_H^*, c_{2H}^*) > T_L(y_H^*, n_H^*, c_{2H}^*)$. We show that this holds as long as the upward incentive constraint is slack – θ_L strictly prefers the allocation $(c_{1L}^*, y_L^*, n_L^*, c_{2L}^*)$ to $(c_{1H}^*, y_H^*, n_H^*, c_{2H}^*)$, i.e.:

$$\begin{aligned} \bar{u}_L &= u(c_{1L}^*) + h(1 - \frac{y_L^*}{\theta_L} - bn_L^*) + \beta n_L^{*\eta} u(c_{2L}^*) \\ &> u(c_{1H}^*) + h(1 - \frac{y_H^*}{\theta_L} - bn_H^*) + \beta n_H^{*\eta} u(c_{2H}^*). \end{aligned}$$

Using the budget constraints, we get

$$\begin{aligned} \bar{u}_L &= u(y_L^* - T_L(y_L^*, n_L^*, c_{2L}^*) - \frac{1}{R} n_L^* c_{2L}^*) + h(1 - \frac{y_L^*}{\theta_L} - bn_L^*) + \beta n_L^{*\eta} u(c_{2L}^*) \\ &> u(y_H^* - T_H(y_H^*, n_H^*, c_{2H}^*) - \frac{1}{R} n_H^* c_{2H}^*) + h(1 - \frac{y_H^*}{\theta_L} - bn_H^*) + \beta n_H^{*\eta} u(c_{2H}^*) \end{aligned}$$

By the definition of $T_L$we have

$$\bar{u}_L = u(y_H^* - T_L(y_H^*, n_H^*, c_{2H}^*) - \frac{1}{R} n_H^* c_{2H}^*) + h(1 - \frac{y_H^*}{\theta_L} - bn_H^*) + \beta n_H^{*\eta} u(c_{2H}^*)$$

Hence, we have that $T_H(y_H^*, n_H^*, c_{2H}^*) > T_L(y_H^*, n_H^*, c_{2H}^*)$. Given the tax function, the consumer's problem is the following:

$$\begin{aligned} \max_{c_1,y,n,c_2} \quad & u(c_1) + h(1 - \frac{y}{\theta} - bn) + \beta n^{\eta} u(c_2) \\ & \text{s.t. } c_1 + \frac{1}{R} n c_2 \le y - T(y, n) \end{aligned}$$

From above, we know that $T(y_H^*, n_H^*, c_{2H}^*) = T_H(y_H^*, n_H^*, c_{2H}^*)$. Hence, type θ_H can afford $(c_{1H}^*, y_H^*, n_H^*, c_{2H}^*)$ and $u(c_{1H}^*, y_H^*, n_H^*, c_{2H}^*; \theta_H) = \bar{u}_H$. Let (c_1, y, n, c_2) be any allocation that satisfies $c_1 + \frac{1}{R} n c_2 = y - T(y, n, c_2)$. Then,

$$c_1 + \frac{1}{R} n c_2 = y - \max\{T_L(y, n, c_2), T_H(y, n, c_2)\} \le y - T_H(y, n, c_2)$$

But by definition of T_H, $u(c_1, y, n, c_2; \theta_H)$ can be at most $\bar{u}_H$. Using a similar argument we can show that type θ_L can afford $(c_{1L}^*, y_L^*, n_L^*, c_{2L}^*)$ and $u(c_{1L}^*, y_L^*, n_L^*, c_{2L}^*; \theta_L) = \bar{u}_L$. Moreover, any allocation that satisfies the budget constraint has utility at most $\bar{u}_L$. ■

The new finding here is that the planner chooses to tax the low type at the margin for having more children – $\frac{\partial T_L}{\partial n}(y_L^*, n_L^*, c_{2L}^*) > 0$. In the Mirrlees model without fertility choice, for incentive reasons, the planner wants to make sure that the low type consumes more leisure (relative to consumption) than he would in a full information world – this makes it easier to get the high type to truthfully admit his type. This is accomplished by having a positive marginal labor tax rate for the low type. Here, there is an additional incentive effect that must be taken care of. This is for the planner to make sure that the low type doesn't use too much of his time free from work raising children. This would also make it more appealing to the high type to lie. To offset this here, the planner also charges a positive tax rate on children for the low type. These two effects taken together ensure that the low type has low consumption and fertility and high leisure thereby separating from the high type.

This feature – a positive tax rate on fertility – is somewhat special to the case where there is no future labor supply after the first period. Indeed, one can show that if the costs of children are in terms of goods only, the tax rate on fertility is negative in the infinite horizon case for incentive reasons.

References

Alvarez, F. (1999): "Social Mobility: The Barro–Becker Children Meet the Laitner–Loury Dynasties," *Review of Economic Dynamics*, 2, 65–103.

Farhi, E. and I. Werning (2010): "Insurance and Taxation over the Life Cycle," Working Paper, MIT.

Fernandes, A. and C. Phelan (2000): "A Recursive Formulation for Repeated Agency with History Dependence," *Journal of Economic Theory*, 91, 223–247.

Golosov, M., M. Troshkin, and A. Tsyvinski (2010): "Optimal Dynamic Taxes," Working Paper, Yale University.

Pavan, A., I. Segal, and J. Toikka (2009): "Dynamic Mechanism Design: Incentive Compatibility, Profit Maximization and Information Disclosure," Working Paper, Northwestern University.

Williams, N. (2009): "Persistent Private Information," Working Paper, University of Wisconsin.

Zhang, Y. (2009): "Dynamic Contracting with Persistent Shocks," *Journal of Economic Theory*, 144, 635–675.