Risk Sharing, Inequality, and Fertility *

Roozbeh Hosseini
Arizona State University
rhossein@asu.edu

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

Ali Shourideh
University of Minnesota
shour004@umn.edu

June 26, 2010

Abstract

We use an extended Barro-Becker model of endogenous fertility, in which parents are heterogeneous in their labor productivity, to study the efficient degree of consumption inequality in the long run when parents productivity is private information. We show that a feature of the informationally constrained optimal insurance contract is that there is a stationary distribution over per capita continuation utilities – there is an efficient amount of long run inequality. This contrasts with much of the earlier literature on dynamic contracting where 'immiseration' occurs. Further, the model has interesting and novel implications for the policies that can be used to implement the efficient allocation. Two examples of this are: 1) estate taxes are positive and 2) there are positive taxes on family size.

Keywords: Private Information, Risk Sharing, Long run Inequality, Endogenous Fertility.

*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, Colubmia, Carnegie Mellon, Iowa State, Ohio State, Texas Austin, Western Ontario, Wharton, Yale, 2009 SED summer meeting, 2009 Minnesota Macro Workshop and 2010 Cowles Summer Conference for comments.

1 Introduction

A key question in normative public finance is the extent to which it is socially efficient to insure agents against shocks to their circumstances. The basic trade-off is one between providing incentives for productive agents to work hard – thereby making the pie big – versus transferring more resources to less productive agents to insure them. This is the problem first analyzed in Mirrlees (1971) where he characterized the solution to a problem of this form in a static setting. This analysis is at the heart of a deep and important question in economics – What is the optimal amount of inequality in society? Mirrlees provides one answer to this question along with a way of implementing his answer to this question – a tax and transfer scheme based on non-linear income taxes that makes up the basis of the optimal social safety net.

More recently, a series of authors (e.g., Green (1987), Thomas and Worrall (1990) and Atkeson and Lucas (1992)) have extended this analysis to cover dynamic settings – agents are more productive some times and less others. A common result from this literature is that the socially efficient level of insurance (ex ante and under commitment) involves an asymmetry between how good and bad shocks are treated. In particular, when an agent is hit with a bad shock, the decrease in what he can expect in the future is more than the corresponding increase after a good shock – there is a negative drift in expected future utility. This feature of the optimal contract has become known as 'immiseration.' Immiseration, although interesting on its own, is more important as an indicator of a more severe problem in the models; there is not a stationary distribution over continuation utilities – the optimal amount of inequality in society grows without bound over time.[1] This weakness means the models cannot be used to answer questions such as: Is there too much inequality in society under the current system?

Two recent papers, Phelan (2006) and Farhi and Werning (2007) have given a different view of the immiseration result. This is to interpret different periods in the model as different generations. In this case, current period agents care about the future because of dynastic altruism a la Barro (1974). Under this interpretation, social insurance is comprised of two conceptually different components: 1) Insurance

[1]Immiseration and the lack of a meaningful stationary distribution are not equivalent in general. There are other examples in the literature (e.g., see Khan and Ravikumar (2001) and Williams (2009)) that show that immiseration need not hold. However, in those examples there is no stationary distribution over continuation utilities – variance grows without bound.

against the uncertainty coming from current generation productivity shocks, 2) Insurance against the uncertainty coming from the shock of what family you are born into – what future utility was promised to your parents (e.g., through the intergenerational transmission of wealth). Under this interpretation, immiseration means that optimal insurance for the current generation features a lower utility level for the next generation.

These authors go on to show that optimal contract features a stationary level of inequality if society values the welfare of children strictly more than their parents do. When this is true, the amount of this long run inequality depends on the difference between societal and parental altruism.

In an intergenerational setting with dynastic altruism such as that studied by these authors, a natural question arises: To what extent are these results altered when the size of generations – i.e., fertility – is itself endogenous (e.g., as in Barro and Becker (1989) and Becker and Barro (1988))? This question is the focus of this paper.

We show that the explicit inclusion of fertility choice in the model alters the qualitative character of the optimal allocations in two important ways. First, we show that even when the planner does *not* put extra weight on future generations, there is a stationary distribution in per capita variables – there is an optimal amount of long run per capita inequality and no immiseration in per capita terms. In addition to this, since fertility is explicitly included, the model has implications about the properties of fertility. Because of this, the model also has implications about the best way to design policies relating to family size (e.g., child care deductions, tax credits for children, education subsidies, etc.)

From a mechanical point of view, the inclusion of fertility gives the planner an extra instrument to use to induce current agents to truthfully reveal their productivities. That is, the planner can use both family size and continuation utility of future generations to induce truthful revelation. In the normal case (i.e., without fertility choice), in order to induce truth telling today, the planner (optimally) chooses to 'spread' out continuation utilities so as to be able to offer insurance in current consumption. The incentives for the planner to do this are present after every possible history. Because it is cheaper to provide incentives in the future when continuation utilities are lower this outward pressure is asymmetric and has a negative bias. Thus, continuation utilities are pushed to their lower bound – inequality becomes greater and greater over time and immiseration occurs.

In contrast, when fertility is endogenous, this optimal degree of spreading in continuation utility for the parent can be thought of as being provided through two distinct sources – spreading of per child continuation utility and spreading in family sizes. In general both of these instruments are used to provide incentives, but the way that they are used is different. While dynasty size can grow or shrink without bound for different realizations, we find that there is a natural limit to the amount of spreading that occurs through per child continuation utilities. Specifically, per child continuation utilities lie in a bounded set. Thus, even if the promised utility to a parent is very low, the continuation utility for their children is bounded below. Similarly, even if the promised utility to a parent is very high, the continuation utility for their children is bounded above. These results form the basis for showing that a stationary distribution exists.

This property, boundedness of per child continuation utilities, is shown by exploiting an interesting kind of history independence in the model. This is what we call the 'resetting' property. There are two versions of this that are important for our results. The first concerns the behavior of continuation utility for children when promised utility for a parent is very low. After some point, reducing promised utility to the parent no longer reduces per child continuation utility – continuation utility for children is bounded below. An implication of this is that if promised utility to the parent is low enough, continuation utility for his children will automatically be higher, no matter what productivity shock is realized. The second version says something similar for high promised utilities for parents. A particularly striking version of this concerns the behavior of the model whenever a family experiences the highest value of the shock: There is a value of continuation utility, w_0, such that all children are assigned to this level no matter what promised utility is. That is, continuation utility gets 'reset' to w_0 subsequent to every realization of a high shock. So, even if a family has a very long series of good shocks, the utility of the children does not continue to grow but stays fixed at w_0.

This reasoning concerning the limits of long run inequality in per capita variables differs from what is found in Farhi and Werning (2007) in two ways. The first of these is the basic reason for the breakdown in the immiseration result. In Farhi and Werning (2007), it is because of a difference between social and private discounting – society puts more weight on future generations than parents do. Here, immiseration breaks down even when social and private discounting are the same because of resetting at the top and bottom of the promised utility distribution. The second difference concerns the movements of per capita consumption over time. The version of the Inverse Euler Equation that holds in the Farhi and Werning (2007) world when society is more patient than individuals implies that consumption has a mean reversion property. In our model, there is a lower bound on continuation utilities for children which is independent of both the promised utility to parents and the current productivity shock. Moreover, as discussed above, the resetting property at the top implies that continuation utility reverts to w_0 each time a high shock is realized. Thus, the types of intergenerational mobility that are present in the two models are quite different.

In contrast to these results about the limits to spreading through per child continuation utility, we find that there is no upper bound on the amount of spreading that occurs through the choice of family size. If the discount factor is equal to the inverse of the interest rate, we show that along *almost* any subset of the family tree, population dies out.[2] However, this does not necessarily imply that population shrinks, since this property holds even when mean population is growing. Rather, some strands of the dynasty tree die out and others expand. Those that are growing are exactly those sub-populations that have had the best 'luck.'

From a mechanical point of view, the key technical difference between our results and that from earlier work is that here, bounds on continuation utility arise naturally from the form of the contracting problem rather than from being exogenously imposed. Contracting problems in which social and private discount factors are different (such as those studied in Phelan (2006) and Farhi and Werning (2007)) can equivalently be thought of as problems where the social and private discount factors are the same, but there are lower bounds on the continuation utility levels of future generations. As such, they are closely linked to the approach followed in Atkeson and Lucas (1995) and Sleet and Yeltekin (2006). Here, the endogenous bounds arise because of the inclusion and optimal exploitation by the planner of a new choice variable, family size.

There are several other interesting differences between the two approaches. For example, one of the key ways that Phelan (2006) and Farhi and Werning (2007), differ from earlier work is that in the socially efficient scheme, the Inverse Euler

[2] This does not hold if the discount factor is not the inverse of the interest rate, see the discussion in Section 4.

Equation need not hold. Indeed, in those papers, the inter-temporal wedge can even be negative.[3] The negative intertemporal wedge can be interpreted (in some implementations) as requiring a negative estate tax. Hence, these authors argue that, in order to overcome immiseration, a negative inter-temporal wedge may be necessary. However, this result does not hold in our model, since a version of the Inverse Euler Equation holds. This implies that there will always be a positive 'wedge' in the FOC determining savings and hence, estate taxes are always implicitly positive.

Finally, an interesting new feature that emerges is the dependence of taxes on family size. What we find is that for everyone other than the highest type, there is a positive wedge on the fertility-consumption margin – fertility is discouraged to better provide incentives for truthful revelation.

Our paper is related to the large literature on dynamic contracting including Green (1987), Thomas and Worrall (1990), Atkeson and Lucas (1992) (and 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.

We also contribute to a line of research that investigates the relationship between income and fertility (see Jones et al. (2008) for a recent summary). In fact, our paper has some novel implications about fertility per se. First, the socially efficient allocation is characterized by a negative income-fertility relationship—independently on specific assumptions on curvature in utility. This suggests that intergenerational income risk and intra-generational risk sharing may be important factors to explain the observed negative income-fertility relationship. Second, very few papers have analyzed ability heterogeneity and intergenerational transmission of wealth in dynastic models with fertility choice. Our paper is most related to Alvarez (1999) who analyzes intergenerational income risk but assumes that it is uninsurable.

In section 2 we present a two period, two shock version of the model to show

[3] Farhi and Werning (2010) in the same environment show that estate taxes have to be progressive.

the basic results in a simple setting. Section 3 contains the description of the general model with private information. In section 4 we study the properties of the model relating to long run inequality. In section 5, we discuss some extensions and complimentary results.

2 An Example and Intuition

In this section we illustrate the key idea for our results in two steps. First, we study our basic incentive problem in a two period model and derive a property, which we call 'resetting.' This property shows that there is a (high) level of utility that is assigned to all children of workers that have high productivity independent of the level of promised utility to the worker. This provides a strong intuition for why adding fertility to the model has such a large impact on the asymptotic behavior of continuation utility – there is a continual recycling of utility levels up to this 'resetting value' each time a high value of the shock is realized. This by itself does not imply that there is a stationary distribution over continuation utilities, but it is one important step in the argument.

Second, we show that the reason that this occurs in the Barro and Becker (1989) model of dynastic altruism is because of a homotheticity property of utility in this model. [4]

In sum, the key feature that fertility brings to the contracting environment is a distinction between total and per capita continuation utility. The implications of this difference are particularly sharp in the Barro and Becker utility case, but they are not limited to it.

2.1 A Two Period Example

Immiseration concerns the limiting behavior of continuation utility as a history of shocks is realized. In a stationary environment, this is determined by the properties of the policy function describing the relationship between future utility, W', as a function of current promised utility, W_0 (and the current shock). Immiseration is

[4] In the Supplementary Appendix we show that even for more general utility functions, a similar result will hold with fertility in the model as long as a certain combination of elasticities is bounded away from infinity. See Hosseini et al. (2010).

then the statement that the only stationary point of this mapping is for continuation utility to converge to its lower bound.

To gain some intuition, we will mimic this in a two period example by reinterpreting W' as second period utility and W_0 as ex ante utility. Then, a necessary condition for immiseration is: when W_0 is low, W' is even lower. This is the version that we will explore in this section.

To this end, consider a two-period economy populated by a continuum of parents with mass 1 who live for one period. Each parent receives a random productivity θ in the set $\Theta = \{\theta_L, \theta_H\}$ with $\theta_H > \theta_L$. At date 1, each parent's productivity, θ, is realized. After this, they consume, work and decide about the number of children. The cost of having a child is in terms of leisure. Every child requires b units of leisure to raise. The coefficient b can be thought of as time spent raising children (or the market value of maternity leave for women). The child lives for one period and consumes out of the savings done by their parents. The parent's utility function is:

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

in which l is hours worked, n is the number of children and c_t is consumption per person in period t. From this, it can be seen that the parent has an altruistic utility function where the degree of altruism is determined by β.

We assume that a worker of productivity $\theta \in \Theta$ who works for l hours has effective labor supply of θl and that both θ and l are private information of the parent. As is typically assumed, we assume that the planner can observe θl. In what follows we denote the aggregate consumption of all children by $C_2 = n_2 c_2$.

Suppose each parent is promised an ex ante utility W_0 at date zero and that the planner has access to a saving technology at rate R. Thus, the planner wishes to allocate resources efficiently subject to the constraints that - 1) Ex ante utility to each parent is at least W_0, 2) Each 'type' is willing to reveal itself - IC.[5]

$$\min_{c_1(\theta), n(\theta), l(\theta), c_2(\theta)} \sum_{\theta_L, \theta_H} \pi(\theta) \left(c_1(\theta) + \frac{1}{R} n(\theta) c_2(\theta) \right) - \sum_{\theta_L, \theta_H} \pi(\theta) \theta l(\theta) \tag{1}$$

[5]Both here, and throughout the remainder of the paper, we will assume that only downward incentive constraints bind. Under certain conditions it can be shown focusing on the downward incentive constraints is sufficient. (See Hosseini et al. (2010).)

s.t.

$$\sum_{\theta_L,\theta_H} \left[\pi(\theta)\left(u(c_1(\theta)) + h\left(1 - l(\theta) - bn(\theta)\right) + \beta n(\theta)^\eta u(c_2(\theta))\right)\right] \geq W_0 \tag{2}$$

$$\begin{aligned} u(c_1(\theta_H)) \quad + \quad & h\left(1 - l(\theta_H) - bn(\theta_H)\right) + \beta n(\theta_H)^\eta u(c_2(\theta_H)) \geq u(c_1(\theta_L)) + \\ & h\left(1 - \frac{\theta_L}{\theta_H} l(\theta_L) - bn(\theta_L)\right) + \beta n(\theta_L)^\eta u(c_2(\theta_L)) \end{aligned} \tag{3}$$

After manipulating the first order conditions for the allocation of the highest type, we obtain:

$$\eta u(c_2(W_0, \theta_H)) = u'(c_2(W_0, \theta_H))c_2(W_0, \theta_H) + b\theta_H R u'(c_2(W_0, \theta_H)). \tag{4}$$

This is an equation in per child consumption for children of parents with the highest shock. Notice that none of the other endogenous variables of the system appear in this equation. Similarly, W_0 does not appear in the equation. Because of this, it follows that the level of consumption for these children, $c_2(W_0, \theta_H)$, is independent of W_0.

We will call this the 'resetting' property – i.e., per capita consumption for the children of parents with the highest shock is reset to a level that is independent of state variables.

There are two key features in the model that are important in deriving this result. The first of these comes from our assumption that no one pretends to be the highest type θ_H.[6] Because of this, it follows from the usual argument that the allocations of this type are undistorted.

The second important feature comes from the fact that the problem is 'homogeneous/homothetic' in aggregate second period consumption, $C_2 = nc_2$, and family size, n. Because of this, $\frac{C_2}{n} = c_2$ is independent of W_0 in undistorted allocations.[7]

In sum, $\lim_{W_0 \to -\infty} C_2(W_0, \theta_H) = 0$, $\lim_{W_0 \to -\infty} n(W_0, \theta_H) = 0$, but $\frac{C_2(W_0,\theta_H)}{n(W_0,\theta_H)}$ is constant.

[6] We can show that at the full information efficient allocations the downward constraints are binding and upward constraints are slack. We also verify the slackness of upward constraints in out numerical example (in infinite horizon environment). In general we cannot prove that only downward constraints are binding because the preferences do not exhibit single-crossing property.

[7] As it turns out, in the full information analog of this problem, this line of reasoning holds for all types, not just the highest one. I.e., with full information, the consumption of a child of a parent of type θ is given by $c_2(W_0, \theta)$ which is independent of W_0.

The next step is to use this result to say something about immiseration. There are two ways to interpret continuation utility in our setting: continuation utility from the point of view of the parent, $\beta n^{\eta} u(c_2)$, and, continuation utility from the point of each child, $u(c_2)$. When fertility is exogenous these two alternative notions are equivalent.

From our discussion above, the 'resetting' property implies that $u(c_2(W_0, \theta_H)) = u(c_2(\theta_H))$ is bounded away from the lower bound of utility and hence there is no immiseration in this sense. However, since $u(c_2(W_0, \theta_H))$ is bounded below and $n(W_0, \theta_H) \rightarrow 0$, it follows that $\beta n(W_0, \theta_H)^{\eta} u(c_2(W_0, \theta_H))$ converges to its lower bound.[8] We summarize this discussion as a Proposition:

Proposition 1 *1.* $c_2(W_0, \theta_H) = \frac{C_2(W_0, \theta_H)}{n(W_0, \theta_H)}$ *is independent of* W_0*;*

2. $u(c_2(W_0, \theta_H))$ *is bounded below;*

3. $\beta n(W_0, \theta_H)^{\eta} u(c_2(W_0, \theta_H))$ *converges to its lower bound.*

It turns out that similar results also hold in a larger class of environments including settings with a goods cost for children, and/or with taste shock rather than productivity shock (see Hosseini et al. (2010)).

Thus, there is a sense in which there is no immiseration – from the point of view of the children – and a sense in which there is immiseration – from the point of view of the parents. As can be seen from this discussion, the key feature, when fertility is included as a choice variable, is the difference between aggregate and per capita variables. While it is hard to think about infinite horizon and stationarity in a two period example, (2) and (3) provide partial intuition. For example, (2) implies that, per capita utility of children does not have a downward trend (as a function of W_0) – there is no immiseration from the point of view of the children. Interpreting (3) is even more difficult, but it, along with a statement that $\beta n(W_0, \theta_H)^{\eta} u(c_2(W_0, \theta_H))$ is below W_0 when W_0 is low enough implies that a form of immiseration *does* hold from the point of view of the parents.

This discussion is complicated by two additional factors when considering the infinite horizon model. The first of these is that to show that a stationary distribution

[8] There are two cases of relevance here. The first is $u \geq 0$ and $0 \leq \eta \leq 1$. In this case, $\beta n(W_0, \theta_H)^{\eta} u(c_2(W_0, \theta_H)) \rightarrow 0$. The second case is when $u \leq 0$ and $\eta < 0$. In this case $\beta n(W_0, \theta_H)^{\eta} u(c_2(W_0, \theta_H)) \rightarrow -\infty$.

exists (in per capita variables), it is not enough to show that there is no immiseration for the highest type. It must also be shown that utility is bounded below for other shocks too. This is a key step of the main result in the paper discussed below.

Second, showing that utility is bounded below for the highest type is also not sufficient for another reason. This is that the proportions of the population that are children of the highest type is itself endogenous. I.e., the result in the Proposition would not have much bite if, for example, $n(W_0, \theta_H) = 0$ for all W_0. This is also discussed below.

2.2 Resetting – Intuition via Homotheticity

Some intuition for the 'resetting' property can be obtained by reformulating the planer's problem.

As a first step, rewrite the problem above by letting $m = 1 - l - bn$ be parents leisure:

$$\min_{c_1(\theta),n(\theta),m(\theta),C_2(\theta)} \sum_{\theta\in\Theta} \pi(\theta)\left(c_1(\theta) + \theta m(\theta)\right) + \sum_{\theta\in\Theta} \pi(\theta)\left(b\theta n(\theta) + \frac{1}{R}C_2(\theta)\right) \tag{5}$$

s.t.

$$\sum_{\theta\in\Theta} \pi(\theta)\left(u(c_1(\theta)) + h\left(m(\theta)\right)\right) + \sum_{\theta\in\Theta} \pi(\theta)\beta n(\theta)^{\eta} u\left(\frac{C_2(\theta)}{n(\theta)}\right) \geq W_0 \tag{6}$$

$$\begin{aligned} u(c_1(\theta_H)) \quad + \quad & h\left(m(\theta_H)\right) + \beta n(\theta_H)^{\eta} u\left(\frac{C_2(\theta_H)}{n(\theta_H)}\right) \geq u(c_1(\theta_L)) + \\ & h\left(\frac{\theta_L}{\theta_H} m(\theta_L) + (1 - \frac{\theta_L}{\theta_H})\left(1 - bn(\theta_L)\right)\right) + \beta n(\theta_L)^{\eta} u\left(\frac{C_2(\theta_L)}{n(\theta_L)}\right) \end{aligned} \tag{7}$$

The first term in the objective function is the planner's expenditure on parents' consumption and leisure (denominated in parents' consumption). The second term is the total expenditure on children: their total consumption and time spent parenting (again, denominated in parents' consumption).

For an allocation to be the solution to this problem, there should not be a way to adjust $n(\theta_H)$ and $C_2(\theta_H)$, while holding the other variables fixed, which lowers cost while still satisfying the constraints.

As can be seen from this problem, there are no interactions between the variables

$n(\theta_H)$ and $C_2(\theta_H)$ and the other variables. That is, they enter additively in the objective function and together, but separate from all other variables in the constraints (i.e., only through the terms based on $\beta n(\theta_H)^\eta u\left(\frac{C_2(\theta_H)}{n(\theta_H)}\right)$). Because of this, it follows that, given the other variables in the problem, the optimal choice of $(n(\theta_H), C_2(\theta_H))$ must solve the sub-problem:

$$\begin{aligned} \min_{C_2,n} \quad & b\theta_H n + \frac{1}{R}C_2 \\ \text{s.t.} \quad & n^\eta u\left(\frac{C_2}{n}\right) = W(W_0, \theta_H) \end{aligned} \tag{8}$$

The resetting property for high productivity parents can be understood by studying this problem. As can be seen, the objective function in this problem is homogeneous of degree one, while the constraint set is homogeneous of degree η. This is analogous to an expenditure minimization problem with a homothetic utility function. One property that problems of this form have is linear income expansion paths. In this case, this means that the ratio $\frac{C_2(W(W_0,\theta_H),\theta_H)}{n(W(W_0,\theta_H),\theta_H)}$ does not depend on $W(W_0, \theta_H)$ (and therefore, does not depend on W_0). Instead, it only depends on technology and preference parameters. Drawing an analogy with consumer demand theory, relative demand, $\frac{C_2}{n}$, only depends on relative prices, $bR\theta$, and not on promised utility. Therefore, the resetting property that we find is an immediate consequence of the homotheticity of the utility function in the Barro and Becker formulation of dynastic altruism.

This same argument does not hold for the low type $\theta = \theta_L$ because $n(\theta_L)$ and $C_2(\theta_L)$ do not separate from the other variables in the maximization problem. Mathematically, this is because of the fact that $n(\theta_L)$ also enters the leisure term of a high type who pretends to be a low type in the incentive constraint. This effect is absent in the full information version of this problem and hence, in that case, there is 'resetting' for all types – with full info, there are values of per capita consumption, $c_2(\theta)$ such that $c_2(W_0, \theta) = c_2(\theta)$ for all W_0.

As the above discussion shows, the resetting property relies on the way parents discount the utility of children. Following Barro and Becker (1989) and Becker and Barro (1988) we have made two assumptions. One, the way parents care about utility of children is multiplicatively separable in the number of children and the consumption of each child. Second, the component of the utility that depends on the number of children is homothetic (n^η). The discussion above indicates that this homotheticity is important in getting the resetting property. If this function is not homothetic, the per capita consumption of each child whose parent receives a high shock may depend on the promised utility of the parent (W_0). However, as it turns out, it can be shown that under a very general condition per capita consumption, and hence, continuation utility, of each child remains bounded away from zero. In the Supplementary Appendix (Hosseini et al. (2010)), we give a general set of conditions under which there is no 'immiseration'.

3 The Infinite Horizon Model

In this section, we will extend the model in section 2 to an infinite horizon setting. The set of possible types is given by $\Theta = \{\theta_1 < \cdots < \theta_I\}$. As above, the planner can observe the output for each agent but not hours worked nor productivity. Using the revelation principle, we will only focus on direct mechanisms in which each agent is asked to reveal his true type in each period. As is typical in problems like these, it can be shown that the full information optimal allocation does not satisfy incentive compatibility. Although the argument is more complex than in the usual case, we show (see Hosseini et al. (2010)) that under the full information allocation, a higher productivity type would want to pretend to be a lower productivity type.

In addition to this, in Mirrleesian environments with private information where a single crossing property holds, one can show only downward incentive constraints bind. We do not currently have a proof that the only incentive constraints that ever bind are the downward ones. In keeping with what others have done (e.g., Phelan (1998) and Golosov and Tsyvinski (2007)), we assume that agents can only report a level of productivity that is less than or equal to their true type.[9] In the Supplementary Appendix, we give a sufficient condition for this to be true. See Hosseini et al. (2010).

Under this assumption, we can restrict reporting strategies, σ, to satisfy $\sigma_t(\theta^t) \leq \theta_t$. (Here, for every history θ^t, $\sigma_t(\theta^t)$ is agent's report of its productivity in period t and $\sigma^t(\theta^t)$ is the history of the reports.) Moreover, because of our assumed restriction on reports, we have $\sigma_t(\theta^t) \leq \theta_t$. Call the set of restricted reports Σ. Then, an allocation

[9] In numerically calculated examples, this assumption is redundant.

is said to be *incentive compatible* if

$$\sum_{t,\theta^t} \beta^t \pi(\theta^t) N_t(\theta^{t-1})^\eta \left[u\left(\frac{C_t(\theta^t)}{N_t(\theta^{t-1})} \right) + h\left(1 - \frac{L_t(\theta^t)}{N_t(\theta^{t-1})} - b\frac{N_{t+1}(\theta^t)}{N_t(\theta^{t-1})} \right) \right] \geq \tag{9}$$

$$\sum_{t,\theta^t} \beta^t \pi(\theta^t) N_t(\sigma^{t-1}(\theta^{t-1}))^\eta \left[u\left(\frac{C_t(\sigma^t(\theta^t))}{N_t(\sigma_{t-1}(\theta^{t-1}))} \right) + \right.$$

$$\left. h\left(1 - \frac{\sigma_t(\theta^t) L_t(\sigma^t(\theta^t))}{\theta_t N_t(\sigma^{t-1}(\theta^{t-1}))} - b\frac{N_{t+1}(\sigma^t(\theta^t))}{N_t(\sigma^{t-1}(\theta^{t-1}))} \right) \right] \quad \forall \ \sigma \in \Sigma$$

Hence, the planning problem becomes the following:

$$\sum_{t,\theta^t} \beta^t \pi(\theta^t) N_t(\theta^{t-1})^\eta \left[u\left(\frac{C_t(\theta^t)}{N_t(\theta^{t-1})} \right) + h\left(1 - \frac{L_t(\theta^t)}{N_t(\theta^{t-1})} - b\frac{N_{t+1}(\theta^t)}{N_t(\theta^{t-1})} \right) \right]$$

s.t.

$$\sum_{t,\theta^t}^{\infty} \frac{1}{R^t} \pi(\theta^t) \left[C_t(\theta^t) - \theta_t L_t(\theta^t) \right] \leq K_0$$

$$\sum_{t,\theta^t} \beta^t \pi(\theta^t) N_t(\theta^{t-1})^\eta \left[u\left(\frac{C_t(\theta^t)}{N_t(\theta^{t-1})} \right) + h\left(1 - \frac{L_t(\theta^t)}{N_t(\theta^{t-1})} - b\frac{N_{t+1}(\theta^t)}{N_t(\theta^{t-1})} \right) \right] \geq$$

$$\sum_{t,\theta^t} \beta^t \pi(\theta^t) N_t(\sigma^{t-1}(\theta^{t-1}))^\eta \left[u\left(\frac{C_t(\sigma^t(\theta^t))}{N_t(\sigma_{t-1}(\theta^{t-1}))} \right) + \right.$$

$$\left. h\left(1 - \frac{\sigma_t(\theta^t) L_t(\sigma^t(\theta^t))}{\theta_t N_t(\sigma^{t-1}(\theta^{t-1}))} - b\frac{N_{t+1}(\sigma^t(\theta^t))}{N_t(\sigma^{t-1}(\theta^{t-1}))} \right) \right] \quad \forall \ \sigma \in \Sigma$$

Using standard arguments, we can show that the above problem is equivalent to the following functional equation:

$$V(N, W) = \min_{C(\theta), L(\theta), N'(\theta)} \sum_{\theta \in \Theta} \pi(\theta) \left[C(\theta) - \theta L(\theta) + \frac{1}{R} V(N'(\theta), W'(\theta)) \right] \qquad (P1) \tag{10}$$

s.t.

$$\sum_{\theta \in \Theta} \pi(\theta) \left[N^\eta \left(u\left(\frac{C(\theta)}{N} \right) + h\left(1 - \frac{L(\theta)}{N} - b\frac{N'(\theta)}{N} \right) \right) + \beta W'(\theta) \right] \geq W$$

$$N^\eta \left(u\left(\frac{C(\theta)}{N}\right) + h\left(1 - \frac{L(\theta)}{N} - b\frac{N'(\theta)}{N}\right)\right) + \beta W'(\theta) \geq \tag{11}$$
$$N^\eta \left(u\left(\frac{C(\hat{\theta})}{N}\right) + h\left(1 - \frac{\hat{\theta}L(\hat{\theta})}{\theta N} - b\frac{N'(\hat{\theta})}{N}\right)\right) + \beta W'(\hat{\theta}) \qquad \forall \theta > \hat{\theta}.$$

As we can see, the problem is homogeneous in N and therefore as before, if we define $v(N, w) = \frac{V(N, N^\eta w)}{N}$, $v(\cdot, \cdot)$ will not depend on N and satisfies the following functional equation:

$$v(w) = \min_{c(\theta), l(\theta), n(\theta), w'(\theta)} \sum_{\theta \in \Theta} \pi(\theta) \left(c(\theta) - \theta l(\theta) + \frac{1}{R} n(\theta) v(w'(\theta)) \right) \tag{P1'}$$

s.t.

$$\sum_{\theta \in \Theta} \pi(\theta) \left(u(c(\theta)) + h(1 - l(\theta) - bn(\theta)) + \beta n(\theta)^\eta w'(\theta) \right) \geq w$$

$$u(c(\theta)) + h(1 - l(\theta) - bn(\theta)) \quad + \quad \beta n(\theta)^\eta w'(\theta) \geq u(c(\hat{\theta})) + h(1 - \frac{\hat{\theta} l(\hat{\theta})}{\theta} - bn(\hat{\theta})) + \beta n(\hat{\theta})^\eta w'(\hat{\theta}), \ \forall \ \theta > \hat{\theta}$$

In what follows, we will assume that the solution to the minimization problem, (P1), has several convenient mathematical properties. These include strict convexity and differentiability of the value function as well as the uniqueness of the policy functions. Normally, these properties can be derived from primitives by showing that $V(N, W)$ is strictly convex, that the constraint set is convex, etc. Because of the presence of the incentive compatibility constraints, the usual lines of argument will not work (due to the non-convexity of the constraint set). In some contracting problems, these issues can be partially resolved. For example, in some cases, a change of variables can be designed so that convexity of the constraint set is guaranteed. Here, because of the way that fertility and labor supply enter the problem, this will no longer work. An alternative way to resolve this issue is by allowing for randomization. Allowing for randomization, makes all the constraints linear in the probability distributions and therefore the constraint correspondence is convex. This is the method used in Phelan and Townsend (1991) and Doepke and Townsend (2006) (see also Acemoglu et al. (2008)). This is not quite enough for us since it only implies convexity of V, not strict convexity, and hence, uniqueness of the policy function cannot be guaranteed.[10] Because of this, we simply assume that V has the needed convexity properties. Similar considerations hold for the differentiability of V. The following lemma on V provides useful later in the paper:

Lemma 2 *If $V(N, W)$ is continuously differentiable and strictly convex, then $v(w)$ is continuously differentiable and strictly convex. Moreover, $\eta v'(w)w - v(w)$ is strictly increasing.*

See Supplementary Appendix in Hosseini et al. (2010) for the proof.

In addition, for the purposes of characterizing the solution, we will want to use the FOC's from this planning problem in some cases. This requires that the solution is interior. The usual approach to guarantee interiority is to use Inada conditions. We use a version of these here to guarantee that c, $1 - l - bn$, and n are interior. The version that we use is stronger than usual and necessary because of the inclusion of private information and fertility.

Assumption 3 *Assume that both u and h are bounded above by 0, and unbounded below. Note that this implies that $\eta < 0$ must hold for concavity of overall utility and hence, an Inada condition on n is automatically satisfied. Finally, we assume that $h(1) < 0$.*[11]

Under this assumption, it follows that consumption, leisure and fertility are all strictly positive. This is not enough to guarantee that the solution is interior however, since hours worked might be zero. Indeed, there is no way to guarantee that $l > 0$ in this model. This is because of the way hours spent raising children enter the problem. Because of this feature of the model, it might be true that the marginal value of leisure exceeds the marginal product of an hour of work even when $l = 0$. The usual way of handling this problem by assuming that $h'(1) = 0$ will not work in this case since we know that $n > 0$. Hence, the marginal value of leisure at zero work will always be positive, even if $h'(1) = 0$. Because of this, when continuation utility is sufficiently high, it is always optimal for work to be zero.

In addition to this, in some cases, there are types that never work. This will be true when it is more efficient for a type to produce goods through the indirect method

[10] In our numerical examples the value function is convex even without the use of lotteries. In the Hosseini et al. (2010), we study a special case where we can show that the constraint correspondence is convex.

[11] This would hold, for example, if $h(\ell) = \frac{\ell^{1-\sigma}}{1-\sigma}$ with $\sigma > 1$.

of having children and having their children work in the future than through the direct method of working themselves. This will hold for a worker with productivity θ, if $\theta < \frac{E(\theta)}{bR}$. That is, $l(w,\theta) = 0$ for all w if $\theta < \frac{E(\theta)}{bR}$. For this reason, we will rule this situation out by making the following assumption:

Assumption 4 *Assume that, for all i, $\theta_i > \frac{E(\theta)}{bR}$.*

This assumption does not guarantee that $l(w,\theta) > 0$ for all w, but it can be shown that when continuation utility is low enough, $l > 0$. As we will show below, this is sufficient to guarantee that a stationary distribution exists.

In what follows, we will simply assume that $l > 0$ for most of the paper. We will return to this issue below when we show that a stationary distribution exists.

4 Properties of the Model

In this section, we lay out the basic properties of the model. These are:

1. A version of the resetting property for the infinite horizon version of the model,
2. A result stating that there is a stationary distribution over per capita variables, and
3. A version of the Inverse Euler Equation adapted to include endogenous fertility.

Taken together these imply that, when $\beta R = 1$, there is no immiseration in per capita terms but, there is immiseration in dynasty size. When $\beta R > 1$, this need not hold.

4.1 The Resetting Property

We have shown in the context of a two period model, children's consumption is independent of parent's promised utility when $\theta = \theta_H$. Here we will show that a similar property holds in the infinite horizon version of the model. This can be derived from the first order conditions of the recursive formulation. Taking first order conditions with respect to $n(\theta_I)$, $l(\theta_I)$ and $w'(\theta_I)$ respectively, gives us the following equations:

$$\begin{aligned}\pi(\theta_I)\frac{1}{R}v(w'(\theta_I)) &= b\left(\lambda\pi(\theta_I)+\sum_{\hat{\theta}<\theta_I}\mu(\theta_I,\hat{\theta})\right)\left[h'\left(1-l(\theta_I)-bn(\theta_I)\right)+\right.\\ &\qquad \left.\beta\eta\left(n(\theta_I)\right)^{\eta-1}w'(\theta_I)\right]\end{aligned}$$

$$\pi(\theta_I)\theta_I+\left(\lambda\pi(\theta_I)+\sum_{\hat{\theta}<\theta_I}\mu(\theta_I,\hat{\theta})\right)h'\left(1-l(\theta_I)-bn(\theta_I)\right)=0$$

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

Combining these gives

$$v(w'(\theta_I))-\eta w'(\theta_I)v'(w'(\theta_I))=-bR\theta_I. \tag{12}$$

We can see that $w'(w,\theta_I)$ is independent of promised continuation utility. That is, $w'(w,\theta_I)=w'(\hat{w},\theta_I)$ for all w, $\hat{w}$. Denote by w_0 this level of promised continuation utility – $w_0=w'(w,\theta_I)$.

The resetting property means that once a parent receives a high productivity shock, the per capita allocation for her descendants is independent of the parents level of wealth – an extreme version of social mobility holds.

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 if $n(w,\theta)$ is independent of (w,θ). 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 not a.s. convergence to the lower bound of continuation utilities. When $n(w,\theta)$ is not constant, the argument involves more steps. Assume that n is bounded above and below – $0<a\leq n(w,\theta)\leq a'$. Then, the fraction of descendants being assigned to w_0 next period is at least $\frac{\pi(\theta_I)a}{(1-\pi(\theta_I))a'}$. Again then, we see that there will not be a.s. immiseration. We summarize this discussion in a Proposition.

Proposition 5 *Assume that v is continuously differentiable and that there is a unique solution to 12. Then, continuation utility has a 'resetting' property, $w'(w, \theta_I) = w_0$ for all w.*

Intuitively, the reason that the resetting property holds here mirrors the argument given above in the two period 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.

Next, we argue that a similar property holds when continuation utility is low enough for any type. That is, even as promised utility, w, gets lower and lower, continuation utility, $w'(w, \theta)$, is bounded away from $-\infty$.

To this end, we show that as $w \to -\infty$, the optimal allocation converges to $c = 0$, $l = 1$, $n = 0$. None of the incentive constraints are binding at this allocation and hence, the optimal allocation has properties similar to those in the full information case. Formally:

Proposition 6 *Suppose that V is continuously differentiable and strictly convex. Then there exists a $\underline{w}_i \in \mathbb{R}$, such that*

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

See Appendix A.1 for the proof.

A key step in the proof is to show that when promised utility is sufficiently low, incentive problems, 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 h is unbounded below.

This property is one of the key technical findings in the paper. It can also be shown that, with utility unbounded below, this also holds in models with exogenous fertility. Loosely speaking, as w gets smaller, the allocations look more and more similar to full information allocations, whether fertility is endogenous or exogenous. What makes an endogenous fertility model different from an exogenous one is the properties of full information allocations – continuation utility is bounded below (by shock-specific resetting values for per child continuation utility) when fertility is endogenous.

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

Corollary 7 *Suppose that V is continuously differentiable and strictly convex. 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.*[12]

4.2 Stationary Distributions

The results from the previous section effectively rule out *a.s.* immiseration as long as n is bounded away from 0. This is not quite enough to show that a stationary distribution exists however. 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 and could be growing (or shrinking), we must also show that the growth rate of population is also stationary. 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, $T\Psi$:

$$T(\Psi)(A) = \int_w \sum_\theta \pi(\theta) \mathbf{1}_{\{(\theta,w);w'(\theta,w)\in A\}}(w,\theta) n(\theta,w) d\Psi(w) \quad (13)$$
$$\forall A : \text{Borel Set in } \mathbb{R}$$

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

[12] Proof can be found in the Supplementary Appendix. See Hosseini et al. (2010).

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

1. For all $w \in [\underline{w}, \overline{w}]$, there is a solution to problem P1';
2. For all $w \in [\underline{w}, \overline{w}]$, $w'(w, \theta) \in [\underline{w}, \overline{w}]$;
3. $n(w, \theta)$ and $w'(w, \theta)$ are continuous functions of w on $[\underline{w}, \overline{w}]$;
4. $\gamma(\Psi)$ is bounded away from zero for the probabilities on $[\underline{w}, \overline{w}]$.

We proceed by defining $\underline{w}$ and $\overline{w}$.

For any fixed $w < 0$, consider the problem:

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

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

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

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 $\overline{w}$ be this fixed point. Since $g(w) < w$ in a neighborhood of 0, it follows that $\overline{w} < 0$. Following Corollary 7, choose $\underline{w} = \underline{w}(\overline{w})$.

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

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

Now, we are ready to prove our main result about the existence of a stationary distribution.[13] Let $M([\underline{w}, \overline{w}])$ be the set of regular probability measures on $[\underline{w}, \bar{w}]$.

Theorem 8 *Assume that for all $w \in [\underline{w}, \overline{w}]$, 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. 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.

[13]The model analyzed here has a lot in common with the 'Ak' model with overall population (N) playing the role of k. Thus, in situations where $\gamma > 1$, there is no issue about increasing scarcity of fixed factors such as land. It is straightforward to add fixed factors to the model – e.g., $Y = F(L, D)$ where L is aggregate hours and D, the amount of land is fixed – but the proof we give below that a stationary distribution exists would have to be modified in technically non-trivial ways.

One version of this adaptation is fairly simple however. This is in the case that the mean population growth rate from the stationary distribution is $\gamma = 1$. In this case, it can be shown that, by setting the initial population, N, so that $F_1(N \times L, C) = E(\theta)$ where L is the mean per capita effective labor supply – $L = \int_w \sum_\theta \theta h(w, \theta) d\Psi$ –,Ψ is still a stationary distribution over continuation utilities with the fixed factor included in the model. The assumption that N is fixed in a steady state is a natural one since there is decreasing returns to scale in $F(\cdot, \cdot)$ with respect to labor. A similar result hold in Barro and Becker (1989).

In computed examples we have found that a stationary distribution exists for all cases – even when $\gamma \neq 1$.

By definition,

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

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

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

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

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 is in contrast to the models with exogenous fertility where a shrinking fraction of the population will have an ever growing fraction of aggregate consumption.[14]

The resetting property at the top has important implications about intergenerational social mobility. In fact, it makes sure that any smart parent will have children with a high level of wealth - as proxied by continuation utility. Finally, there is a lower bound on how much of this mobility occurs:

Remark 9 *Suppose that $\bar{w} = w_0$. Choose $A > 0$ so that $\frac{n(w,\theta_I)}{n(w,\theta)} \geq A$ for all w and θ. Suppose that $l(w, \theta_I) > 0$, for all $w \in [\underline{w}, w_0]$, then for any $\Psi \in M([\underline{w}, \bar{w}])$, we have:*

$$\hat{T}(\Psi)\left(\{w_0\}\right) \geq \frac{\pi_I A}{1 - \pi_I + \pi_I A}.$$

[14]There is a technical difficulty with extending this Theorem to settings with a continuum of types. A sufficient condition for the result to hold is that $w'(w, \theta)$ is increasing in θ.

See Supplementary Appendix (Hosseini et al. (2010))for proof.

The resetting property – at the top and the bottom – also has major implications about intergenerational transmission of wealth. That is, children's wealth – proxied by promised utility – is independent of parent's wealth when either parents wealth is very low or their productivity is very high. This is different than the previous work by Farhi and Werning (2007). However, since we have a limited set of theoretical results concerning the policy functions, further quantitative work is needed in order to see which model performs better relative to the data on intergenerational wealth mobility – such as the analysis provided by Charles and Hurst (2003).

Theorem 8, although the main theorem of the paper, says very little about uniqueness and stability as well as its derivation. The main problem is with endogeneity of population. This feature of the model, makes it very hard to show results regarding uniqueness or stability. In the Supplementary Appendix, we give an example of an environment with two values of shocks two productivity. In this case, under the resetting assumption, we are able to characterize one stationary distribution and show that given the class of distributions considered, the stationary distribution is unique. This procedure, as described in the Supplementary Appendix, can be used to construct at least one stationary distribution. The main idea for the construction is to start from full mass at the resetting value $\underline{w}_I$ and iterate the economy until convergence.

The difficulty in proving uniqueness and stability of the stationary distribution, depends heavily on the fact that fertility is endogenous. Endogeneity of fertility, implies that the transition function for promised value is not Markov. That is, the set of possible per capita promised values in the next period is not of unit measure. This in turn implies that this transition function can have multiple eigenvalues and eigenvectors each corresponding to a population growth rate γ and a stationary distribution Ψ. Therefore, we suspect that there are example economies in which stationary distribution is not unique.

4.3 Inverse Euler Equation and a Martingale Property

An important feature of dynamic Mirrleesian models with private information is the Inverse Euler Equation. Golosov et al. (2003) extend the original result of Rogerson (1985) and show that in a dynamic Mirrleesian model with private information, when utility is separable in consumption and leisure, the Inverse Euler equation holds when processes for productivity come from a general class. Here we will show that a version of the Inverse Euler equation holds. To do so, consider problem (P1). Suppose the multiplier on promise keeping is λ and the multiplier on 11 is $\mu(\theta, \hat{\theta})$. Then the first order condition with respect to $W'(\theta)$ is the following:

$$\pi(\theta)\frac{1}{R}V_W(N'(\theta), W'(\theta)) + \lambda\pi(\theta)\beta + \beta\sum_{\theta>\hat{\theta}}\mu(\theta, \hat{\theta}) - \beta\sum_{\theta<\hat{\theta}}\mu(\hat{\theta}, \theta) = 0.$$

Define $\mu(\theta, \hat{\theta}) = 0$, if $\hat{\theta} \geq \theta$. Summing the above equations over all θ's, we have

$$\frac{1}{R}\sum_{\theta}\pi(\theta)V_W(N'(\theta), W'(\theta)) + \beta\lambda + \beta\sum_{\theta}\sum_{\hat{\theta}}\mu(\theta, \hat{\theta}) - \beta\sum_{\theta}\sum_{\hat{\theta}}\mu(\hat{\theta}, \theta) = 0.$$

Moreover, from the Envelope Condition:

$$V_W(N, W) = -\lambda.$$

Therefore, we have

$$\sum_{\theta}\pi(\theta)V_W(N'(\theta), W'(\theta)) = \beta R V_W(N, W).$$

Now consider the first order condition with respect to $C(\theta)$:

$$\pi(\theta) + \lambda\pi(\beta)N^{\eta-1}u'\left(\frac{C(\theta)}{N}\right) \quad + \quad N^{\eta-1}u'\left(\frac{C(\theta)}{N}\right)\sum_{\hat{\theta}}\mu(\theta, \hat{\theta}) - N^{\eta-1}u'\left(\frac{C(\theta)}{N}\right)\sum_{\hat{\theta}}\mu(\hat{\theta}, \theta) = 0.$$

Thus,

$$V_W(N'(\theta), W'(\theta)) = \frac{\beta R}{N^{\eta-1}u'\left(\frac{C(\theta)}{N}\right)}.$$

which implies that

$$V_W(N_{t+1}(\theta^t), W_t'(\theta^t)) = \frac{\beta R}{N_t(\theta^{t-1})^{\eta-1}u'\left(\frac{C_t(\theta^t)}{N_t(\theta^{t-1})}\right)}.$$

Hence, we can derive the Inverse Euler Equation:

$$E\left[\frac{1}{N_{t+1}(\theta^t)^{\eta-1}u'\left(\frac{C_{t+1}(\theta^{t+1})}{N_{t+1}(\theta^t)}\right)}|\theta^t\right] = \frac{\beta R}{N_t(\theta^{t-1})^{\eta-1}u'\left(\frac{C_t(\theta^t)}{N_t(\theta^{t-1})}\right)}. \tag{14}$$

An intuition for this equation is worth mentioning. Consider decreasing per capita consumption of an agent with history θ^t and saving that unit. There will be R units available the next day that can be distributed among the descendants. We increase consumption of agents of type θ^{t+1} by $\epsilon(\theta_{t+1})$ such that:

$$\begin{aligned} &n_t(\theta^t)\sum_\theta \pi(\theta)\epsilon(\theta) = R \\ &u'(c_{t+1}(\theta^t,\theta))\epsilon(\theta) = u'(c_{t+1}(\theta^t,\theta'))\epsilon(\theta') = \Delta. \end{aligned}$$

The first is the resource constraint implied by redistributing the available resources. The second one makes sure that the incentives are aligned. In fact it implies that the change in the utility of all types are the same and there is no incentive to lie. The above equations imply that

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

Since the change in utility from this perturbation must be zero, we must have $\beta\Delta = u'(c_t(\theta^t))$. Replacing in the above equation leads to equation (14). We summarize this as a Proposition:

Proposition 10 *If the optimal allocation is interior and V is continuously differentiable, the solution satisfies a version of the Inverse Euler Equation:*

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

Moreover, $E_t N_{t+1}^{1-\eta}v'(w_{t+1}) = \beta R N_t^{1-\eta}v'(w_t)$. Hence, if $\beta R = 1$, $\frac{1}{N_{t+1}(\theta^t)^{\eta-1}u'\left(\frac{C_{t+1}(\theta^{t+1})}{N_{t+1}(\theta^t)}\right)}$ and $N_t^{1-\eta}v'(w_t)$ are non-negative Martingales.

If $\beta R = 1$, we see from above that $X_t = N_t^{1-\eta} v'(w_t)$ is a non-negative martingale. Thus, the martingale convergence theorem implies that there exists a non-negative random variable with finite mean, X_∞, such that $X_t \to X_\infty$ $a.s.$

As is standard in this literature, to provide incentives for truthful revelation of types, we must have 'spreading' in $(N'(\theta))^{1-\eta} v'(w'(\theta))$ (details in Hosseini et al. (2010)) as long as some incentive constraint is binding.[15] Thomas and Worrall (1990) have shown that in an environment where incentive constraints are always binding, spreading leads to immiseration. We can show a similar result in our environment, under some restrictions:

Theorem 11 *If* $\beta R = 1$ *and* $\Psi^*(\{w; \exists i \neq j \in \{1, \cdots, I\}, \mu(w; i, j) > 0\}) = 1$, *then* $N_t \to 0$ $a.s.$

The condition above ensures that there is always spreading in $N_t^{\eta-1} v'(w_t)$ when the economy starts from Ψ^* as initial distribution for w. In this case, the same proof as in Proposition 3 of Thomas and Worrall (1990) goes through and the above theorem holds. In fact, spreading implies that X_t converges to zero in almost all sample paths. Since $w_t$is stationary, N_t converges to zero almost surely.

The failure of the above condition implies that, there exists a subset of promised utilities A such that $\Psi^*(A) > 0$ and $\forall w \in A, \forall i, j,\ \mu(w, i, j) = 0$. For all $w \in A$, based on the analysis in the Supplemental Appendix, there is no spreading. The evolution of N_t in this case depends on the details of the policy function $w'(w, \theta)$. For example, suppose that for all $w \in A, \theta \in \Theta$, $w'(w, \theta) \in A$. Then if at some point in time $w_t \in A$, then $w_{t'} \in A$ for all subsequent periods t' and $N_{t'}$ will evolve so that $N_{t'}^{1-\eta} v'(w_{t'})$ is a fixed number - equal to $N_t^{1-\eta} v'(w_t)$. Since $w_{t'} \in [\underline{w}, \bar{w}]$, $N_{t'}$ will also be a finite number and is bigger than zero. In this case, the population would not be shrinking or growing indefinitely following these sequences of shocks and this happens for a positive measure of long run histories. This case, is similar to an example given in Phelan (1998) in which case a positive fraction of agents end up with infinite consumption and a positive fraction of agents end up with zero consumption. Kocherlakota (2010) constructs a similar example for a Mirrleesian economy.

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

[15] When promised utility of the parent is very high, it is possible that all types work zero hours. In this case, all types receive the same allocation and none of the incentive constraints are binding. See discussion in section 3.

from the more standard approach with exogenous fertility.

Finally, the fact that $N_{t+1}(\theta^t) \to 0$ $a.s.$ does not mean that fertility converges to zero almost surely, rather, it means that it is less than replacement (i.e., $n < 1$). Indeed, in computed examples, it can be shown that for certain parameter configurations (with $\beta R = 1$), $E(N_{t+1}(\theta^t)) \to \infty$ (i.e., $\gamma_{\Psi^*} > 1$). 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.

Similarly, it can be shown that when $\beta R > 1$, a stationary distribution over per capita variables still exists (see Theorem 8) but it need not follow that $N'(\theta^t) \to 0$ $a.s.$ In fact, numerical examples can be constructed in which $N_t \to \infty$ $a.s.$ (see the Supplementary Appendix in Hosseini et al. (2010)). In the numerical example, we solve the optimal contracting problem for an example with two values of shocks. For that example, we can calculate the Markov process for n_t induced by the policy functions $w'(w, \theta)$ and $n(w, \theta)$. If the economy starts from the stationary distribution, it can be shown that this Markov process is irreducible and acyclical and therefore has a unique stationary distribution Φ^*. Moreover, in our example $\int \log n d\Phi^* > 0$. Therefore,by theorem 14.7 in Stokey et al. (1989), the strong law of large number holds for $\log n_t$:

$$\frac{1}{T}\sum_{t=1}^{T} \log n_t \to \int \log n d\Phi^* > 0, as\ T \to \infty,\ a.s.$$

Notice that $N_{t+1} = N_t \times n_t$ and therefore $N_{t+1} = \log N_t + \log n_t$. This implies that $\frac{1}{t} \log N_t \to \int \log n d\Phi^*$ $a.s.$ and since $\int \log n d\Phi^* > 0$, $\log N_t \to \infty$ $a.s.$ Therefore, through our numerical example, we can see that when $\beta R > 1$, a per capita stationary distribution exists, population grows at a positive rate, and almost all dynasties survive. This is in contrast with Atkeson and Lucas (1992) where consumption inequality grows without bound for any value of β and R.

5 Extensions and Complementary Results

In this section, we discuss some complementary results. These are:

1. Implementing the optimal allocation through a tax system, and,

2. Differences between social and private discounting.

5.1 Implementation: A Two Period Example

Here, we discuss implementing the efficient allocations described above through 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. Similar results can be shown for the analogous 'wedges' in the infinite horizon setting.

As in the example in Section 2, 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.

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

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

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 12 *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. See Appendix. ■

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.

5.1.1 Positive Estate Taxes

In the above example, since all individual uncertainty is realized in the first period, there is no need for taxation of estate/capital. However, we know from Golosov et al. (2003) that if there is subsequent realization of individual shocks, the optimal allocation features a positive wedge on saving. The same logic applies here since we have a version of Inverse Euler Equation. Consider the model in section 3. Recall, from section 4 that we have

$$\frac{\beta R}{u'(c_t(\theta^t))} = \sum_{\theta_{t+1}\in\Theta} \pi(\theta_{t+1}) \frac{1}{n_t(\theta^t)^{\eta-1} u'(c_{t+1}(\theta^t, \theta_{t+1}))}$$

When, $c_{t+1}(\theta^t, \theta_{t+1})$ varies with realization of θ_{t+1}, the Jensen's inequality implies that

$$\frac{\beta R}{u'(c_t(\theta^t))} = \sum_{\theta_{t+1}\in\Theta} \pi(\theta_{t+1}) \frac{1}{n_t(\theta^t)^{\eta-1} u'(c_{t+1}(\theta^t, \theta_{t+1}))} < \frac{1}{\sum_{\theta\in\Theta} \pi(\theta_{t+1}) n_t(\theta^t)^{\eta-1} u'(c_{t+1}(\theta^t, \theta_{t+1}))}$$

and therefore

$$\beta R \sum_{\theta\in\Theta} \pi(\theta_{t+1}) n_t(\theta^t)^{\eta-1} u'(c_{t+1}(\theta^t, \theta_{t+1})) < u'(c_t(\theta^t))$$

The above equation implies that there is a positive wedge on saving. This positive wedge, in turn, translates into positive marginal tax rates on bequests. In fact, we think that it is relatively straightforward to extend the tax system in Werning (2010) in order to implement the optimal allocation in our environment and that tax system features positive taxes on bequests. This is in contrast with Farhi and Werning (2010), where they have shown that in order to implement an optimal allocation that has a stationary distribution, the planner must subsidize bequests.

5.2 Social vs. Private Discounting

In this section we consider a version of the two period example discussed in section 2 motivated by the papers by Phelan (2006) and Farhi and Werning (2007). This is to include a difference between the discount rate used by private agents and that used by the planner. Thus, we want to analyze the solution to the following planning problem:

$$\max_{c_1(\theta),n(\theta),l(\theta),c_2(\theta)} \sum_{\theta_L,\theta_H} \left[\pi(\theta)\left(u(c_1(\theta)) + h\left(1 - l(\theta) - bn(\theta)\right) + \hat{\beta}n(\theta)^{\eta}u(c_2(\theta))\right)\right]$$

s.t.

$$\sum_{\theta_L,\theta_H} \pi(\theta)\left(c_1(\theta) + \frac{1}{R}n(\theta)c_2(\theta)\right) \leq \sum_{\theta_L,\theta_H} \pi(\theta)\theta l(\theta) + K_0 \tag{15}$$

$$\begin{aligned} u(c_1(\theta_H)) + h(1 - l(\theta_H) - bn(\theta_H)) \quad + \quad & \beta n(\theta_H)^{\eta}u(c_2(\theta_H)) \geq u(c_1(\theta_L)) + \\ & h\left(1 - \frac{\theta_L}{\theta_H}l(\theta_L) - bn(\theta_L)\right) + \beta n(\theta_L)^{\eta}u(c_2(\theta_L)) \end{aligned} \tag{16}$$

This can be thought of as finding an alternative Pareto Optimal allocation when the planner puts more weight on children than individual parents do. As before, we have the following result:

$$\eta u(c_2(\theta_H)) = u'(c_2(\theta_H))c_2(\theta_H) + b\theta_H Ru'(c_2(\theta_H)).$$

From this, we can see that there is still 'resetting' at the top – $c_2(\theta_H)$ does not depend on K_0. The reasoning behind this is the same as that given in section 2.2, i.e., the planner always has an incentive of choosing the mix between C_2 and n so as to minimize the cost of providing any given level of utility in the second period.

Because of this, the same homogeneity/homotheticity logic still holds (and it does not depend on $\hat{\beta}$). Further, $c_2(\theta_H)$ does not depend on $\hat{\beta}$.

This suggests that our argument showing that there is a stationary distribution over continuation utilities will go through in this more general case. The one potential difficulty is proving an analog of theorem 8.

In addition, we have:

$$\left(\hat{\beta}-\beta\right)\frac{u'(c_1(\theta_H))}{\beta\lambda}+1=\frac{1}{\beta R}n(\theta_H)^{1-\eta}\frac{u'(c_1(\theta_H))}{u'(c_2(\theta_H))}$$

where λ is the Lagrange multiplier on the resource constraint.

As can be seen from this, when $\hat{\beta}=\beta$, this gives the usual Euler equation. When $\hat{\beta}>\beta$ this equation shows that, in general, there is an extra force to increase n. This is because $c_2(\theta_H)$ is independent of $\hat{\beta}$ and the term $\left(\hat{\beta}-\beta\right)\frac{u'(c_1(\theta_H))}{\beta\lambda}$ is strictly positive. Formally, holding $c_1(\theta_H)$ fixed, increasing $\hat{\beta}$ increases $n(\theta_H)$.

Intuitively, when $\hat{\beta}>\beta$ the planner wants to increase second period utility (relative to the $\hat{\beta}=\beta$ case). Since $u(c_2(\theta_H))$ does not depend on $\hat{\beta}$ the only channel available to do this is through increasing $n(\theta_H)$.

In sum, when the planner is more patient than private agents, he will encourage more investment both by increasing population size and increasing savings. Thus, this approach has important implications for population policy over and above what it implies about long run inequality.

Appendix

A Proofs

A.1 Proof of Proposition 6

We first prove the following lemma:

Lemma 13 *Suppose Assumptions 3 and 4 hold, then the value function and the policy functions satisfy the following properties:*

$$\begin{aligned}
\lim_{w\to-\infty} v(w) &= -\sum_i \pi_i\theta_i \\
\lim_{w\to-\infty} c(w,\theta_i) &= 0 \\
\lim_{w\to-\infty} n(w,\theta_i) &= 0 \\
\lim_{w\to-\infty} l(w,\theta_i) &= 1
\end{aligned}$$

Proof. Consider the following set of function:

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

Moreover define the following mapping on S as

$$T\hat{v}(w) = \min \sum_j \pi_j \left[c_j - \theta_j l_j + \frac{1}{R} n_j \hat{v}(w'_j) \right]$$

s.t.

$$\sum_j \pi_j \left[u(c_j) + h(1 - l_j - bn_j) + \beta n_j^\eta w'_j \right] \geq w$$

$$u(c_j) + h(1 - l_j - bn_j) + \beta n_j^\eta w'_j \geq u(c_i) + h(1 - \frac{\theta_i l_i}{\theta_j} - bn_i) + \beta n_i^\eta w'_i, \ \forall j > i$$

$$\begin{aligned}
l_j + bn_j &\leq 1 \\
c_j, l_j, n_j &\geq 0
\end{aligned}$$

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 $w_n \to -\infty$ such that for some i, $n(w_n, \theta_i) \to \bar{n}_i > 0$. For each $j \neq i$, define $\bar{n}_j = \liminf_{n\to\infty} n(w_n, \theta_j)$, then we must have

$$\liminf_{n\to\infty} T\hat{v}(w_n) \geq \sum_j \pi_j [-\theta_j(1 - b\bar{n}_j) + \frac{1}{R}\bar{n}_j \left[-\sum_k \pi_k\theta_k \right]$$

Note that by Assumption 4, we have

$$b\theta_j > \frac{1}{R}\sum_k \pi_k\theta_k, : \forall j$$

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

$$-\theta_j + b\bar{n}_j\theta_j - \frac{1}{R}\bar{n}_j\sum_k \pi_k\theta_k \geq -\theta_j$$

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

$$\liminf_{n\to\infty} T\hat{v}(w_n) > -\sum_j \pi_j\theta_j$$

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

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

If h is bounded above and below, define

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

By construction,

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

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 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

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

Hence, we have that

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

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, $h(1 - \frac{\theta_i l_m(\theta_i)}{\theta_j} - bn_m(\theta_i))$, converges to $h(1 - \frac{\theta_i}{\theta_j})$. This implies that for m large enough

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

and for m large enough the allocation is incentive compatible.

Therefore, The utility from the constructed allocation is the following:

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

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,

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

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

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

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

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

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

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

By Lemma 13, we must have:

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

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

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

In addition,

$$\begin{aligned}
\pi_{n-1}\lambda(w) + \sum_{j<n-1} \mu(n-1,j;w) - \mu(n,n-1;w) &= \frac{\pi_{n-1}}{u'(c(w,\theta_{n-1}))} < \frac{\epsilon}{N} \\
\Rightarrow \mu(n-1,j;w) &< \frac{2\epsilon}{N}
\end{aligned}$$

By an inductive argument, we have

$$\mu(i,j;w) < \frac{a_i \epsilon}{N}$$

where $a_{n-1} = 1, a_{n-2} = 2, a_{n-i} = a_{n-1} + \cdots + a_{n-i+1} + 1$. If we pick N so that $a_1 < N$, we have that

$$\mu(i,j;w) < \epsilon, \ \forall w < W.$$

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

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

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 5(resetting at top) we know that there is a w_0 such that:

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

Moreover, from the first order conditions, we know that

$$\eta v'(w'(w,\theta_1))w'(w,\theta_1) - v(w'(w,\theta_1)) \leq bR\theta_1.$$

Therefore, by the intermediate value theorem, there exists a unique $\underline{w}_i > -\infty$ which satisfies

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

Moreover, by substituting first order conditions, we get

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

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

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

Continuity of v' implies that

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

■

A.2 Implementation

A.2.1 Distortions

The constrained efficient allocation $c_1^*(\theta), l^*(\theta), n^*(\theta), c_2^*(\theta)$ solves the following problem

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

s.t.

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

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

Assuming the solution is interior, it satisfies the following first order conditions:

$$
\begin{aligned}
u'(c^*_{1H})(1+\frac{\mu}{\pi_H}) &= \lambda \\
u'(c^*_{1L})(1-\frac{\mu}{\pi_L}) &= \lambda \\
h'(1-l^*_H-bn^*_H)(1+\frac{\mu}{\pi_H}) &= \lambda\theta_H \\
h'(1-l^*_L-bn^*_L)-\frac{\mu}{\pi_L}\frac{\theta_L}{\theta_H}h'(1-\frac{\theta_L l^*_L}{\theta_H}-bn^*_L) &= \lambda\theta_L \\
\left[-bh'(1-l^*_H-bn^*_H)+\beta\eta n_H^{*\eta-1}u(c^*_{2H})\right](1+\frac{\mu}{\pi_H}) &= \lambda\frac{1}{R}\frac{\pi_H}{\pi_L}c^*_{2H} \\
&\quad \left[-bh'(1-l^*_L-bn^*_L)+\beta\eta n_L^{*\eta-1}u(c^*_{2L})\right] \\
-\left[-bh'(1-\frac{\theta_L l^*_L}{\theta_H}-bn^*_L)+\beta\eta n_L^{*\eta-1}u(c^*_{2L})\right]\frac{\mu}{\pi_L} &= \lambda\frac{1}{R}c^*_{2L}
\end{aligned}
$$

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

$$\max u(c_1)+h(1-l-bn)+\beta n^\eta u(c_2)$$

s.t.

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

As a first step, we assume that T is differentiable and that y is interior for both types. Then the FOCs are the following:

$$
\begin{aligned}
u'(c_1) &= \lambda_1 \\
h'(1-l-bn) &= \lambda_1\theta(1-T_y(\theta l,n,c_2)) \\
R\lambda_1 &= \lambda_2 \\
-bh'(1-l-bn)+\beta\eta n^{\eta-1}u(c_2) &= \lambda_2 c_2+\lambda_1 T_n(\theta l,n,c_2) \\
\beta n^\eta u'(c_2) &= n\lambda_2+\lambda_1 T_{c_2}(\theta l,n,c_2)
\end{aligned}
$$

Comparing the FOC's for the planner with these, we see immediately that $T_n(\theta_H l^*_H, n^*_H, c^*_{2H}) = T_y(\theta_H l^*_H, n^*_H, c^*_{2H}) = T_{c_2}(\theta_H l^*_H, n^*_H, c^*_{2H}) = 0$ – there are no (marginal) distortions on the decisions of the agent with the high shock. Moreover, from the FOC's of the planner's problem we get:

$$\begin{aligned}\left[-bh'(1-l_L^*-bn_L^*)\pi_L + bh'(1-\frac{\theta_L l_L^*}{\theta_H}-bn_L^*)\mu\right]\frac{1}{\pi_L-\mu} &+ \beta\eta n_L^{*\eta-1}u(c_{2L}^*) \\ &= \frac{1}{R}u'(c_{1L}^*)c_{2L}^*.\end{aligned}$$

We know that

$$\begin{aligned}1-\frac{\theta_L l_L^*}{\theta_H}-bn_L^* &> 1-l_L^*-bn_L^* \\ \Rightarrow h'(1-\frac{\theta_L l_L^*}{\theta_H}-bn_L^*)\mu &< h'(1-l_L^*-bn_L^*)\mu \\ bh'(1-\frac{\theta_L l_L^*}{\theta_H}-bn_L^*)\mu - bh'(1-l_L^*-bn_L^*)\pi_L &< bh'(1-l_L^*-bn_L^*)\mu \\ &\quad -bh'(1-l_L^*-bn_L^*)\pi_L \\ \left[-bh'(1-l_L^*-bn_L^*)\pi_L + bh'(1-\frac{\theta_L l_L^*}{\theta_H}-bn_L^*)\mu\right]\frac{1}{\pi_L-\mu} &< -bh'(1-l_L^*-bn_L^*).\end{aligned}$$

Hence,

$$\frac{1}{R}u'(c_{1L}^*)c_{2L}^* - \beta\eta n_L^{*\eta-1}u(c_{2L}^*) < -bh'(1-l_L^*-bn_L^*).$$

From the FOC of the consumer's problem,we have

$$0 < T_n(\theta_L l_L^*, n_L^*, c_{2L}^*) = -bh'(1-l_L^*-bn_L^*) - \frac{1}{R}u'(c_{1L}^*)c_{2L}^* + \beta n_L^{*\eta-1}u(c_{2L}^*).$$

Next, we turn to $T_y(\theta_L l_L^*, n_L^*, c_{2L}^*)$. From above, we have:

$$\begin{aligned}h'(1-l_L^*-bn_L^*)\pi_L - \mu\frac{\theta_L}{\theta_H}h'(1-\frac{\theta_L l_L^*}{\theta_H}-bn_L^*) &= \lambda\theta_L\pi_L \\ h'(1-l_L^*-bn_L^*)\left[\pi_L-\mu\frac{\theta_L}{\theta_H}\right] &< \lambda\theta_L\pi_L \\ h'(1-l_L^*-bn_L^*)\left[\pi_L-\mu\frac{\theta_L}{\theta_H}\right] &< \theta_L u'(c_{1L}^*)(\pi_L-\mu) \\ h'(1-l_L^*-bn_L^*) &< \theta_L u'(c_{1L}^*)\end{aligned}$$

Thus, from the FOC's of the agent's problem, we see that $T_y(\theta_L l_L^*, n_L^*, c_{2L}^*) > 0$.

Finally, since there are only shocks in the first period, it is never optimal to distort the savings decision for either type. Because of this, it follows that $T_{c_2}(\theta_L l_L^*, n_L^*, c_{2L}^*) = 0$.

A.2.2 Proof of Remark 12

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:

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

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,

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

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 differentiability of T and properties of marginal taxes follow from the discussion above.

References

Acemoglu, D., M. Golosov, and A. Tsyvinski (2008): "Political Economy of Mechanisms," *Econometrica*, 76, 619–641.

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.

——— (1995): "Efficiency and Equality in a Simple Model of Efficient Unemployment Insurance," *Journal of Economic Theory*, 66, 64–88.

Barro, R. J. (1974): "Are Government Bonds Net Wealth?" *Journal of Political Economy*, 82, 1095–1117.

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.

Charles, K. and E. Hurst (2003): "The Correlation of Wealth Across Generations," *Journal of Political Economy*, 111, 1155–1182.

Doepke, M. and R. M. Townsend (2006): "Dynamic Mechanism Design with Hidden Income and Hidden Actions," *Journal of Economic Theory*, 126, 235–285.

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 (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 (2010): "Risk Sharing, Inequality and Fertility, Supplemental Appendix," http://www.econ.umn.edu/~shour004/appendix2.pdf.

Jones, L., A. Schoonbroodt, and M. Tertilt (2008): "Fertility theories: can they explain the negative fertility-income relationship?" National Bureau of Economic Research Cambridge, Mass., USA.

KHAN, A. AND B. RAVIKUMAR (2001): "Growth and Risk-Sharing with Private Information," *Journal of Monetary Economics*, 47, 499–521.

KOCHERLAKOTA, N. R. (2010): *The New Dynamic Public Finance*, Princeton University Press.

MIRRLEES, J. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175–208.

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.

PHELAN, C. AND R. M. TOWNSEND (1991): "Computing Multi-Period, Information-Constrained Optima," *The Review of Economic Studies*, 58, 853–881.

ROGERSON, W. (1985): "Repeated Moral Hazard," *Econometrica*, 53, 69–76.

RUDIN, W. (1991): *Functional Analysis*, McGraw-Hill.

SLEET, C. AND Ş. YELTEKIN (2006): "Credibility and endogenous societal discounting," *Review of Economic Dynamics*, 9, 410–437.

STOKEY, N. L., R. E. LUCAS, AND E. C. PRESCOTT (1989): *Recursive Methods in Economic Dynamics.*

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.

WERNING, I. (2010): "Nonlinear Capital Taxation," Working Paper, MIT.

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