Essays on Dynamic Public Policy

A Thesis Submitted to the Faculty of the Graduate School of the

University of Minnesota

by

Hakki Yazici

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Narayana R. Kocherlakota, Adviser

AUGUST, 2009

to my parents, Cengiz and Neziha ...

Acknowledgements

This dissertation would not come into existence if it were not for the invaluable help and support of numerous individuals. First and foremost, I am grateful to my adviser Narayana R. Kocherlakota for his continuous guidance and support; I have learned an enormous amount from him on many aspects of research and professional life. I thank Larry E. Jones, Chris J. Phelan, and Paul Povel for their services on my exam committees and insightful comments on the work in progress. Arpad Abraham, Laurence Ales, Cristina Arellano, V.V. Chari, John T. Dalton, Seda Ertac, Kenichi Fukushima, Turkmen Goksel, Jonathan Heathcote, Patrick Kehoe, David Lagakos, Tommy Leung, Pricila Maziero, Guillermo Ordonez, Ayca Ozdogan, Fabrizio Perri, Arthur J. Rolnick, Jose-Victor Rios-Rull, Daniel Samano, Anderson Schneider, Ali Shourideh, Ctirad Slavik, Adam Slawski, Richard Todd, Michael E. Waugh, Kevin Wiseman, Cengiz Yazici, Kuzey Yilmaz, the members of the Public Economics workshop at the University of Minnesota, and the seminar participants at the Federal Reserve Bank of Minneapolis, the SED meetings in Cambridge, Carnegie Mellon University, University College London, University of Zurich, EIEF, Sabanci University, and Koc University provided several helpful comments and suggestions; thank you all. Finally, a special thanks goes to Mariana Ferreiro for her priceless friendship.

Abstract

This dissertation considers optimal government policy when agents in the economy are privately informed about their various characteristics. The first chapter deals with optimal education policy. Raising children is an important productive activity for a society since children's outcomes depend on their parents' investments. This paper develops an intergenerational framework in which adult outcomes are determined by parental investment, analyzes Pareto efficient allocations, and derives implications for policy. There are two key frictions: first, parental altruism type is private information; second, it is impossible for society to monitor parental investment. The main characterization result is that in any ex post Pareto efficient allocation, in any generation, society should transfer extra resources to all poor parents. This implies all agents, including the poor, should live above a certain welfare level, independent of whether or not society cares about them. Regarding implementation, the paper considers a market structure in which parents cannot sign contracts binding their descendants. Under such a market structure, implementing any Pareto efficient allocation requires government intervention. A feature shared by all Pareto efficient income tax schedules is that income taxes of agents with currently low income are negative.

The second chapter analyzes efficient allocation of resources in an economy in which agents are initially heterogeneous with regard to their wealth levels and whether they have ideas or not. An agent with an idea can start a business that generates random returns. Agents have private information about (1) their initial types, (2) how they allocate their resources, and (3) the realized returns. The unobservability of returns creates a novel motive for subsidizing agents who have ideas but lack resources to invest in them. The unobservability of initial types and actions implies that the subsidy that poor agents with ideas receive is limited by incentive compatibility: the society should provide other agents with enough incentives so that they do not claim to be poor and have ideas. The paper then provides an implementation of the constrained-efficient allocation in an incomplete markets setup that is similar to the U.S. Small Business Administration's Business Loan Program.

Contents

Chapter 1

Efficient Investment in Children and Implications for Policy

1.1 Introduction

It is mostly parents who choose the level of human capital investment their children receive. For instance, parents choose how many books to buy their children, how much time to spend reading to them, or whether to pay for private tutors for their children. It is empirically an established fact that these parental investments are significant in determining adult outcomes.[1] Hence, parental investments are important productive activities from society's perspective. However, it is hard to monitor the level of

[1] Todd and Wolpin (2006) is a recent paper that provides evidence on the importance of parental investment in explaining children's cognitive test score gaps. This provides an indirect evidence on the importance of parental inputs in determining adult outcomes since, as it is documented by Leibowitz (1974), Neal and Johnson (1996), and Keane and Wolpin (1997), cognitive tests taken during adolescence are predictive of adult labor market outcomes. For more direct support of the effect of parental investments in determining future outcomes see Cunha and Heckman (2006).

investment each child receives from her parents. This implies that society potentially faces an agency problem regarding investment in children: society has to provide parents with the right incentives so as to make them invest in their children.

This paper develops an intergenerational model in which parental investment in children is subject to such an agency problem in order to accomplish two specific goals. The first goal is to analyze the structure of the entire set of Pareto efficient allocations. The main characterization result is that, independent of social preferences, in all generations, poor parents should receive subsidies from the rest of the society. This implies all agents, including the poor, should live above a certain welfare level, independent of whether or not society cares about them. The second goal is to explore the types of market arrangements and policies that attain Pareto efficiency.

Specifically, the paper considers a dynastic model with heterogenous intergenerational altruism. In any generation, there are two types of parents: *altruistic* parents care about their children whereas *selfish* do not. Agents live for two periods. In the childhood period, they simply receive human capital investment from their parents which determines their output next period, when they become parents.[2] Investment in children is in terms of forgone consumption, and there is diminishing marginal returns to investment. In parenthood, agents first realize their altruism type and then choose how to divide family resources between consumption and investment in children.

I make two crucial informational assumptions. First, parents privately observe whether they are altruistic or not. Second, parental investment in a child is private

[2]Becker and Tomes (1979), Loury (1981), and Becker and Tomes (1986) are seminal papers that model adult output as being determined by parental human capital investments undertaken during childhood.

information at the time of investment. Society learns the level of human capital investment a child receives from her parents by observing the child's adult outcome, and, thus, with a lag. The unobservability of parental investment implies that there is an agency problem regarding parental investment in children. Society cannot force parents to invest in their children; parents must be provided with incentives to do so. Since parents are already replaced by their children by the time investment becomes observable, the only way to provide parents with incentives is by rewarding or punishing their children in terms of consumption.

In the absence of informational frictions, Pareto efficient resource allocation involves two separate steps: (1) productive efficiency requires equating marginal returns to parental investment across children within any generation; (2) distributive efficiency then requires making transfers between agents so as to achieve the desired consumption distribution, which depends on the welfare criterion of society. However, I show that, under informational frictions, productive efficiency and distributive efficiency cannot be completely separated. Attaining productive efficiency implies certain properties that are common to all Pareto efficient distributions of consumption, independent of any social welfare criterion. The first major goal of this paper is to establish these properties.

I show in any Pareto efficient allocation, for all generations, all poor parents receive subsidies. That is, society should transfer additional resources to parents who have not received much human capital investment from their parents, and, hence, have low output. This then implies all agents at all periods live above a certain welfare level. The intuition is simple. Without a subsidy, the amount of investment children receive would be limited by parental output. Due to diminishing marginal returns to investment, this means children of poor parents would have relatively high marginal returns. By redistributing resources from rich parents to poor ones, society reallocates investment from low return children to high return children. As a result, average returns to childhood investment increase, and, hence, there are more resources next period to be distributed among those society cares about. However, since parental investment is unobservable, in order to make parents invest, society has to provide them with incentives in terms of their children's consumption, because otherwise parents would consume the subsidies intended for investment in children. Therefore, subsidies toward the poor implies welfare transfers toward them, which implies, all parents, including the poor, live above a certain welfare level.

Consequently, subsidizing the poor is required by Pareto efficiency because it enhances productive efficiency. Since there are no assumptions about any particular social welfare function, equality or insurance considerations play *no* role in the subsidy result. In this sense the subsidy result in the current paper is distinct from most of the redistribution results in public finance as these results are generally conditional on assumptions about social welfare functions that value equity. This distinction is a very important one because the result in this paper establishes that subsidizing the poor is a necessary condition to be efficient whereas previous results merely say redistributing to the poor is needed if we assume that society has a particular social welfare criterion.

The second major goal of the paper is to find actual market structures and tax systems that implement Pareto efficient allocations. The paper first considers a market in which initial generation parents can sign exclusive contracts ex ante that are legally binding for all their descendants. A version of the First Welfare Theorem holds: the market equilibrium is Pareto efficient.[3] As a result, if such a market is in operation, any government intervention would be based on redistributive motives rather than efficiency considerations. However, such a market structure is highly unrealistic as it crucially requires enforcement of contracts which, when signed by an agent, bind all future descendants.

The paper then considers implementing Pareto efficient allocations in a more realistic market setup in which contracts signed by parents do not bind their children. Agents are allowed to sign all other types of contracts that respect informational constraints. The market equilibrium is Pareto inefficient, which implies government intervention is essential under such a market structure. An interesting property of this market arrangement is that there is a subset of Pareto efficient allocations for which there is *no* tax implementation. Then, the paper focuses on Pareto efficient allocations that can be implemented with income taxes and derives an important property shared by all Pareto efficient income tax systems: in any generation, income taxes of agents with currently low income are negative. This result is essentially a translation of the subsidy result about Pareto efficient allocations into a result on Pareto efficient income taxation.

Finally, in section 1.6, I pay particular attention to the Pareto efficient allocation that arises when the social objective puts direct welfare weights only on initial generation parents. The aim of this section is to investigate whether or not the celebrated immiseration result, which essentially says that in dynamic moral hazard environments almost all agents' consumption diverges to the lower bound, holds in

[3] The result that ex ante efficient allocations under private information can be decentralized as through ex ante contracts is originally due to Prescott and Townsend (1984) for static economies. Atkeson and Lucas (1992) and Golosov and Tsyvinski (2007) provide similar decentralizations for different dynamic private information economies.

the current environment. I show that the unobservability of investment creates a novel force which breaks down the immiseration result.

Becker and Tomes (1979), Loury (1981), and Becker and Tomes (1986) are seminal papers that develop dynastic models in which individuals' adult income is a function of the parental investment they receive during childhood. These papers consider equilibria under different incomplete market structures with the aim of establishing investments within family as an important source of income inequality. A recent paper in the same tradition is Aiyagari, Greenwood, and Seshadri (2002) which quantitatively compares the performance of different market structures in terms of welfare, aggregate output, and inequality in a model in which parental inputs are important determinants of adult outcomes. The main distinction of the current paper from all these papers is that, instead of assuming the existence of market incompleteness exogenously, I endogenize market incompleteness by assuming that there are key informational frictions regarding investment in children. Furthermore, all the aforementioned papers are essentially positive analyzes, whereas the current paper focuses on the properties of Pareto efficient allocations and implications for optimal policy.[4]

Banerjee and Newman (1991), Galor and Zeira (1993), and Aghion and Bolton (1997) are also related to the current paper. These papers study the link between the distribution of wealth and productive activity in dynastic economies in which there is individual level production with diminishing returns to scale. Even though they do not focus on policy analysis, an implicit conclusion common to all these papers is that redistribution towards the poor can enhance productive efficiency.[5] The main

[4]An exception is Loury (1981) which provides some policy experiments. Furthermore, Seshadri and Yuki (2004) also investigates numerical policy experiments in an incomplete markets setup. However, unlike the current paper, neither of these papers analyze optimal policy.

[5]Indeed, Aghion and Bolton (1997) explicitly carries out such a normative analysis. The paper

difference with the current paper is that I focus on Pareto efficient allocations and the resulting optimal public policies.

This paper is also related to a number of recent papers that explore constrained efficient allocations in intergenerational settings with private information frictions, such as Phelan (2006), Farhi and Werning (2007), and Farhi and Werning (2008). The most important difference between the papers above and the current one is that the papers above focus on social planning problems that maximize specific social welfare functions whereas I analyze the properties shared by the entire set of Pareto efficient allocations. This is a very important distinction because analyzing all Pareto efficient allocations allows me to, as Stiglitz (1987) puts it, "separate out efficiency considerations from the value judgements associated with choices among Pareto efficient points." As a result, I isolate and focus on the implications of productive efficiency while the efficiency concepts used by previous papers include social preference for equality.

The rest of the paper is structured as follows. In section 2, I introduce the model formally. Section 3 characterizes properties common to all ex post Pareto efficient allocations. Section 4 discusses two different implementations of Pareto efficient allocations. Finally, section 5 concludes.

takes output maximization as the social objective and shows that when there are moral hazard problems due to unobservable effort, redistributing resources from the rich to the poor may increase total output, boosting economic growth.

1.2 Model

1.2.1 Environment

The economy is populated by a continuum of unit measure of dynasties who live for a countable infinity of periods, $t = 1, 2, ...$ Each agent within a dynasty lives for two periods: childhood and parenthood. In each period t, each parent has a child. In period $t + 1$, the child becomes a parent and replaces the period t parent.[6] The economy starts at the beginning of period one with the first members of each dynasty: initial parent generation and their children.

A period t parent has expected utility preferences, with utility function:

$$U_t = u(c_t) + \beta_t U_{t+1},$$

where c_t is period t parent's consumption, $u(\cdot)$ is instantaneous utility function, β_t is period t parent's altruism factor, and U_{t+1} is the utility of the offspring of the period t parent. Agents have to consume a non-negative amount, i.e. $c_t \geq 0$. The above specification of altruism is consistent with agents having a preference over the entire future consumption of their dynasty:

$$U_t = \sum_{\tau=t}^{\infty} \mathbb{E}_t[\prod_{s=t}^{\tau-1} \beta_s u(c_\tau)].$$

I make the following assumptions about the instantaneous utility function.

[6] The terms period t parent and period t adult will be used interchangeably to refer to agents who have children in period t.

Assumption 1 *The instantaneous utility function* $u : \Re_+ \to \Re$ *satisfies:*

a. $u', -u'' > 0$.

b. $\lim_{c \to 0} u'(c) < \infty$ *and* $\lim_{c \to \infty} u'(c) = 0$.

b. $u(0) = \kappa > -\infty$.

The first part of the assumption is standard: instantaneous utility functions is strictly increasing and strictly concave in consumption. The second part of the assumption simply says that the lowest period utility that an agent can receive is finite.

A period t parent chooses how to allocate family resources between own consumption, c_t, and human capital investment in the child, n_t. So, investment in children is in terms of forgone consumption.[7] Investment cannot be negative, i.e., $n_t \geq 0$. When invested n_t units in period t, a child produces $y_{t+1} = f(n_t)$ units in period $t+1$ as an adult. During childhood, an agent is merely a production unit operated by the parent and does not make any decisions. I make the following somewhat standard assumptions on the production function f.

Assumption 2 *The production function* $f : \Re_+ \to \Re_+$ *satisfies:*

a. $f', -f'' > 0$.

b. $\lim_{n \to 0} f'(n) = \infty$.

c. $f(0) = 0$.

The first part of the assumption is standard and says children's outcomes are strictly increasing and strictly concave in the amount of investment they receive from their

[7] I follow Loury (1981), Becker and Tomes (1986), and Cunha and Heckman (2006), among others, in assuming that investment in children is in terms of forgone consumption.

parents. The second part is simply the Inada condition. Finally, the last condition means if children do not receive any investment then their adult outcomes are zero.

Parents are heterogenous with respect to the level of their altruism towards their children. Some of them care about their children $\beta = \bar{\beta}$, and some do not, $\beta = \underline{\beta}$, where $0 \leq \underline{\beta} < \bar{\beta} < 1$. Define $H = \{\underline{\beta}, \bar{\beta}\}$ to be the set of all altruism factors. Parents with altruism factor $\bar{\beta}$ are called altruistic, and those with altruism factor $\underline{\beta}$ are called selfish. Agents learn how altruistic they are only after becoming parents. Consequently, a period t child learns how altruistic she is in the beginning of period $t+1$, before making any decisions. The probability that a child in period t is going to be an altruistic parent in period $t+1$ is $\mu(\bar{\beta})$. With probability $\mu(\underline{\beta})$, the child becomes a selfish parent. The distribution of β is i.i.d. across dynasties and time.

In period t, dynasties are different from each other only with respect to the realizations of altruism factors up to that period. Therefore, in period t, each dynasty is identified with its realized history of altruism factors, $h^t = (\beta_1, \beta_2, ..., \beta_t)$. Define H^t to be the set of all histories at time t. Let the notation $h^{t+j} \succeq h^t$ mean that h^{t+j} follows h^t (i.e. its first t components are equal to h^t). With some abuse of notation, define $\mu(h^t)$ to be the ex ante probability of a period t parent having history h^t. By the Law of Large Numbers, $\mu(h^t)$ is also the fraction of dynasties with history h^t at time t.

To simplify notation, I define $\delta_t(h^\tau) = \prod_{s=t}^{\tau-1} \beta_s$. In words, $\delta_t(h^\tau)$ is the cumulative altruism towards agent with history h^τ from her period t ancestor. For initial parents, the cumulative altruism towards generation τ history h^τ descendant is denoted by $\delta(h^\tau) = \prod_{s=1}^{\tau-1} \beta_s$.

I make two key informational assumptions. First, each parent privately observes her altruism factor. I believe this is a reasonable assumption as it is hard, if not possible, for outsiders to know how much a parent cares about her child. Furthermore, how a parent allocates family resources between consumption and investment in the child is not publicly observable in the period in which the decision is made. Society learns the level of human capital investment in a child by observing the child's adult income next period, and thus, with a lag. When the child becomes an adult society observes her output and deduces the investment she received retrospectively.

An *allocation* in this economy is defined to be $x \equiv (c_t, n_t)_{t \geq 1}$, where

$$c_t, n_t : H^t \rightarrow \Re_+.$$

Here $c_t(h^t)$ and $n_t(h^t)$ are period t consumption and investment levels for a dynasty with an altruism history of h^t. An allocation x is *feasible* if for all t,

$$\sum_{h^t} \mu(h^t)[c_t(h^t) + n_t(h^t)] \leq \sum_{h^{t-1}} \mu(h^{t-1}) f(n_{t-1}(h^{t-1})), \tag{1.2.1}$$

$$c_t(h^t), n_t(h^t) \geq 0, \tag{1.2.2}$$

where $f(n_0) > 0$ is given. The first condition merely says that for any generation t the sum of aggregate consumption and aggregate investment should be no greater than aggregate output.[8] The second condition ensures that consumption and investment are non-negative.

[8]Observe that one can write the feasibility condition as in (1.2.1), even though c_t and n_t are private information. The reason is that the sum $c_t + n_t$ is publicly observable.

1.2.2 Incentive-compatibility

There are two sources of private information in the model. First, there is hidden information: parents privately observe their altruism types. Second, parents are involved in hidden action: their consumption and human capital investment choices are unobservable. Hence, in this world, parents can deviate from an allocation recommended by the planner in three ways: they can lie about their altruism factor, they can choose an investment level that is different from what the planner recommended, or they can double deviate by doing both at the same time. A powerful Revelation Principle due to Myerson (1982) says that if an allocation is implemented as an equilibrium of an arbitrary mechanism, then it can also be implemented as a truth-telling and obedient equilibrium of a direct revelation mechanism. Therefore, without loss of generality, the rest of the paper restricts attention to allocations that can be implemented as truth-telling and obedient equilibria of direct revelation mechanisms.

In this subsection, I first define what it means for an allocation to be implemented as a truth-telling and obedient equilibrium of a direct revelation mechanism. Then, I define a subset of the set of all direct revelation mechanisms which I call the set of worst punishment direct revelation mechanisms. I prove that an allocation is implementable as a truth-telling and obedient equilibrium of a direct revelation mechanism if and only if it is implementable as a truth-telling and obedient equilibrium of a worst punishment direct revelation mechanism. This implies that one can further restrict attention to allocations that can be implemented as a truth-telling and obedient equilibrium of worst punishment direct revelation mechanisms. This is an important simplification because it is enough to check a single inequality to see whether an allocation can be implemented as a truth-telling equilibrium of a worst punishment direct revelation mechanism. This inequality is the incentive compatibility condition and the allocations that satisfy this condition are called incentive compatible.

In the last part of the subsection, I show that one can replace the incentive compatibility condition by a set of much simpler conditions, often referred to as temporary incentive compatibility constraints in the dynamic contracting literature.[9] Consequently, these inequalities completely represent the restrictions that informational problems place on the set of allocations achievable by the society.

A *direct revelation mechanism* is $(C, N) \equiv (C_t, N_t)_{t\geq 1}$, where $C_t, N_t : H^t \times \Re_+^{t-1} \to \Re_+$. Here, C_t and N_t are the outcome functions. The planner transfers a total of $C_t(h^t, \tilde{n}^{t-1}) + N_t(h^t, \tilde{n}^{t-1})$ to the dynasty with history $(h^t, \tilde{n}^{t-1})$ of reported altruism types and observed human capital investment levels, and recommends the parent to spend $N_t(h^t, \tilde{n}^{t-1})$ units on the child and the rest on her own consumption. Note that the outcome functions depend on the past observed history of investment levels as well as the history of reported types. This is due to the feature of the model that investment becomes publicly observable with one period lag.

A *dynastic strategy* is $(\hat{\sigma}, \hat{n}) \equiv (\hat{\sigma}_t, \hat{n}_t)_{t\geq 1}$, where $\hat{\sigma}_t : H^t \times \Re_+^{t-1} \to H$ and $\hat{n}_t : H^t \times \Re_+^{t-1} \to [C_t(h^t, \tilde{n}^{t-1}) + N_t(h^t, \tilde{n}^{t-1})]$. Here, $\hat{\sigma}_t(h^t, \tilde{n}^{t-1})$ is the report of the parent with history (h^t, n^{t-1}), and $\hat{n}_t(h^t, \tilde{n}^{t-1})$ is the investment of the same parent in her children. Define $\hat{\sigma}^t(h^t, \tilde{n}^{t-1})$ as the history of reports along $(h^t, \tilde{n}^{t-1})$ and $\hat{n}^t(h^t, \tilde{n}^{t-1})$ as the history of investment levels along the same path. Let Γ be the set of all dynastic strategies. Finally, define (σ, N) as the truth-telling and obedient dynastic strategy.

Define $W(\hat{\sigma}, \hat{n}|C, N)$ to be the expected lifetime payoff to a dynasty under the

[9] The term temporary incentive compatibility condition is due to Green (1987).

dynastic strategy $(\hat{\sigma}, \hat{n})$, given the direct revelation mechanism, (C, N) :

$$W(\hat{\sigma}, \hat{n}|C, N) =$$

$$\sum_{t=1}^{\infty} \sum_{h^t} \mu(h^t)\delta(h^t)u\Big[C_t\Big(\hat{\sigma}^t(\cdot), \hat{n}^{t-1}(\cdot)\Big) + N_t\Big(\hat{\sigma}^t(\cdot), \hat{n}^{t-1}(\cdot)\Big) - \hat{n}_t\Big(h^t, \hat{n}^{t-1}(\cdot)\Big)\Big].$$

An allocation x is implemented as a truth-telling and obedient equilibrium of a direct revelation mechanism (C, N) if two conditions hold. First, the payoff to each dynasty of telling the truth and obeying recommendations should be at least as large as the payoff from any other strategy. Second, consumption and human capital investment allocations along the equilibrium path of the direct revelation game should correspond to the allocation x. The formal definition is given below.

Definition 1 *An allocation x is implemented as a truth-telling and obedient equilibrium of a direct revelation mechanism if and only if*

a. For all $(\hat{\sigma}, \hat{n}) \in \Gamma$,

$$W(\sigma, H|C, N) \geq W(\hat{\sigma}, \hat{n}|C, N).$$

b. For all h^t,

$$C_t(h^t, N^{t-1}(\cdot)) = c_t(h^t),$$

$$N_t(h^t, N^{t-1}(\cdot)) = n_t(h^t).$$

By the Revelation Principle, Definition 1, together with the feasibility constraints (1.2.1) and (1.2.1), characterizes the set of all allocations that are achievable by society.

I define a *worst punishment direct revelation mechanism* to be a direct revelation mechanism that sets consumption and investment levels to zero at all nodes following a deviation from recommended investment. For each direct revelation mechanism, (C, N), there exists a corresponding worst punishment direct revelation mechanism, (C', N'), defined as follows:

For any $t \geq 1$ and $(h^t, \tilde{n}^{t-1}) \in H^t \times \Re_+^{t-1}$,

$$
\begin{aligned}
Z'_t(h^t, \tilde{n}^{t-1}) &= Z_t(h^t, \tilde{n}^{t-1}), \quad && \text{if } \tilde{n}^{t-1} = N^{t-1}(h^{t-1}, N^{t-2}(\cdot)); \\
Z'_\tau(h^\tau, \tilde{n}^{\tau-1}) &= 0, && \text{if else,}
\end{aligned}
$$

for $Z = C, N$. Observe that for any period t, one can divide the public histories of the direct revelation game, $(h^t, \tilde{n}^{t-1})$, into two groups. First, there are paths along which there have been no deviations from the recommended investment levels up to t. Second, there are histories in which there has been at least one deviation in the first $t-1$ generations. I call the first group of histories the obedience path. So, (C', N') is identical to (C, N) on the path of obedience. The mechanisms differ in their reaction to dynasties that deviate from the recommended investment level, i.e., off the obedience path. The worst punishment direct revelation mechanisms punish disobeying dynasties in the harshest way possible. Note that the set of all worst punishment direct revelation mechanisms is a subset of the set of all direct revelation mechanisms.

Lemma 1 ***(Worst punishment)*** *An allocation x is implemented as a truth-telling* *and obedient equilibrium of a direct revelation mechanism* (C, N) *if and only if it is implemented as a truth-telling and obedient equilibrium of the corresponding worst punishment direct revelation mechanism,* (C', N').

Proof.

Relegated to Appendix 1.8.1. ■

Consequently, the rest of the paper restricts attention to worst punishment direct revelation mechanisms. Observe that worst punishment mechanisms are much simpler objects compared to general direct revelation mechanisms. In fact, a worst punishment mechanism offers a dynasty an allocation x along the obedience path and identically zero transfers off the obedience path. To see this, let $C'_t(h^t, N'^{t-1}(h^{t-1}, N'^{t-2}(\cdot))) = c_t(h^t)$ and $N'_t(h^t, N'^{t-1}(h^{t-1}, N'^{t-2}(\cdot))) = n_t(h^t)$. In the rest of the paper, I refer to a worst punishment mechanism by the allocation it corresponds to on the obedience path since all worst punishment mechanisms are the same off the obedience path.

Restricting attention to worst punishment mechanisms is an important simplification because it allows one not to worry about the nodes that are reached after a deviation from an action recommended by the planner. At all such nodes, the planner will be transferring the dynasty zero units no matter what the dynasty reports. Consequently, reports following a deviation from recommended action are payoff irrelevant. Also, human capital investment has to be zero at any node following a deviation. These imply that one can redefine reporting and investment strategies independent of the history of investments.

Redefine a dynastic reporting strategy as a sequence of functions, $(\hat{\sigma}_t)_{t\geq 1}$, where $\hat{\sigma}_t : H^t \to H$. $\hat{\sigma}_t(h^t)$ gives a dynasty's report at node h^t, independent of the levels of investment made in the past. Now, I similarly redefine human capital investment as a function of the history of past types only. First, observe an implication of Lemma 1 : it is optimal from the point of view of a dynasty that deviates from the recommended investment level to consume all the resources given by the planner in the period of deviation. This is straightforward since under a worst punishment mechanism a dynasty's payoff following a deviation from recommended action is zero independent of the investment level the dynasty deviates to.

Therefore, without losing generality, we can restrict attention to investment strategies in which, at any period and node, the dynasty either invests at the level recommended by the planner to the reported type or invests zero. As a result, $\hat{\phi} = (\hat{\phi}_t)_{t\geq 1}$, where $\hat{\phi}_t : H^t \to \{0, 1\}$, is sufficient to represent a dynasty's investment strategy. Here, $\hat{\phi}_t(h^t) = 1$ means a dynasty that follows strategy $\hat{\phi}$ invests at the level recommended by the planner to the reported type at node h^t, $n_t(\hat{\sigma}^t(h^t))$, whereas $\hat{\phi}_t(h^t) = 0$ implies the dynasty chooses to invest zero.

Redefine $(\hat{\sigma}, \hat{\phi}) \equiv (\hat{\sigma}_t, \hat{\phi}_t)_{t\geq 1}$ as a dynastic strategy. Let (σ, ϕ) be the truth-telling and obedient strategy, where $\sigma_t(h^t) = \beta_t$ and $\phi_t(h^t) = 1$, for all t, h^t. Now, define the expected dynastic payoff under strategy $(\hat{\sigma}, \hat{\phi})$ given the worst punishment direct revelation mechanism x :

$$W(\hat{\sigma}, \hat{\phi}|x) = \sum_{t=1}^{\infty} \sum_{h^t} \mu(h^t)\delta(h^t)u\Big[\Big(\min_{h^\tau \prec h^t} \hat{\phi}_\tau(h^\tau)\Big)\Big(c_t(\hat{\sigma}^t(h^t)) + (1 - \hat{\phi}(h^t))n_t(\hat{\sigma}^t(h^t))\Big)\Big]. \tag{1.2.3}$$

Finally, I can define what it means for an allocation to be incentive compatible. An allocation x is incentive compatible if and only if there exists a worst punishment direct revelation mechanism x that implements it. The formal definition is given below.

Definition 2 ***(Incentive compatibility)*** *An allocation x is incentive compatible if and only if for all $(\hat{\sigma}, \hat{\phi})$,*

$$W(\sigma, \phi | x) \geq W(\hat{\sigma}, \hat{\phi} | x). \tag{1.2.4}$$

An allocation that is feasible and incentive compatible is called *incentive feasible.* These are the allocations society can achieve.

The remainder of this subsection shows that the incentive-compatibility conditions represented by (1.2.4) can be replaced with temporary incentive constraints. Given allocation x, define $V(h^t|x)$ to be the continuation utility to dynasty h^t from period $t+1$ on when truth-telling and obedience is followed in all the nodes succeeding h^t and there has been no detection of disobedience in the past:

$$V(h^t|x) = \sum_{\tau=t+1}^{\infty} \sum_{h^\tau \succ h^t} \mu(h^\tau | h^t) \delta_{t+1}(h^\tau) u(c_\tau(h^\tau)).$$

If a dynasty member has already deviated in the past, then the continuation utility from period $t+1$ onwards is equal to the expected discounted utility of consuming zero in all nodes from period $t+1$ onwards, which is equal to $\frac{1}{1-\mathbb{E}\beta}\kappa$, where $\mathbb{E}$ denotes the expectation operator.

Proposition 1 ***(Temporary incentive compatibility)*** *An allocation x satisfies* *incentive constraints* (1.2.4) *if and only if it satisfies for all* $t, h^{t-1}, \beta, \beta^o$

$$(IC_L) \qquad u(c_t(h^{t-1},\beta)) + \beta V(h^{t-1},\beta|x) \geq u(c_t(h^{t-1},\beta^o)) + \beta V(h^{t-1},\beta^o),$$

$$(IC_D) \qquad u(c_t(h^{t-1},\beta)) + \beta V(h^{t-1},\beta|x) \geq u(c_t(h^{t-1},\beta) + n_t(h^{t-1},\beta)) + \beta \tfrac{1}{1-\mathbb{E}\beta}\kappa,$$

$$(IC_{LD}) \qquad u(c_t(h^{t-1},\beta)) + \beta V(h^{t-1},\beta|x) \geq u(c_t(h^{t-1},\beta^o) + n_t(h^{t-1},\beta^o)) + \beta \tfrac{1}{1-\mathbb{E}\beta}\kappa. \tag{1.2.5}$$

Proof.

Relegated to Appendix 1.8.1. ■

The rest of the paper uses the three sets of inequalities in (1.2.5) to represent incentive compatibility of allocations. The first set of incentive-compatibility conditions, type L, is the usual one that says no parent should find it profitable to lie to be of the other type at any time and node. Type D incentive constraints say that parents should not find it profitable to disobey the investment level recommended by the planner. Finally, type LD constraints say that under an incentive-compatible allocation, parents should not be getting better off by double deviating: lying to be of the other type and investing zero in the child.

Observe that type D and type LD incentive constraints are novel in the dynamic contracting literature. These incentive constraints are very simple compared to the ones that usually arise in the presence of hidden action. This is due to the assumption that private information regarding actions (investment in children in the current context) is temporary.

1.2.3 Pareto Efficient Allocations

Now, I define Pareto efficient allocations for this economy.[10] In that regard, define $U(\beta^t|x) = u(c_1(\beta^t)) + \beta V(\beta^t|x)$ to be the dynastic welfare of agent with history h^t under allocation x.

Definition 3 *An allocation x^* is Pareto efficient if it is incentive feasible, delivers welfare $(U(\beta^t|x^*))_{t,\beta^t}$ and there is no other incentive feasible allocation x that delivers $(U(\beta^t|x))_{t,\beta^t}$ with $U(\beta^t|x) \geq U(\beta^t|x^*)$ for all β^t, t and at least one of these inequalities holding strictly.*

Pareto efficient allocations solve the following social planner's problem:

$$\max_x \sum_{t=1}^{\infty} \sum_{\beta^t} \mu(\beta^t) U(\bar{\beta}|x) \tag{1.2.6}$$

subject to for any t, β^t

$$U(\beta^t|x) \geq \bar{v}(\beta^t)$$

and subject to incentive-feasibility.

Observe that the social planning problem is indexed by $(\bar{v}(\beta^t))_{t,\beta^t \neq \bar{\beta}}$. Therefore, by changing the stochastic process $(\bar{v}(\beta^t))_{t,\beta^t \neq \bar{\beta}}$, we can trace the infinite-dimensional Pareto frontier of this economy.[11]

[10] Throughout the paper, the term Pareto efficient allocation refers to Pareto efficient allocation under informational problems.

[11] Obviously, not all such processes define a Pareto efficient allocation. For a stochastic process $(\bar{v}(\beta^t))_{t,\beta^t \neq \bar{\beta}}$ to define a Pareto efficient allocation, there should be at least one incentive feasible allocation delivering each agent a welfare level that is weakly greater than the welfare level assigned to her by that stochastic process.

1.3 Characterization of Pareto Efficient Allocations

The aim of this section is to analyze the properties that are common to all Pareto efficient allocations. Put differently, I am interested in features that are shared by all allocations on the Pareto frontier of the economy. I assume $\underline{\beta} = 0$ in the main body of the paper. This assumption is mainly for expositional purposes and is not needed for any of the results that follow.[12]

Before analyzing the Pareto problem of the previous section, I first analyze efficiency under two different benchmark environments, which makes it easier to understand the role of informational assumptions and the resulting agency problem.

1.3.1 Two Benchmark Cases

Full Information Benchmark. First, consider an economy that is identical to the one described in section 2 except for all economic activity is public information. I want to characterize the whole set of Pareto efficient allocations. An allocation x is *full information Pareto efficient* if it is feasible, delivers welfare $(U(h^t|x^*))_{t,h^t}$, and there is no other feasible allocation x that delivers $(U(h^t|x))_{t,h^t}$ with $U(h^t|x) \geq U(h^t|x^*)$ for all h^t, t and at least one of these inequalities holding strictly.

Full information Pareto efficient allocations solve the following social planner's problem:

$$\max_x \sum_{t=1}^{\infty} \sum_{h^t} \mu(h^t) U(\bar{\beta}|x) \tag{1.3.1}$$

[12] I show in subsection 5.3 of the paper that the main result of the paper holds as long as $\underline{\beta}$ is below an upper bound $\tilde{\underline{\beta}}$.

subject to for any t, h^t

$$U(h^t|x) \geq \bar{v}(h^t)$$

and subject to feasibility.

It follows simply from the first-order optimality conditions of the above problem that in any full information Pareto efficient allocation marginal returns to human capital investment in children is equated across children within any generation. This implies that in any generation all children receive same investment independent of which dynasty they belong to, i.e., $n_t(h^t)$ is independent of h^t. Since there are no ex ante productivity differences among children and there is diminishing marginal returns to investment, productive efficiency calls for an equal division of investment across children.

Under the full information assumption, productive efficiency is separated from distributive efficiency. Society invests in children in the productively efficient way and then distributes the output among agents in the way social norms call for.

Observable Investment Benchmark. Consider the same economy as before, but, now, the only source of private information is parental altruism types. In such a world, due to parental types being unobservable, only a subset of the set of all feasible allocation, those which are incentive compatible, are achievable by society. When the only source of private information is parental type, an allocation is incentive compatible if IC_L is satisfied in all generations and histories. Observe that the incentive constraints IC_D and IC_{LD} are not required for incentive compatibility when investment is observable. This is simply due to the fact that in this case the only way a parent can deviate is by lying to be of a different type; since investment is observable, parents have to obey the social planner's recommendation about how to allocate family resources.

Below I define a planner's problem, the solution of which gives the set of all observable investment Pareto efficient allocations as one varies $(\bar{v}(h^t))_{t,h^t \neq \bar{\beta}}$ appropriately. Observe that the only difference between the problem below and the social planning problem in (1.3.1) are the incentive constraints, IC_L.

$$\max_x \sum_{t=1}^{\infty} \sum_{h^t} \mu(h^t) U(\bar{\beta}|x) \tag{1.3.2}$$

subject to for any t, h^t

$$U(h^t|x) \geq \bar{v}(h^t)$$

and subject to feasibility and $IC_L(h^{t-1}, \beta)$, for all h^{t-1}, β, β^o.

Similar to the full information case, in any observable investment Pareto efficient allocation, human capital investment in children is independent of their dynastic history. Since investment is still observable, there is nothing preventing society from ensuring that marginal returns to investment are equated across children. To see how this follows from the social planner's problem above, observe that the only way investment enters the problem is through the resource constraints; it does not enter incentive constraints. Therefore, productive efficiency and distributive efficiency are separated in the model with observable investment as well. The main difference from the full information case is that, due to types being unobservable, society has to take into account incentive constraints IC_L when distributing consumption among agents.

Consequently, both in the full information and observable investment benchmark cases, investment in children is independent of history. This is due to the fact that the agency problem regarding investment in children disappears when investment is observable.

1.3.2 Characterization of Pareto Efficient Allocations

This subsection analyzes Pareto efficient allocations in the economy introduced in section 2 when investment in children, as well as altruism types of parents, is private information. The goal is to prove that in any Pareto efficient allocation, for any generation, poor parents are subsidized, and, hence, all agents, including the poor, live above a certain welfare level.

The first step in analyzing Pareto efficient allocations is simplifying the incentive compatibility constraints. In that regard, Lemma 2 below shows one can disregard some incentive-compatibility conditions from the outset. First, given an allocation, if an altruistic parent does not lie to be a selfish parent and the latter does not deviate by disobeying the recommended investment, then the altruistic parent does not deviate by disobeying either. Hence, incentive constraints regarding deviations in which a currently altruistic parent disobeys (type L and type LD constraints) are automatically satisfied. Second, observe that for any h^{t-1}, if x satisfies $IC_{LD}(h^{t-1}, \underline{\beta})$, then it also satisfies $IC_L(h^{t-1}, \underline{\beta})$. In words, if selfish parents do not deviate by lying and disobeying, then they do not deviate by only lying either. This follows straightforwardly from the facts that selfish parents do not care about their descendants and investment in children is non-negative.

Lemma 2 *For any h^{t-1}, incentive constraints $IC_D(h^{t-1}, \bar{\beta})$, $IC_{LD}(h^{t-1}, \bar{\beta})$, and $IC_L(h^{t-1}, \underline{\beta})$*

are redundant.

Proof.

Relegated to Appendix 1.8.2. ■

As a result, in the rest of the paper, incentive compatibility conditions in which altruistic parents disobey and those in which selfish parents only lie are going to be omitted without loss of generality.

It is evident from the Pareto problems that social and parental objectives are not necessarily aligned. When parental investment is unobservable, this implies that society faces an agency problem regarding childhood investment. As a result, if society wants parents to invest in their children, it has to provide them with incentives to do so. This implies that resource cost is not the only social cost of investment in children; there is also an incentive cost. Since this incentive cost is potentially different for parents with different histories, it is no longer Pareto optimal to equate marginal returns to investment across children in any generation. To see the incentive cost of investment mechanically, observe that investment enters not only into the resource constraints but also incentive compatibility constraints in the social planner's problem. Consequently, productive efficiency and distributive efficiency are not separated when there is an agency problem in childhood investment.

Lemma 3 below which establishes two important properties of parental investment in children that are common to all Pareto efficient allocations. The subsidy result follows directly from these properties. The first property is that the Pareto efficient level of investment in children of selfish parents is zero in any generation. This follows directly from IC_D and $\underline{\beta} = 0$:

$$u(c_t^*(h^{t-1}, \beta)) \geq u(c_t^*(h^{t-1}, \beta) + n_t^*(h^{t-1}, \beta))$$

$$\Rightarrow n_t^*(h^{t-1}, \beta) = 0.$$

Since deviation from the recommended investment level is detected only after the children become adults, the only way society can make parents invest in their children is by punishing the children of deviating parents. However, since selfish parents do not care about their children at all, there is no reward/punishment mechanism that can make them invest in their children.

The second property of Pareto efficient investment in children is that, in all generations, children of all altruistic parents receive strictly positive investment. That is, even if society does not care at all about the welfare of an altruistic parent or any of her successors, it should still make sure that her child receives parental investment. Clearly, no notion of equality is responsible for the result. It follows solely from productive efficiency. Since there is diminishing marginal returns, society wants to exploit investment opportunities in all children. Smoothing investment across children in a period means higher average returns and, hence, higher output next generation, which the society can distribute among those it cares about.

Lemma 3 *For any allocation x^* in the set of Pareto efficient allocations,*

1. $n_t^*(h^{t-1}, \underline{\beta}) = 0$, *for all* h^{t-1}.

2. $n_t^*(h^{t-1}, \bar{\beta}) > 0$, *for all* h^{t-1}.

Proof.

Relegated to Appendix 1.8.2. ■

Now, I show the main result of this section. First, define

$$\Delta_t(h^t) \equiv n_t(h^t) + c_t(h^t) - f(n_{t-1}(h^{t-1})).$$

In words, $\Delta_t(h^t)$ is the net transfer that an agent with history h^t receives from society under allocation x. I say that an agent h^t is *subsidized* under allocation x if $\Delta_t(h^t) > 0$. I make one more definition. I say that an agent h^t lives in misery under allocation x if $c_\tau(h^\tau) = 0$, for all $h^\tau \succeq h^t$. In words, an agent lives in misery if own and all future descendants' consumption is equal to zero.

The main result of this section is Proposition 2 below. It has three parts. First, children of selfish parents have zero output when they become adults. This follows from part one of Lemma 3 : children of selfish parents do not receive any human capital investment.

Second, in any generation, poor agents receive subsidy from the rest of society, independent of the social welfare criterion. The intuition is very simple. According to part two of Lemma 3, productive efficiency requires all altruistic parents to invest in their children. This then implies that altruistic parents with no output should be subsidized so that they can invest in their children. Selfish agents with no output are subsidized because parental altruism is private information.

Third, in any Pareto efficient allocation, all agents live strictly above misery. Exploiting investment opportunities in children requires society to reward people who operate them, their parents, due to the private information nature of the investment. Then, selfish parents should also receive consumption, since otherwise they would lie to be altruistic and consume the transfers intended for altruistic agents. Consequently, Pareto efficiency requires society to ensure that none of its members live in misery, even when some of these agents are not valued at all by society.

Proposition 2 *In any Pareto efficient allocation x^*,*

1. *If $h^{t-1} = (h^{t-2}, \underline{\beta})$, then $y_t^*(h^{t-1}, \beta) = 0$.*

2. *If $y_t^*(h^t) = 0$, then $\Delta_t^*(h^t) > 0$.*

3. *No agent lives in misery.*

Proof.

Part 1. By Lemma 3 part one, $n_{t-1}^*(h^{t-2}, \underline{\beta}) = 0$, which implies $y_t^*(h^{t-2}, \bar{\beta}, \beta) = 0$.

Part 2. By Lemma 3 part two, $n_t^*(h^{t-1}, \bar{\beta}) > 0$. If $y_t^*(h^{t-1}, \beta) = 0$, then $\Delta_t^*(h^{t-1}, \bar{\beta}) > 0$. Then, from incentive-compatibility condition $IC_{LD}(h^{t-1}, \underline{\beta})$, it follows that $\Delta_t^*(h^{t-1}, \underline{\beta}) > 0$ too.

Part 3. $n_t^*(h^{t-1}, \bar{\beta}) > 0$ and $IC_{LD}(h^{t-1}, \underline{\beta})$ imply $c_t^*(h_{t-1}, \underline{\beta}) > 0$. Thus, no selfish agent is miserable. Since the altruism factor is privately observed, $IC_L(h^{t-1}, \bar{\beta})$ implies altruistic agents are not miserable either. ■

Note that this subsidy result is a powerful one, as it does not depend on society's preferences. Even if a society does not care at all about its poor citizens, it still should subsidize them on productive efficiency grounds. Similarly, independent of the social welfare criterion, society should guarantee a minimum level of welfare for all its members. Equity or insurance considerations play no role in these results. The results depend crucially on two features of the model. First, all parents in the model "operate" individual level "production technologies," their children, with decreasing returns to scale. Obviously, there would be no issue of productive efficiency if the current model were an endowment economy or there were constant returns to scale to investment in children; in that case, there would be no need to subsidize any agent unless it is required by a particular social welfare criterion. Second, parental investment is unobservable, which implies that in order to exploit these investment opportunities, society has to provide parents with incentives, and, thus, welfare.

1.4 Implementation

This section is concerned with the following question. What kinds of markets and policies are sufficient to attain Pareto efficiency? I analyze the implementation of Pareto efficient allocations in two different market setups.

1.4.1 Market with Dynastic Contracts

The first market structure I consider is one in which generation one parents can sign contracts ex ante, before any uncertainty is realized, and, importantly, these contracts are legally binding for all their descendants. There is a continuum of intermediaries that are owned equally by all dynasties. In the beginning of generation one, before realization of any uncertainty, each intermediary signs a contract with a continuum of initial parents which binds all following generations. The contracts are offered competitively, and initial parents choose the contract with the highest promised ex ante expected utility.

A contract is $x = (c_t, n_t)_{t\geq 1}$, where $c_t, n_t : H^t \to \Re_+$. Here, $c_t(h^t) + n_t(h^t)$ is the total amount of resources transferred to a dynasty who have reported h^t and obeyed the intermediaries' recommendations up to period t. Any dynasty that deviates from the actions recommended by an intermediary receives zero units forever after the deviation is detected. Thus, without loss of generality, I confine intermediaries to offer worst punishment direct revelation mechanisms. Facing such a contract, each dynasty chooses a participation strategy, $(\hat{\sigma}_t, \hat{\phi}_t)_{t\geq 1}$, receives either $c_t(\hat{\sigma}_t(h^t)) + n_t(\hat{\sigma}_t(h^t))$ or zero depending on past deviation from recommendation, and chooses how to allocate the total amount received. Intermediaries take into account that each generation a parent can choose a different allocation of total resources handed to her than what is intended by the intermediary. The incentive compatibility conditions summarized by (1.2.5) in section 2 completely characterize the set of allocations intermediaries can offer to dynasties through period one contracts.

Intermediaries trade one period claims with each other. Denote by q_t the price of a claim a_t that pays one unit of consumption good in period $t+1$. I consider symmetric equilibria in which all intermediaries are price takers. Letting $\underline{U}$ be the equilibrium utility each dynasty receives, representative intermediary's problem is

$$\max_{x,a} \sum_{t=1}^{\infty} \sum_{h^t} \mu(h^t) \prod_{s=1}^{t-1} q_s \Big[f(n_{t-1}(h^{t-1})) - c_t(h^t) - n_t(h^t) \Big] \tag{1.4.1}$$

subject to (1.2.5) and $\sum_h \mu_h U(h|x) \geq \underline{U}$.

A formal definition of market equilibrium is as follows:

Definition 4 *A competitive equilibrium is prices $(q_t)_{t\geq 1}$, allocations $(c_t, n_t)_{t\geq 1}$, trades of one period claims $(a_t)_{t\geq 1}$, and a utility level $\underline{U}$ such that*

1. *Taking $(q_t)_{t\geq 1}$ and $\underline{U}$ as given, $(c_t, n_t, a_t)_{t\geq 1}$ solves the intermediaries' problem* (1.4.1);

2. *Initial generation parents choose contracts that offer the highest utility;*

3. *Claims market clears every period.*

I now prove a version of the First Welfare Theorem for this private information economy. This result was originally proven by Prescott and Townsend (1984) for a static private information economy. Atkeson and Lucas (1992) and Golosov and Tsyvinski (2007) prove the same result for different dynamic private information economies. The result follows from the observation that the intermediary's problem is the dual of the social planner's problem of maximizing ex ante dynastic utility. The proof uses a property of the equilibrium that is easy to show: in equilibrium, perfect competition drives intermediary's profits to zero.

Proposition 3 *The competitive equilibrium in which initial parents can sign contracts that bind all their descendants is Pareto efficient.*

Proof.

It suffices to show that equilibrium x solves the social planner's problem of maximizing ex ante utility. Suppose this is not true. That the claims market clears every period implies that the equilibrium allocation satisfies aggregate feasibility. Incentive compatibility of the equilibrium allocation follows directly from the intermediary's problem. Hence, it must be that ex ante expected utility dynasties receive under the efficient allocation, call it U^*, is strictly greater than the utility under equilibrium allocation $\underline{U}$. Observe that the efficient allocation is available to the intermediary since $U^* > \underline{U}$. Now, consider a third allocation x' identical to the efficient one except for $c_1'(\bar{\beta}) = c_1^*(\bar{\beta}) - \epsilon$. For $\epsilon > 0$ small, the ex ante utility this allocation offers, U', is still strictly greater than $\underline{U}$, and x' is incentive compatible. Observe that under x' the intermediary makes strictly positive profits, which is a contradiction. ■

The analysis above suggests private markets achieve Pareto efficiency without government intervention. As a result, any intervention would be based on redistributive motives rather than efficiency considerations. Put differently, government should intervene only if social norms call for a different Pareto efficient allocation than the one achieved under private markets. [13] However, observe that efficiency of private markets crucially depends on a very strong assumption about the set of contracts available: period one parents can write contracts that legally bind all their descendants. This assumption is highly unrealistic, as in most countries around the world contractual arrangements signed by individuals are not binding for their children. The next subsection analyzes a different, more realistic market structure in which there is no intergenerational trade.

1.4.2 Market without Intergenerational Trade

Consider a market setup in which contracts signed by parents do not bind their children. Agents are allowed to sign all other types of contracts that respect informational constraints. The aim is to provide an implementation of Pareto efficient allocations in this market setup. Towards this goal, let $T = (T_t)_{t \geq 1}$ be an income tax system where

$$T_t : \Re_+^t \rightarrow \Re.$$

[13] Under this market structure, markets only attain the Pareto efficient allocation that maximizes the ex ante welfare of initial generation parents. This result is in the same spirit as Bernheim (1989) which shows that many Pareto efficient allocations cannot be achieved in dynastic models with altruistically linked individuals.

Here, $T_t(y^t)$ is the income tax levied on a period t parent with a dynastic history of incomes $y^t = (y_1, ..., y_t)$. Therefore, a dynasty's problem now reads:

$$\max_x \sum_\beta \mu(\beta) U(\beta|x) \tag{1.4.2}$$

subject to for all t and h^t,

$$c_t(h^t) + n_t(h^t) \leq y_t(h^t) - T_t(y^t(h^t)).$$

Definition 5 *An equilibrium with income taxes in a market without intergenerational borrowing and lending is income taxes T and an allocation x such that*

1. *Given T, x solves the dynastic problem;*

2. *Government budget balances every period.*

It is easy to see that the unique equilibrium allocation under laissez-faire is equivalent to the autarkic allocation. In that allocation, call it x^a, $n_t^a(h^t) = 0$, for all $t \geq 1$, $h^t \neq (\bar{\beta}, ..., \bar{\beta})$. By Proposition 3, this is clearly Pareto inefficient.

Proposition 4 *The laissez-faire market equilibrium without intergenerational borrowing and lending is Pareto inefficient.*

Under the equilibrium allocation, all agents who have at least one selfish ancestor have zero income. Therefore, altruistic parents who have at least one selfish ancestor cannot invest in their children, which means society cannot exploit very high return investment opportunities in children of altruistic parents. This breaks productive efficiency and creates Pareto inefficiency. As a result, as long as intergenerational borrowing and lending is absent, government intervention in the economy is essential since there is room for Pareto improvement over the existing market. A natural next step is to see what a government should do in order to ensure Pareto efficiency.

An interesting property of this market structure is that there is *no* tax implementation for some Pareto efficient allocations. To see this, remember that in the current model an agent's income is determined by parental investment during her childhood, i.e., $y_t = f(n_t(h^{t-1}))$. As a result, current income, y_t, does not reveal any information to the government about the agent's current type, β_t. But, then, the tax system cannot differentiate between agents with the same history of ancestors, h^{t-1}. Thus, period t parents with histories $(h^{t-1}, \bar{\beta})$ and $(h^{t-1}, \underline{\beta})$ receive the same transfers in any allocation that is implemented with income taxes. This implies that Pareto efficient allocations in which, in some period t, transfers depend on agents' period t types cannot be implemented through taxes on income under the market structure with no intergenerational trade.

If there were an intergenerational market of some kind, it could be possible for the government to deduce information about whether an agent is altruistic or not by observing her actions in the market. However, when such markets are absent, government has no way of screening altruistic and selfish parents with the same dynastic histories. As a result, Pareto efficient allocations in which selfish and altruistic parents with the same histories receive different transfers cannot be implemented through taxes. Fortunately, it is possible to completely characterize the set of Pareto efficient allocations that can be implemented by income taxes, as shown by Proposition 5.

Define $Y_t^* = \{y^t \in \Re_+^t | y_\tau = f(n_{\tau-1}^*(h^{\tau-1})), \text{for all } \tau \leq t \text{ and } h^{\tau-1} \preceq h^{t-1}, \text{ for some } h^{t-1}\}$. In words, y^t is in Y_t^* if there exists some type h^{t-1} that receives income history y^t in the Pareto efficient allocation.

Now set the tax system as follows:

$$\begin{aligned} T_t^*(y^t) &= -\Delta_t^*(h^{t-1}), \quad \text{if } y^t \in Y_t^*; \\ T_t^*(y^t) &= y^t, \quad \text{if else.} \end{aligned} \tag{1.4.3}$$

Proposition 5 *A Pareto efficient allocation x^* can be implemented in a market with no intergenerational borrowing and lending through taxes on income if and only if $(\bar{v}(h^t))_{t\geq 1, h^t \neq \bar{\beta}}$ is such that $IC_{LD}(h^{t-1}, \underline{\beta})$ holds with equality for all h^{t-1}.*

Proof.

$(\Rightarrow)$: Suppose a Pareto efficient allocation x^* can be implemented by an income tax system in a market with no intergenerational trade. This implies $c_t^*(h^{t-1}, \beta) + n_t^*(h^{t-1}, \beta)$ is independent of β, for all h^{t-1}, which implies $IC_{LD}(h^{t-1}, \underline{\beta})$ holds with equality for all h^{t-1}.

$(\Leftarrow)$: Now, suppose $IC_{LD}(h^{t-1}, \underline{\beta})$ holds with equality for all h^{t-1} in a Pareto efficient allocation x^*. Then, since $IC_D(h^{t-1}, \underline{\beta})$ binds for all h^{t-1} in any Pareto efficient allocation, $c_t^*(h^{t-1}, \beta) + n_t^*(h^{t-1}, \beta)$ is independent of β. Set the taxes as in (1.4.3).

Consider the dynastic problem with taxes. Define a dynasty's income strategy as $y = (y_t)_{t\geq 1}$, where $y_t : H^{t-1} \to \Re_+$. A dynastic income strategy gives a dynastic children investment strategy since income is a one-to-one function of childhood investment. The investment strategy then implies a consumption strategy through the budget constraint. Therefore, given taxes, the only object a dynasty is choosing is its income strategy. Let y^* be the income strategy in which $y_t^*(h^{t-1}) = f(n_{t-1}^*(h^{t-1}))$.

First, observe that if a dynasty chooses an income strategy y' such that for all t, h^{t-1}, $y_t'(h^{t-1}) = f(n_{t-1}^*(\sigma_{t-1}'(h^{t-1})))$, then by construction of the tax system, flow budget constraints imply that the dynasty receives allocation $(c_t^*(\sigma_t'(h^t)), n_t^*(\sigma_t'(h^t)))$, where $c_t^*(\sigma_t'(h^t))$ is the consumption of history h^t dynastic member. Hence, the expected discounted dynastic welfare under y' would be equal to $W(\sigma', \phi|x^*)$.

Second, suppose a dynasty chooses income strategy y. Suppose in some period t in some node h^t, y calls for setting human capital investment in the child to some level such that the dynastic income history next period is not going to be in Y_{t+1}^*. By the construction of income taxes, this implies, beginning from generation $t+1$ on, all output of the dynasty will be seized by the government. Hence, it is optimal for the period t member of the dynasty to consume all resources.

As a result, any potentially optimal income strategy y can be decomposed into two parts: an income strategy y' such that for all t, h^{t-1}, $y_t'(h^{t-1}) = f(n_{t-1}^*(\sigma_{t-1}'(h^{t-1})))$, and a decision made at each node about whether to deviate from the income path that occurs under y'. Denote the latter decision by $\nu' = (\nu_t')_{t\geq 1}$, where $\nu_t'(h^t) = 1$ means at node h^t the dynasty makes an amount of investment in children that makes the dynastic income equal to $y_{t+1}'(h^t)$. Thus, the expected discounted dynastic welfare under y is equal to $W(\sigma', \phi'|x^*)$, for $\phi' = \nu'$.

Then, under the tax system T^*, if a dynasty chooses y^*, then the dynasty receives (c^*, h^*) and the expected discounted utility, $W(\sigma, \phi|x^*)$. If a dynasty chooses any other y, then the dynastic utility is $W(\sigma', \phi'|x^*.)$ Then, by incentive compatibility of x^*, the dynasty chooses y^*. ■

Section 3 already proved that $IC_D(h^{t-1}, \underline{\beta})$ binds for all h^{t-1} in any Pareto efficient allocation. If it is also the case that $IC_{LD}(h^{t-1}, \underline{\beta})$ holds with equality for all h^{t-1}, then $c_t^*(h^{t-1}, \beta) + n_t^*(h^{t-1}, \beta)$ is independent of β, which means $\Delta_t^*(h^{t-1}, \beta)$ is independent of β.

In the rest of the subsection, I confine analysis to Pareto efficient allocations that can be implemented with income taxes and derive an important property of Pareto efficient taxes. The corollary below essentially translates the subsidy result about Pareto efficient allocations derived in section 3 into a result on income taxes. Remember that by Proposition 3, in any Pareto efficient allocation, x^*, $\Delta_t^*(h^{t-2}, \underline{\beta}, h) > 0$, for all h^{t-2}. Remember also that by the same proposition, $n_t^*(h^{t-1}, \underline{\beta}) = 0$, for all h^{t-1}. Combining these two results, I get the following corollary.

Corollary 1 *Under the market structure with no intergenerational trade, in all Pareto efficient income tax systems, all agents with zero income receive strictly positive subsidies, i.e.,* $T_t^*(y^{t-1}, 0) < 0$, *for all* t *and* $y^{t-1} \in Y_{t-1}^*$.

1.5 Discussion

In the first two subsections of this section, I discuss whether the main characterization result of the paper survives generalizations of the model along two important dimensions. The last subsection provides a robustness check: the main result of the paper does not depend on the assumption that $\underline{\beta}$ is exactly zero.

1.5.1 Public Investment in Children

Throughout the paper I assume that government can only indirectly regulate investment in children through subsidizing their parents. In real life, on the other hand, governments also try to control childhood investment directly through in-kind subsidies.[14]

A natural question then is: What happens to the main result if we equip the government in the model with in-kind subsidies? One way to answer this question is to expand the "production function" of children by introducing an additional investment good, g, which is publicly observable.[15] The adult output of a child is now equal to $f(n, g)$, where n is the investment in children that is carried out privately by the parent as before. Now, the government has two ways to regulate investment in children. First, it can provide a parent with resources and incentives to invest in her child. Second, the government can directly invest in the child through publicly observable investment good.[16] The latter way represents in-kind subsidies in the model.

Whether or not the main result of the paper still holds depends on how good public investment is in terms of substituting parental investment. To see this, consider the extreme case in which parental and public investments in children are perfect substitutes and $f(n, g) = (n + g)^{\alpha}$, $\alpha \in (0, 1)$. In this case, the government can completely control the level of investment children receive through public investment. The set of

[14]Head Start is an example of an in-kind subsidy program in the U.S. It is a preschool program that aims to prepare disadvantaged children to school. See Currie (2001) for more information about Head Start and various other early childhood education programs.

[15]It is important to note that the only aspect that is public about this good is information. Each child receives her publicly observable investment separately and this investment affects only her adult output. Put differently, there is no public good aspect to g.

[16]Alternative to direct public provision, one can also think that the government is transferring consumption goods to the parent and the transfer is conditional on investing a specific amount of g in the child. The exact way in-kind subsidy is provided is not important for the question I am after.

Pareto efficient allocations is identical to the one in the observable benchmark case studies in subsection 3.1. All children receive the same level of investment which implies, in any period, all agents have the same level of output. Therefore, there is no poor to subsidize. Furthermore, since investment is completely controlled by the government, there is no need to provide parents with incentives to invest in their children. As a result, some people do live in misery depending on the social welfare function. Consequently, if public investments can completely substitute parental investments without any extra costs, then the main characterization result does not hold.

If, on the other hand, we assume that some components of parental investment are essential in raising children and cannot be substituted by public investment, then one can show that the main result of the paper still holds provided that we still have the assumption that there are arbitrarily high returns to parental investment as the level of this investment decreases to zero, i.e. $\lim_{n \to 0} \partial f(n,g)/\partial n = \infty$.

1.5.2 Persistent Ability

The main characterization result of the paper follows from the intermediate result that the children of all altruistic parents, rich or poor, should receive some level of investment in any Pareto efficient allocation. In this subsection I ask the question: How does this intermediate result depend on the simplifying assumption that there are equal returns to investing in poor and rich children?

One way to answer this question is to modify the model so that adult output is not only a function of parental investment but it is also a function of the child's ability, θ, and ability is persistent over generations. So, suppose parents know their children's θ before they make the investment decision. A child with ability θ and investment n produces $f(n, \theta)$ next period. Suppose also that the children of parents who are highly able are more likely to be highly able themselves. Also, for simplicity, suppose children's ability is public information. In such a world, marginal returns to investing in rich children is on average higher than marginal returns to investing in poor children at any investment level. Therefore, keeping everything else the same, one would expect that, on average, poor children should receive less investment compared to the level of investment they receive in the original model where poor and rich children have equal returns to investment. As a result, the amount subsidies poor receive is smaller than the amount they receive in the original model. However, as long as Inada condition holds at all ability levels, meaning $\lim_{n \to 0} \partial f(n, \theta)/\partial n = \infty$ for all θ, children of all altruistic parents should receive investment, which implies the main result of the paper still holds.

1.5.3 A Robustness Check: $\underline{\beta} > 0$

I show the main characterization result of the paper under the assumption that selfish parents do not care at all about their children, i.e., $\underline{\beta} = 0$. This subsection shows that this specific assumption is not required for the main result. The crucial assumption is that $\underline{\beta}$ is sufficiently close to zero. More precisely, I show that, in any finite horizon version of the model economy, there is an upper bound for $\underline{\beta}$, call it $\tilde{\underline{\beta}}$, such that, as long as $\underline{\beta}$ is weakly below $\tilde{\underline{\beta}}$, the main characterization result holds.

Proposition 6 *For a version of the model with arbitrary finite length, there exists* $\tilde{\underline{\beta}}$ *such that if* $\underline{\beta} \leq \tilde{\underline{\beta}}$*, then in any Pareto efficient allocation* x^* :

1. *For any t, there exists $\underline{y}_t$ such that if $h^{t-1} = (h^{t-2}, \underline{\beta})$, then $y_t^*(h^{t-1}, \beta) < \underline{y}_t$.*

2. *If $y_t^*(h^t) < \underline{y}_t$, then $\Delta_t^*(h^t) > 0$.*

3. *No agent lives in misery.*

Proof.

Relegated to Appendix 1.8.3. ■

1.6 Ex ante Efficient Allocation and Long-Run Inequality

In this section I consider the following family of social welfare functions parameterized by generational Pareto weight ρ :

$$\sum_{t=1}^{\infty} \sum_{h^t} \mu(h^t) \rho^{t-1} U(h^t|x).$$

In particular, I focus on the case where $\rho = 0$. This is the case in which society cares about only period one parents directly; all other generations are valued solely indirectly through their parents.[17] I refer to the particular Pareto efficient allocation that arises from maximizing this social welfare function as the ex ante efficient allocation. The previous literature on dynamic contracting mostly focused on characterizing ex ante efficient allocation and one arguably robust result stated in this literature is the celebrated *immiseration result.* This result, which was initially shown by Green

[17]Note that the social planning problem with the objective function in (1.6) corresponds to the Pareto problem in which $\bar{v}(h^t) = 0$, for all $t > 1$ and for all $h^t \neq \bar{\beta}$.

(1987) and Thomas and Worrall (1990) independently, states that in dynamic moral hazard environments almost all agents' consumption diverges to the lower bound (assumed to be zero in the current paper).[18] In other words, in the long run, almost all agents live at misery.

In this section, I show that the immiseration result does not hold in the current environment. Phelan (2006) and Farhi and Werning (2007) also show that immiseration result does not hold; they attain the no-immiseration result by allowing the social planner to put welfare weight on future generations directly.[19] The novelty of my approach is that I attain the no-immiseration result without modifying the social planner's objective.[20]

Proposition 7 *In the ex ante efficient allocation,* $\exists\varepsilon\ \exists\eta$ *such that*

$$\lim_{t\to\infty} Pr\{h^t | c_t^*(h^t) \geq \eta\} \geq \varepsilon.$$

Proof.

Relegated to Appendix 1.8.4. ■

The proposition says that in the ex ante efficient allocation, a strictly positive fraction of agents consume an amount strictly above the lower bound on consumption

[18]Both of these papers are partial equilibrium models. Atkeson and Lucas (1992) shows immiseration result holds under general equilibrium as well.

[19]Putting direct welfare weight on future generations can be shown to be equivalent to putting lower bounds on the continuation utility levels of future generations in a social planning problem with no direct weight on future generations in the sense that the two problems would generate same allocations. In that sense, Phelan (2006) and Farhi and Werning (2007) are closely related to Atkeson and Lucas (1995) and Sleet and Yeltekin (2006) who get rid of immiseration by putting lower bounds on continuation utilities.

[20]A recent paper that also attains no-immiseration result without requiring that the objectives of the planner and the private agents disagree is Hosseini, Jones, and Shourideh (2009). It is the addition of the fertility decision into the planning problem that gets rid of immiseration in their environment.

(zero in the current paper) in the long run. So, immiseration of almost all agents does not occur in the long run.

In order to understand why immiseration does not occur in the current environment, it is helpful to understand why immiseration holds in the previous models. In those models, at ny history, in order to insure current consumption and still induce truth telling, the planner chooses to spread out continuation utilities. Therefore, continuation utilities are constantly pushed outward. Due to the concavity of the utility function, it is cheaper to provide incentives in the future when continuation utilities are lower. This is why continuation utilities are pushed to their lower bound for almost all agents. This same force is present in the current environment. Indeed, it is possible to show that in the version of this model where only altruism factor is private information (parental investment is observable), the same force implies that the immiseration result holds. However, when parental investment is also unobservable, then parents have to be provided with incentives to invest in their children, as it is represented by IC_D and IC_{LD}. This implies that there is a novel force implied by productive efficiency to distribute consumption across agents. This force balances the force that gives rise to immiseration and hence we get the no-immiseration result.

1.7 Conclusion

Parental investment is a key factor in determining children's adult outcomes, which implies it is an important productive activity for society. However, society cannot force parents to invest in their children since it is hard to monitor such investments. Furthermore, social and parental objectives are often not aligned. As a result, society faces an agency problem regarding childhood investment. This paper provides an intergenerational framework with such an agency problem and focuses on the entire set of ex post Pareto efficient allocations. The main characterization result is that, independent of social preferences, in all generations, poor parents receive subsidies. This implies that all parents, including the poor, live above misery.

The paper then provides two market structures which achieve Pareto efficiency. In the first one, parents can sign dynastic contracts that bind all their descendants. A version of the First Welfare Theorem holds for this economy: market equilibrium attains a Pareto efficient allocation. However, this market structure is not practical as in most parts of the world contracts signed by parents are not binding for their children. In the second one, I construct a more realistic market structure in which no intergenerational trade is enforced and show that a government can restore efficiency through taxes on income. A property common to all Pareto efficient income tax structures is that agents with low income pay negative taxes.

Empirical literature often interprets the dependence of children's outcomes on parental incomes as a sign of inefficiency.[21] When the only cost of investing in children is the resource cost, put differently, when there are no informational frictions, the current model confirms this interpretation: in all full information Pareto efficient allocations, investment in children is independent of dynastic history and, hence, parental income. Since children's outcomes are determined by parental investment in children, this means if there are no informational frictions, efficiency requires children's outcomes to be independent of parental incomes. On the other hand, when

[21] See Solon (1992), Zimmerman (1992), and Mazumder (2001) among others for empirical papers that establish intergenerational persistence in outcomes such as earnings and income. See Stokey (1996) for a paper that interprets intergenerational persistence as a sign of inefficiency.

there is an agency problem regarding investment in children, there is an additional incentive cost of investment in children, which depends on the dynastic history of the children. In this case, I show that investment in children and, hence, their outcomes depend on parental incomes in Pareto efficient allocations. Consequently, it is not obvious a priori that intergenerational persistence in outcomes observed in data implies data is generated in an inefficient world. One can see whether the agency problem in childhood investment, not the inefficiency of the society in allocating resources, can account for the persistence in outcomes between generations by taking an appropriately modified version of the current model to data. I believe that this is an exciting area for future research.

1.8 Appendix

1.8.1 Incentive Compatibility

Proof of Lemma 1.

$$\begin{aligned} W(\sigma, N|C', N') &= W(\sigma, N|C, N) \\ &\geq W(\hat{\sigma}, \hat{n}|C, N) \geq W(\hat{\sigma}, \hat{n}|C', N'), \end{aligned}$$

where the first and last inequalities are true by construction and the second inequality is true since (C, N) induces truth-telling and obedience. ■

Proof of Proposition 2.

Obviously, incentive constraints (1.2.4) imply temporary incentive constraints (1.2.5). Therefore, I only show the other direction.

Suppose for contradiction that there exists an allocation x satisfying (1.2.5) and violating (1.2.4). This implies there exists a strategy $(\bar{\sigma}, \bar{\phi})$ such that

$$W(\bar{\sigma}, \bar{\phi}|x) - W(\sigma, \phi|x) = \epsilon > 0.$$

Define a new strategy $(\bar{\sigma}^\tau, \bar{\phi}^\tau)$ that is identical to $(\bar{\sigma}, \bar{\phi})$ up to and including period τ and that, for $t > \tau$, prescribes telling the truth and obeying the recommendation at all nodes, i.e., $\bar{\sigma}^\tau_t(h^t) = \beta_t$ and $\bar{\phi}^\tau_t(h^t) = 1$.

Since $\underline{\beta}, \bar{\beta} < 1$, and assuming the appropriate boundary condition, it is obvious that as $\tau \to \infty$, $W(\bar{\sigma}^\tau, \bar{\phi}^\tau|x)$ converges to $W(\bar{\sigma}, \bar{\phi}|x)$. Thus, there exists a τ such that

$$|W(\bar{\sigma}^\tau, \bar{\phi}^\tau|x) - W(\bar{\sigma}, \bar{\phi}|x)| < \epsilon/2.$$

As a result, if one can prove that $W(\bar{\sigma}^{\tau}, \bar{\phi}^{\tau}|x) \leq W(\sigma, \phi|x)$, one gets a contradiction.

Define a new strategy $(\bar{\sigma}^{\tau-1}, \bar{\phi}^{\tau-1})$ that is identical to $(\bar{\sigma}, \bar{\phi})$ up to and including period $\tau - 1$ and that, for $t > \tau - 1$, prescribes telling the truth and obeying the recommendation at all nodes, i.e., $\bar{\sigma}_t^{\tau-1}(h^t) = \beta_t$ and $\bar{\phi}_t^{\tau-1}(h^t) = 1$.

$$
\begin{aligned}
&W(\bar{\sigma}^{\tau-1}, \bar{\phi}^{\tau-1}|x) - W(\bar{\sigma}^{\tau}, \bar{\phi}^{\tau}|x) \\
&= \sum_{h^\tau} \mu(h^t)\Big\{ \Big[u\Big(\Big(\min_{h^t \prec h^\tau} \bar{\phi}_t(h^t) \Big) c_\tau(\bar{\sigma}^{\tau-1}(h^{\tau-1}), \beta_\tau) \Big) \\
&\quad + \beta_\tau V\Big(\Big(\min_{h^t \prec h^\tau} \bar{\phi}_t(h^t) \Big) (\bar{\sigma}^{\tau-1}(h^{\tau-1}), \beta_\tau | x) \Big) \Big] \\
&\quad - \Big[u\Big(\Big(\min_{h^t \prec h^\tau} \bar{\phi}_t(h^t) \Big) c_\tau(\bar{\sigma}^{\tau}(h^{\tau})) + (1 - \bar{\phi}_\tau(h^\tau)) n_\tau(\bar{\sigma}^{\tau}(h^{\tau})) \Big) \\
&\quad + \beta_\tau V\Big(\Big(\min_{h^t \prec h^\tau} \bar{\phi}_t(h^t) \Big) (\bar{\phi}_\tau(h^\tau))(\bar{\sigma}^{\tau}(h^{\tau})|x) \Big) \Big] \Big\},
\end{aligned}
$$

where $V(0)$ is defined as the expected discounted utility from consuming zero in all future nodes, i.e., $V(0) = \frac{1}{1-\mathbb{E}\beta}\kappa$.

For histories h^τ such that there exists a $h^t \prec h^\tau$ with $\bar{\phi}_t(h^t) = 0$, the value inside the brackets is zero, since both expressions inside the brackets is equal to the expected discounted utility of consuming zero in the current period and in all future nodes.

For histories h^τ such that there is no $h^t \prec h^\tau$ with $\bar{\phi}_t(h^t) = 0$, the value of the expression inside the brackets is greater than or equal to zero due to the temporary incentive constraints (1.2.5). Thus,

$$
W(\bar{\sigma}^{\tau-1}, \bar{\phi}^{\tau-1}|x) \geq W(\bar{\sigma}^{\tau}, \bar{\phi}^{\tau}|x).
$$

Define new strategies $(\bar{\sigma}^s, \bar{\phi}^s)$, for $s \in \{0, 1, ..., \tau - 2\}$, that are identical to $(\bar{\sigma}, \bar{\phi})$ up to and including period s and that prescribe telling the truth and obedience for periods $t > s$. By backwards induction, one can show that

$$W(\bar{\sigma}^0, \bar{\phi}^0|x) \geq W(\bar{\sigma}^\tau, \bar{\phi}^\tau|x).$$

Observing that the strategy $(\bar{\sigma}^0, \bar{\phi}^0|x)$ is by definition the truth telling and obedience strategy, (σ, ϕ), gives the desired contradiction. ■

1.8.2 Pareto Efficient Allocations

Proof of Lemma 2.

First, I prove $IC_D(h^{t-1}, \bar{\beta})$ and $IC_{LD}(h^{t-1}, \bar{\beta})$ are redundant.

$$\begin{aligned} & u(c_t(h^{t-1}, \bar{\beta})) + \bar{\beta} V(h^{t-1}, \bar{\beta}|x) \\ \geq\ & u(c_t(h^{t-1}, \underline{\beta})) + \bar{\beta} V(h^{t-1}, \underline{\beta}|x) \\ \geq\ & u(c_t(h^{t-1}, \underline{\beta})) + \underline{\beta} V(h^{t-1}, \underline{\beta}|x) \\ \geq\ & u(c_t(h^{t-1}, \beta) + n_t(h^{t-1}, \beta)), \forall \beta \in H, \end{aligned}$$

where the second inequality follows from $\bar{\beta} > \underline{\beta}$, the first inequality follows from $IC_L(h^{t-1}, \bar{\beta})$, and the third inequality follows from $IC_D(h^{t-1}, \underline{\beta})$ and $IC_{LD}(h^{t-1}, \underline{\beta})$.

Next, I show $IC_L(h^{t-1}, \underline{\beta})$ is redundant. $IC - C(h^{t-1}, \underline{\beta})$ being satisfied implies $u(c_t(h^{t-1}, \underline{\beta})) \geq u(c_t(h^{t-1}, \bar{\beta}) + n_t(h^{t-1}, \bar{\beta}))$, which implies $u(c_t(h^{t-1}, \underline{\beta})) \geq u(c_t(h^{t-1}, \bar{\beta}))$ since investment has to be non-negative. ■

Proof of Lemma 3.

Part 1. This follows directly from $IC_D(h^{t-1}, \underline{\beta})$.

Part 2. Suppose for a contradiction that for some constrained Pareto efficient allocation x^*, $n_t^*(h'^{t-1}, \bar{\beta}) = 0$, for some h'^{t-1}.

First, I show that $n_t^*(h'^{t-1}, \bar{\beta}) = 0$ implies incentive-constraint $IC_{LD}(h'^{t-1}, \underline{\beta})$ binds. Suppose for a contradiction that it does not. Dropping this incentive-constraint and the non-negativity condition on $n_t(h'^{t-1}, \underline{\beta})$ from the planner's problem and taking the first order condition with respect to $n_t(h'^{t-1}, \underline{\beta})$, one gets $f'(n_t^*(h'^{t-1}, \underline{\beta})) = \lambda_t^*/\lambda_{t+1}^* > 0$. By Inada assumption on the production function, this implies $n_t^*(h'^{t-1}, \underline{\beta}) > 0$, which is a contradiction.

Now, I show that $IC_{LD}(h'^{t-1}, \underline{\beta})$ binding implies $c_t^*(h'^{t-1}, \bar{\beta}) > 0$. Suppose to the contrary that $c_t^*(h'^{t-1}, \bar{\beta}) = 0$. Consider the solution to the planner's problem without the incentive constraint $IC_{LD}(h'^{t-1}, \underline{\beta})$. In the solution, it has to be the case that $c_t^*(h'^{t-1}, \underline{\beta}) \geq 0$. Then, $IC_{LD}(h'^{t-1}, \underline{\beta})$ is automatically satisfied, meaning it is indeed slack, which is a contradiction.

Therefore, $c_t^*(h'^{t-1}, \bar{\beta}) > 0$. Now, consider the allocation $\tilde{x}$, which is the same as the constrained Pareto efficient allocation except for:

$$\tilde{c}_t(h'^{t-1}, \bar{\beta}) = c_t^*(h'^{t-1}, \bar{\beta}) - \epsilon,$$

$$\tilde{n}_t(h'^{t-1}, \bar{\beta}) = n_t^*(h'^{t-1}, \bar{\beta}) + \epsilon = \epsilon,$$

and

$$\tilde{c}_{t+1}(h'^{t-1}, \bar{\beta}, \beta) = c_t^*(h'^{t-1}, \bar{\beta}, \beta) + \epsilon',$$

where $\epsilon > 0$ is small and ϵ' is chosen to keep the welfare of parent with dynastic history $(h'^{t-1}, \bar{\beta})$ unchanged:

$$u(c_t^*(h'^{t-1},\bar{\beta})-\epsilon)+\bar{\beta}\sum_{\beta}\mu(\beta)[u(c_{t+1}^*(h'^{t-1},\bar{\beta},\beta)+\epsilon')+\beta V(h'^{t-1},\bar{\beta},\beta|\tilde{x})] = \\ u(c_t^*(h'^{t-1},\bar{\beta}))+\bar{\beta}\sum_{\beta}\mu(\beta)[u(c_{t+1}^*(h'^{t-1},\bar{\beta},\beta))+\beta V(h'^{t-1},\bar{\beta},\beta|x^*)]. \tag{1.8.1}$$

By construction, $V(h'^{t-1}, \bar{\beta}, \beta|\tilde{x}) = V(h'^{t-1}, \bar{\beta}, \beta|x^*)$. Then, for small ϵ, (1.8.1) implies that the resources needed in period $t+1$ to keep the welfare of parent $(h'^{t-1}, \bar{\beta})$ unchanged are:

$$\epsilon' = \epsilon\frac{u'(c_t^*(h'^{t-1},\bar{\beta})-\epsilon)}{\bar{\beta}\sum_{\beta}\mu(\beta)u'(c_{t+1}^*(h'^{t-1},\bar{\beta},\beta)+\epsilon')}. \tag{1.8.2}$$

New resources created in $t+1$ due to human capital investment in the children of parents with dynastic history $(h'^{t-1}, \bar{\beta})$ is equal to $f(\epsilon)$. For ϵ small, this is equal to $\epsilon f'(\epsilon)$. As a result, in period $t+1$, total resource change is:

$$\epsilon\Big[f'(\epsilon)-\frac{u'(c_t^*(h'^{t-1},\bar{\beta})-\epsilon)}{\bar{\beta}\sum_{\beta}\mu(\beta)u'(c_{t+1}^*(h'^{t-1},\bar{\beta},\beta)+\epsilon')}\Big].$$

As ϵ approaches 0, by the Inada condition on f, the first term in the brackets tends to ∞, whereas the second term goes to $\frac{u'(c_t^*(h'^{t-1},\bar{\beta}))}{\bar{\beta}\sum_{\beta}\mu(\beta)u'(c_{t+1}^*(h'^{t-1},\bar{\beta},\beta))}$), which is a finite number. Therefore, for small ϵ, the allocation $\tilde{x}$ raises some extra resources in period $t+1$. By construction, the welfare levels of all parents besides those with dynastic history $(h'^{t-1}, \bar{\beta}, \beta)$ are unchanged, and the welfare of the latter is strictly increased under $\tilde{x}$. Hence, showing that the new allocation $\tilde{x}$ is incentive compatible amounts to showing that it Pareto improves over x^*, which is a contradiction.

Now I show that $\tilde{x}$ is indeed incentive compatible. Since the only change from the original allocation is following the node $(h'^{t-1}, \bar{\beta})$, and by construction this agent's utility is left unchanged, all incentive-constraints preceding node $(h'^{t-1}, \bar{\beta})$ are satisfied in the new allocation. $IC_{LD}(h'^{t-1}, \underline{\beta})$ is satisfied since $c_t^*(h'^{t-1}, \bar{\beta}) + n_t^*(h'^{t-1}, \bar{\beta})$ is unchanged in the new allocation. $IC_L(h'^{t-1}, \bar{\beta})$ is satisfied since the welfare of parent $(h'^{t-1}, \bar{\beta})$ is unchanged. Since the new allocation adds the same amount, ϵ', on top of $c_{t+1}^*(h'^{t-1}, \bar{\beta}, \beta)$ for both β, $IC_L(h'^{t-1}, \bar{\beta}, \bar{\beta})$ and $IC_{LD}(h'^{t-1}, \bar{\beta}, \underline{\beta})$ are also satisfied. There is no reason for the descendants of parents with history $(h'^{t-1}, \bar{\beta}, \beta)$ to deviate since their allocations remain unchanged in the new allocation. Finally, allocations in all other branches are also kept the same, completing the proof that the new allocation is incentive-compatible. ■

1.8.3 $\underline{\beta} > 0$

Proof of Proposition 6.

I need to find (1)$\underline{y}_t$ for every t and (2) $\tilde{\underline{\beta}}$ such that the proposition holds.

(1) Remember that I already proved that $n_t^*(h^{t-1}, \bar{\beta}) > 0$ for all t, h^{t-1}, when $\underline{\beta} = 0$. Similarly, one can show that $n_t^*(h^{t-1}, \bar{\beta}) > 0$ for all t, h^{t-1}, for all $\underline{\beta} \in [0, \bar{\beta}]$. Then, for any t, let $\underline{y}_t = \min_{\underline{\beta} \in [0,\bar{\beta}]} \min_{h^{t-1}} n_t^*(h^{t-1}, \bar{\beta})$. Clearly, $\underline{y}_t > 0$ for all t.

(2) I first prove the following claim:

Claim: $n_t^*(h^{t-1}, \underline{\beta}) \to 0$ for all t, h^{t-1} as $\underline{\beta} \to 0$.

To see this, suppose this is not true for some t, h^{t-1}. That means there exists $\epsilon > 0$ such that no matter how small $\underline{\beta}$ is $n_t^*(h^{t-1}, \underline{\beta}) \geq \epsilon$. Then, consider the incentive constraint $IC_D(h^{t-1}, \underline{\beta})$:

$$u(c_t^*(h^{t-1}, \underline{\beta})) + \underline{\beta} V(h^{t-1}, \underline{\beta}|x^*) \geq u(c_t^*(h^{t-1}, \underline{\beta}) + n_t^*(h^{t-1}, \underline{\beta})) + \underline{\beta} \frac{1}{1 - \mathbb{E}\beta} \kappa.$$

By rearranging this constraint we get:

$$u(c_t^*(h^{t-1},\underline{\beta})) - u(c_t^*(h^{t-1},\underline{\beta}) + n_t^*(h^{t-1},\underline{\beta})) \geq \underline{\beta}\Big[\frac{1}{1-\mathbb{E}\beta}\kappa - V(h^{t-1},\underline{\beta}|x^*)\Big].$$

Since there is a finite amount of resources in this economy at any point in time, L.H.S. of the above expression is a strictly negative number bounded away from zero for any $\underline{\beta} \in [0,\bar{\beta}]$. R.H.S. of the above expression, on the other hand, converges to 0 as $\underline{\beta} \to 0$. But this is a contradiction.

Given the claim it is straightforward that, for every t, there exists $\tilde{\underline{\beta}}_t$ such that for all $\underline{\beta} \leq \tilde{\underline{\beta}}_t$, $f(n_t^*(h^{t-1},\underline{\beta})) < \underline{n}_{t+1}$. I also choose $\tilde{\underline{\beta}}_t$ small enough so that $IC_{LD}(h^{t-1},\underline{\beta})$ implies that $c_t^*(h^{t-1},\underline{\beta}) - [c_t^*(h^{t-1},\bar{\beta}) + n_t^*(h^{t-1},\bar{\beta})] \geq -\delta_t$ where $\delta_t > 0$ is arbitrarily small. Now let $\tilde{\underline{\beta}} = \min_t \tilde{\underline{\beta}}_t$.

Now I am ready to prove the proposition.

Part 1. $y_t^*(h^{t-2},\underline{\beta},\beta) = f(n_{t-1}^*(h^{t-2},\underline{\beta})) < \underline{y}_t$ where the inequality is true by construction.

Part 2. Suppose $y_t^*(h^{t-1},\beta) < \underline{y}_t$. Then, $\Delta_t^*(h^{t-1},\bar{\beta}) = c_t^*(h^{t-1},\bar{\beta}) + n_t^*(h^{t-1},\bar{\beta}) - y_t^*(h^{t-1},\bar{\beta}) > 0$ since $n_t^*(h^{t-1},\bar{\beta}) > \underline{y}_t$ by construction. Then, $\Delta_t^*(h^{t-1},\underline{\beta}) = c_t^*(h^{t-1},\underline{\beta}) + n_t^*(h^{t-1},\underline{\beta}) - y_t^*(h^{t-1},\bar{\beta}) > 0$ since, by construction, $c_t^*(h^{t-1},\underline{\beta}) - [c_t^*(h^{t-1},\bar{\beta}) + n_t^*(h^{t-1},\bar{\beta})] \geq -\delta_t$ where $\delta_t > 0$ is arbitrarily small and $n_t^*(h^{t-1},\bar{\beta}) > \underline{y}_t$.

Part 3. Similar to proof of Part 3 of Proposition 2. ■

1.8.4 Long-Run Inequality

Proof of Proposition 7.

The proof proceeds in two steps.

First, I prove that $\exists\eta\ \exists\varepsilon\ \forall t\ Pr\{h^{t-1}|n_t^*(h^{t-1},\bar{\beta})\geq\eta\}\geq\varepsilon$.

Suppose for a contradiction that this is false. Hence, $\forall\eta\ \forall\varepsilon\ \exists t_\eta : Pr\{h^{t_\eta-1}|n_{t_\eta}^*(h^{t_\eta-1},\bar{\beta})\geq\eta\}<\epsilon$.

From first order necessary conditions, it is easy to show that

$$u'(c_t^*(h^{t-1},\bar{\beta}))\geq\bar{\beta}\sum_{\beta}\mu_\beta u'(c_{t+1}^*(h^{t-1},\bar{\beta},\beta))f'(n_t^*(h^{t-1},\bar{\beta})), \tag{1.8.3}$$

for all h^{t-1},t.

Letting $\eta\to 0$ and keeping ε fixed, $f'(n_{t_\eta}^*(h^{t_\eta-1},\bar{\beta}))\to\infty$. By the assumptions on the utility function, this implies $u'(c_{t_\eta+1}^*(h^{t_\eta-1},\bar{\beta},\beta))\to 0$, for some $\beta\in H$. This implies $c_{t_\eta+1}^*(h^{t_\eta-1},\bar{\beta},\beta)\to\infty$. So, I have shown so far that consumption in period $t_\eta+1$ converges to ∞ for a strictly positive measure of agents as η converges to zero.

For period $t_\eta+1$ feasibility condition to be satisfied, we need total amount of resources in period $t_\eta+1$ to converge to ∞ as well. However, it is easy to show that with a strictly concave production function like f and with no exogenous technological improvements, the total amount of resources in this economy is bounded from above at any point in time. Thus, we achieve the desired contradiction.

In the second step of the proof, I show that the claim just proved implies that at least ε measure of agents consume at least η units in the long run.

$\exists\eta\ \exists\varepsilon\ \forall t\ Pr\{h^{t-1}|n_t^*(h^{t-1},\bar{\beta})\geq\eta\}\geq\varepsilon$ implies that $\lim_{t\to\infty}Pr\{h^{t-1}|n_t^*(h^{t-1},\bar{\beta})\geq\eta\}\geq\epsilon$. Then, by $IC_{LD}(h^{t-1},\underline{\beta})$, we have $\lim_{t\to\infty}Pr\{h^{t-1}|c_t^*(h^{t-1},\bar{\beta})\geq\eta\}\geq\epsilon\frac{\mu(\bar{\beta})}{\mu(\underline{\beta})}$. ■

Chapter 2

Business Start-Ups and Productive Efficiency

2.1 Introduction

Starting a business requires two main ingredients: a productive idea and resources to invest in that idea. Unfortunately, it is not necessarily the case that whoever has one of these ingredients also has the other one. Consequently, there is a potential mismatch among individuals in a society in terms of who holds productive resources and who can use them most efficiently. In a frictionless world, a solution to this mismatch is private markets: those with ideas (potential start-ups) can borrow from those with resources, invest, and then pay back. This paper explores how a society should cope with this mismatch in an environment in which individuals possess asymmetric information at all stages of the process of starting a business. Ex ante, individuals privately know their wealth levels and whether they have ideas or not; interim, their actions are unobservable to others; and ex post, they privately observe investment returns. The paper shows that under these informational frictions, the society has to subsidize poor agents who start businesses so as to cope with this mismatch. Therefore, the paper provides a novel rationale for governments to subsidize business start-ups.

Individuals in the model economy live for two periods and are risk-neutral. In period one, agents are heterogeneous with respect to wealth levels and whether they have ideas or not. Agents with ideas can create businesses that generate risky returns in the next period and feature diminishing marginal returns to capital. In the absence of informational frictions, efficient resource allocation involves two separate steps: (1) productive efficiency requires transferring resources to poor and productive agents initially to make sure that all productive agents can invest at the socially efficient level; (2) distributive efficiency then requires making transfers between agents so as to achieve the desired consumption distribution, which depends on the welfare criterion of the society.

Unfortunately, it is hardly the case that all relevant information about business start-ups are known publicly.[1] The paper assumes that agents' ex ante types (wealth-idea), how they allocate their resources, and ex post returns to business start-ups are private information. The result that poor agents with ideas should be subsidized depends solely on the assumption of unobservability of returns. The assumption that ex ante types and actions of agents are private information only limits the scope of this subsidy. The role of each informational assumption on the results will be discussed

[1]See Hubbard (1998) for a survey of the literature on informational problems in capital markets.

in detail in section 3.

In order to understand the intuition for the subsidy result, one first needs to know what society cares about in this economy. I assume the social welfare function to be utilitarian with equal weights on every agent. This assumption, together with risk neutrality of agents, implies that society has a preference only for the amount of total consumption, not for how it is distributed across agents. The society is only concerned about agents making right amounts of investment. Therefore, the problem that the society is facing is maximizing production subject to incentive compatibility and feasibility.

The intuition for the subsidy result is simple. Since there are diminishing marginal returns to capital, it is socially optimal to have all agents with ideas invest at the same socially efficient level. However, since returns to start-ups are unobservable, agents cannot write contracts with state-contingent repayment schedules. This market incompleteness then implies that agents can, at most, borrow an amount that they can pay back the next period in the lowest return state.[2] This borrowing constraint binds for poor agents with ideas when they want to invest at the efficient level. If the society can transfer some of its resources to these individuals, it would relax their

[2]Observe that I do not allow for default in the model. Following Diamond (1984), one can add default to this model by assuming that if a start-up continues to operate after period two, this brings a continuation value to the owner; if not, then at least some strictly positive fraction of this value gets destroyed. Then, agents can write state-contingent contracts by conditioning the continuation of a start-up business on the level of repayment. The fear of losing a fraction of the continuation value can make the agent make the payment associated with her true return level. In such a world, θ_l state can be interpreted as a default state. Even though in such a model poor agents with ideas would be able to borrow more than they can in the original model, one can show that this level would still be strictly less than the amount they need to finance socially efficient investment level. Therefore, the subsidy result would still be true under this alternative model. The reason why such an extension can be interesting is because it can make the details of the efficient social contract more realistic, giving rise to a more realistic implementation. A paper along these lines is currently work in progress.

budget constraints, enabling them to produce at a level closer to the social optimal, which is the social objective.

Consequently, this paper focuses solely on productive efficiency, leaving aside distributional concerns. The motivation for subsidy in this model is the need to finance the investment of poor agents with ideas. In fact, due to the choice of social welfare function and risk-neutrality assumption, this is the only reason why subsidy is socially desirable.

If the society knew who were the poor agents with ideas, then it would be very easy to implement the subsidy. However, when there are benefits at stake, such as a subsidy, people can pretend to be poor and to have productive ideas, get the subsidy, and consume it. As a result, the amount of subsidy going to poor agents with ideas is constrained by incentive compatibility: agents should not find it optimal to lie about their wealth and ideas, and use the subsidy for reasons other than investment.

Of course, it is possible that the society can try to understand whether people's ideas are productive, and monitor their wealth and how they use the subsidy. However, these activities are all costly. The assumption that it is impossible to pursue such monitoring activities corresponds to assuming that monitoring costs are prohibitively high.[3] I accept that this is an extreme assumption; however, assuming that agents' wealth, ideas, and actions are perfectly, costlessly observable is also extreme. I focus on the less studied of the extremes. I conjecture that the subsidy result would still be true if I allowed for monitoring technologies as long as the cost of monitoring is not zero.

[3]Formally introducing monitoring technologies to the model that allow the user to observe another agent's wealth, idea, or actions and setting the cost of this technology to infinity restores the current model.

It is important to note that the subsidy result is not an artifact of risk neutrality; it survives even if agents have strictly concave utility functions. However, in that case, society would also have a taste for equality that would force a redistribution from the rich to the poor. Furthermore, since agents would be risk-averse, society would like to smooth their consumption across states and periods. The risk-neutrality assumption makes it possible to abstract away from these additional distributive forces and focus solely on what productive efficiency dictates.

A corollary that follows from the subsidy result is that how productive activity (distribution of investment in the current context) *should be* organized in the economy depends on the distribution of wealth. This result depends crucially on the existence of informational frictions. The result and the assumptions behind it are further explained in section 3.

The paper provides a decentralization of the constrained efficient allocation in an incomplete markets setup where people trade risk-free bonds in a competitive market. Given that markets cannot provide subsidies on their own, an incomplete markets equilibrium under laissez-faire cannot attain constrained efficiency.[4] In order to implement the efficient allocation, the paper introduces two separate institutions to the market environment: a government and a government agency that deals with start-up firms. The government taxes all agents in a lump-sum manner and subsidizes its agency from its budget. The agency then subsidizes some individuals from a pool

[4]Note that I do not allow for markets to open ex ante, meaning before agents know whether they are rich or poor and whether they have ideas or not. If that is allowed, then the interpretation of the optimal contract would be completely different. Instead of calling the transfers in the optimal social contract subsidies, we would call them state-contingent payment schedules of the optimal financial contract written between agents behind the veil of ignorance. This is the implementation technique proposed in Prescott and Townsend (1984). Thus, constrained efficiency requires either markets to open ex ante or government to execute subsidies.

of applicants based only on their level of bond holdings. The tax-subsidy system is chosen such that only agents with ideas get subsidized. A comparison of the implementation with the U.S. Small Business Administration's (SBA) Business Loan Program is provided in section 4.

This is not the first paper to put forth the idea that, under informational frictions, productive efficiency may require subsidizing a certain group of individuals in a society. Aghion and Bolton (1997) also takes output maximization as the social objective and shows that when there are moral hazard problems due to unobservable effort, redistributing resources from the rich to the poor may increase total output, boosting economic growth. However, the underlying mechanisms behind the subsidy results in the current paper and in that paper are completely different. Another difference of the current paper from Aghion and Bolton (1997) is that these authors focus on a particular equilibrium notion, whereas the current paper analyzes constrained-efficient allocations.

Loury (1981), Banerjee and Newman (1991), and Galor and Zeira (1993) are also related to the current paper. These papers share a common result: in the presence of capital market imperfections, the distribution of wealth affects the distribution of investment, and hence aggregate output.[5] This is akin to the following result I derive in this paper: the distribution of wealth affects the distribution of productive activity in the constrained-efficient allocation. However, there is an important distinction between the two results. All the papers mentioned above assume some form of market incompleteness and show that the wealth distribution affects equilibrium the distribution of investment under this assumption. The contribution of the current paper

[5]Aghion, Caroli, and Garcia-Penalosa (1999), section 2, not only proves a similar result but also provides a discussion of related papers.

is that, instead of making arbitrary assumptions on the space of contracts available to agents, it takes as given informational frictions and shows that the distribution of wealth affects productive activity in an economy *even* in the constrained-efficient allocation. Consequently, this paper directly establishes that it is due to informational frictions that the distribution of wealth affects the distribution of productive activity.[6]

Another strand of literature that is related to this paper is on optimal venture capital contracts since both this literature and the current paper consider the question of how to finance business start-ups.[7] In general, venture capital literature focuses on characterizing the structure of optimal contracts in principal-agent relationships in which venture capitalists monitor everything but entrepreneurs' effort. The current paper assumes less transparency between agents by assuming that people cannot monitor each other's investment levels or output realizations. This rules out the existence of venture capital in the current model.[8] Consequently, this paper deals with the complementary problem of how a society should allocate productive resources in an environment in which agents are more opaque and hence less capable of allocating resources themselves.

The rest of the paper is organized as follows. Section 2 introduces the baseline model formally and analyzes the full information benchmark. Section 3 defines and

[6]Banerjee and Newman (1991) is closer to the current paper in the sense that it explicitly models an informational friction that causes the market imperfection. However, it restricts the contract space available to the agents arbitrarily. Therefore, essentially, it also focuses on some exogenously specified equilibrium notion, not on constrained efficiency.

[7]See Admati and Peiderer (1994), Gompers (1995), and Jovanovic and Szentes (2007) for important contributions to this literature.

[8]The paper does not claim that venture capital does not exist in real life or it is not important. However, given that it requires some resources that are limited in supply (like time of experts) and, hence, serves a relatively small portion of business start-ups, an alternative less transparent relationship is also present.

solves for the constrained efficient allocation. Section 4 provides an implementation of constrained efficient allocation similar to the U.S. SBA's loan program. Section 5 studies some extensions and generalizations of the model and shows how robust results are. Finally, section 6 concludes.

2.2 Model

2.2.1 Environment

The economy is populated by a continuum of unit measure of agents who live for two periods. Agents are risk-neutral with the instantaneous utility function $u : \Re \rightarrow \Re$ defined as $u(c) = c$, for $c \geq 0$ and $u(c) = -\infty$, for $c < 0$.[9] They are expected utility maximizers with

$$E_1\{u(c_1) + \beta u(c_2)\},$$

where c_t is period t consumption and $\beta \in (0, 1)$ is the discount factor.

At the beginning of period one, some agents are born with ideas and some without. Let i denote whether an agent has an idea or not. Those who have ideas are called $i = 1$ types, and those who do not are called $i = 0$ types. Let $I = \{0, 1\}$. The fraction of agents born with (without) an idea is η_1 (η_0). In order to produce, an agent must have an idea. Agents are also born with different levels of initial endowment of the only consumption good, $w \in W = \{p, r\}$, $p < r$. Fraction ζ_w are born with initial

[9] Allowing for negative consumption but setting utility derived from it to negative infinity is a convenient way of securing non-negativity of consumption in the solution. The reason for requiring non-negativity of consumption is as follows. The only reason why there is any exchange between agents in this model is to finance investment. If utility function were $u(c) = c$ for all c, then any agent can finance her own investment by consuming a negative amount. Then, even under autarky full information efficiency would be achieved.

wealth level w. There is no endowment in period two. So, there are four types of agents initially, at the beginning of period one: $\{(p,0),(p,1),(r,0),(r,1)\}$.

Agent of type (w,i) operates the following production technology:

$$y = i\theta k^{\alpha}, \alpha \in (0,1),$$ [10]

where k is the amount invested in period one, θ is the random return on capital, and y is the random output produced in period two. θ is drawn from the set $\Theta = \{\theta_l, \theta_h\}$, where $\theta_l < \theta_h$, according to the probability distribution μ, independently across agents.[11] The probability of drawing θ_l is μ_l and θ_h is μ_h. An agent gets to learn the realization of return after the investment is made. Hence, agents face idiosyncratic investment risk. The term i is in the production function to denote that only agents with ideas can start businesses.

There is also a risk-free, linear storage technology that is available to all agents. An agent who stores s_1 units in period one wakes up with As_1 units in period two. The assumption below says that the storage technology is wasteful.

Assumption 3 $A < 1/\beta$.

The information structure and timing of events are as follows: An agent's initial type, actions, and period two realized returns are private information. The rest of the data of the economy is public information. Given her initial type, an agent chooses how much to consume, invest, and store in period one. Then, in period two, θ is

[10] This specific form of the production function is not needed for any of the results; all the results go through if $y = \theta f(k)$, where $f(0) = 0$ and $f', -f'' > 0$.

[11] The assumption that the cardinality of the set of returns is two is immaterial for any result. All the results go through if Θ has any finite number or a continuum of elements.

realized and hence output is produced, and the agent consumes.[12]

One way to think about resource allocation is to consider a benevolent social planner who chooses allocations for agents. Since consumption-investment choice is unobservable, the planner cannot choose allocations directly. Instead, each period the planner makes transfers between agents based on their reports of their private histories. This way the planner manipulates agents' actions. In addition, there is no outside party, which means the planner cannot save or borrow resources through time.[13]

An allocation in this economy is a vector $(c, k, s, \delta) \equiv (c_1, c_2, k_1, s_1, \delta_1, \delta_2)$, where

$$\begin{aligned} c_1 &: W \times I \rightarrow \Re \\ k_1 &: W \times I \rightarrow \Re_+ \\ s_1 &: W \times I \rightarrow \Re_+ \\ c_2 &: W \times I \times \Theta \rightarrow \Re \\ \delta_1 &: W \times I \rightarrow \Re \\ \delta_2 &: W \times I \times \Theta \rightarrow \Re. \end{aligned}$$

In the above, $c_1(w, i)$, $k_1(w, i)$, and $s_1(w, i)$ refer to period one levels of consumption, investment, and storage of the agent who has initial wealth w and idea i. Similarly, $c_2(w, i, \theta)$ is the consumption level of the agent of type (w, i) who has a realized return θ in period two.[14] $\delta_1(w, i)$ and $\delta_2(w, i, \theta)$ are the levels of transfers received by corresponding types.

[12]Whether θ is realized in period one or two is immaterial for the results; the important thing is that it is realized after the investment decision is made.

[13]All results go through if the planner can borrow and save at a risk-free rate of $1/\beta$.

[14]Since an agent with no idea cannot produce, her period two consumption is independent of θ. So $c_2(w, 0, \theta_l) = c_2(w, 0, \theta_h)$.

Feasibility. An allocation (c, k, s, δ) is *feasible* if

$$\begin{gathered}\sum_{w,i} \zeta_w \eta_i \delta_1(w,i) \leq 0, \\ \sum_{w,i} \sum_\theta \zeta_w \eta_i \mu_\theta \delta_2(w,i,\theta) \leq 0,\end{gathered} \tag{2.2.1}$$

and for every $(w, i) \in W \times I$

$$\begin{gathered}c_1(w,i) + k_1(w,i) + s_1(w,i) \leq w + \delta_1(w,i), \\ c_2(w,i,\theta) \leq i\theta k_1(w,i)^\alpha + As_1(w,i) + \delta_2(w,i,\theta),\end{gathered} \tag{2.2.2}$$

$$k_1(w,i), s_1(w,i) \geq 0.^{15} \tag{2.2.3}$$

Here, (2.2.1) is the *aggregate feasibility* condition, which says that the planner should balance its budget every period. (2.2.2) is *individual feasibility* and stands for the fact that allocation assigned to each agent should be affordable by him. (2.2.3) is just the non-negativity constraint on investment and storage.

Incentive compatibility. Using the terminology of mechanism design literature, there are two sources of private information in the model. First, there is hidden information: an agent's initial type and period two investment returns are observed privately by the agent. Second, agents are involved in hidden action: their consumption and investment levels are hidden. Hence, they can deviate from an allocation recommended by the planner in two ways: they can lie about their private information and/or they can choose an investment level that is different from what the planner recommended. Due to these informational frictions, only incentive-compatible alloca-

[15] As it can be understood from (2.2.2), we assume that there is full depreciation of capital. This assumption is only made for notational simplicity and is not needed for any of the results.

tions are achievable. I invoke a powerful revelation principle introduced by Myerson (1982) and characterize the set of incentive-compatible allocations as follows.

Let $(\tilde{w}, \tilde{i}) \in W \times I$ and $\tilde{\theta} : \Theta \to \Theta$ be agent's period one and period two reporting strategies, respectively. Also, define $(\tilde{k}_1, \tilde{s}_1) \in \Re^2_+$ as agent's investment strategy. Then, $\tilde{\gamma} \equiv (\tilde{w}, \tilde{i}, \tilde{\theta}, \tilde{k}_1, \tilde{s}_1)$ is a complete strategy of agent (w, i).[16] Let Γ be the set of all complete strategy profiles. Given the allocation (c, k, s, δ), for any (w, i), the utility of following a strategy $\tilde{\gamma}$ is:

$$V_{w,i}(\tilde{\gamma}; c, k, s, \delta) \equiv u[w + \delta_1(\tilde{w}, \tilde{i}) - \tilde{k}_1 - \tilde{s}_1] + \beta \sum_{\theta} \mu_\theta u[i\theta \tilde{k}_1^{\alpha} + A\tilde{s}_1 + \delta_2(\tilde{w}, \tilde{i}, \tilde{\theta}(\theta))]$$

Define $\gamma \equiv (w, i, \theta, k_1, s_1)$ to be the strategy consisting of truthful reporting and obeying recommendations, where $\theta(\theta) = \theta$ denotes the truth-telling reporting strategy.

An allocation (c, k, s, δ) is *incentive-compatible* if for each $(w, i) \in W \times I$,

$$V_{w,i}(\gamma; c, k, s, \delta) \geq V_{w,i}(\tilde{\gamma}; c, k, s, \delta), \text{for all } \tilde{\gamma} \in \Gamma. \tag{2.2.4}$$

An allocation that is feasible and incentive-compatible is called *incentive-feasible.*

2.2.2 Benchmark: Full Information Efficiency

The aim of this subsection is to analyze what society can achieve when everything in the economy is publicly observable. Full information efficient allocation turns out to be a useful benchmark for the constrained-efficient allocation. Under the utilitarian

[16] Myerson (1982) calls this *participation strategy.* Also, note that consumption is not a part of the strategy since it is implied by the choice of other actions.

objective, the efficient allocation with full information is the solution to the following problem:

Planner's full information problem.

$$\max_{c,k} \sum_{w,i} \zeta_w \eta_i \Big\{ u(c_1(w,i)) + \beta \sum_{\theta} \mu_\theta u(c_2(w,i,\theta)) \Big\}$$

s.t.

$$\sum_{w,i} \zeta_w \eta_i \Big\{ c_1(w,i) + k_1(w,i) + s_1(w,i) \Big\} \leq \sum_{w} \zeta_w w,$$

$$\sum_{w,i} \zeta_w \eta_i \sum_{\theta} \mu_\theta c_2(w,i,\theta) \leq \sum_{w,i} \zeta_w \eta_i \sum_{\theta} \mu_\theta \Big\{ i\theta k_1(w,i)^\alpha + A s_1(w,i) \Big\},$$

$$k_1(w,i), s_1(w,i) \geq 0, \text{ for all } (w,i) \in W \times I.$$

Since s_1 is wasteful, it is obvious that in the full information efficient allocation $s_1(w,i) = 0$, for all $(w,i) \in W \times I$.

Assuming that total initial wealth in period one is large enough, the first-order optimality condition for investment of agents with an idea reads:

$$1/\beta = \alpha k_1(w,1)^{\alpha-1} \sum_{\theta} \mu_\theta \theta.$$

The left-hand side of the equation is the marginal social cost of investing an additional unit in terms of period two utility. The right-hand side is the marginal social benefit of investment in the same units. This condition defines

$$k^{fi} = \left\{ \beta\alpha \sum_i \mu_i \theta_i \right\}^{\frac{1}{1-\alpha}}$$

as the full information efficient level of investment provided that the following holds:

Assumption 4 *Total resources in the economy in period one are sufficient to finance k^{fi} investment for each $(w, 1)$ agent, or*

$$\eta_1 k^{fi} \leq \sum_w \zeta_w w.$$

Assumption 4 formally states that cumulative initial wealth is sufficiently large.[17]

Lemma 4 *Suppose Assumption 4 holds. Then,*

1. *The full information level of investment for agents with ideas is equal to k^{fi}, irrespective of their wealth. The full information level of investment for agents without ideas is zero.*

2. *The full information level of storage is zero for all agents.*

3. *As long as it provides non-negative consumption to all agents and uses all output, distribution of individual consumption does not matter.*

[17]If Assumption 4 does not hold, then the full information level of investment will be a corner solution: $\frac{\sum_w \zeta_w w}{\eta_1}$, and all the results of the paper go unchanged.

Looking at the objective function of the full information problem, one can see that utilitarian welfare with equal weights and risk neutrality together imply that society has no preference for how total consumption should be distributed, as long as no one gets negative consumption. The society is only concerned about the right agents making the right amounts of investment. Therefore, there is a set of full information efficient allocations that are unique up to the distribution of consumption.

The next section analyzes a problem with exactly the same objective function, but this time with a different constraint set due to private information. As a result, that problem will be one of maximizing production subject to feasibility and incentive compatibility. Thanks to Lemma 4, it is clear now that the challenge that awaits the society under private information is to make agents with ideas invest as close to the full information efficient level as possible.[18]

2.3 Constrained-Efficient Allocation

In analyzing the benchmark case, the only assumption made was about total initial wealth. However, with private information, the comparison of p, r and k^{fi} becomes important. The first assumption about this comparison is the following:

Assumption 5 $p < k^{fi} < r$.

The first part of this assumption, $p < k^{fi}$, says that the initial wealth of the poor is not large enough to cover the full information level of investment. Thus, a poor agent with an idea cannot operate her idea at the most efficient level on her own. If, to the

[18] Of course, without making any agent consume a negative amount.

contrary, $p \geq k^{fi}$ were the case, the economy would reach full information without agents interacting at all. Obviously, this case is neither interesting nor realistic. The second part of the assumption, that $k^{fi} < r$, simply says that a rich agent who has an idea can invest at the efficient level even under autarky. Therefore, Assumption 5 ensures that there is a reason for the planner to intervene in this economy: to ensure that poor agents with ideas invest at the efficient level.

The remainder of this section first defines and then characterizes constrained-efficient allocation.[19]

Definition 6 *An allocation* $(c^*, k^*, s^*, \delta^*)$ *is called constrained-efficient if it solves the following social planner's problem:*

$$\max_{c,k,\delta} \sum_{w,i} \zeta_w \eta_i \Big\{ u(c_1(w,i)) + \beta \sum_{\theta} \mu_\theta u(c_2(w,i,\theta)) \Big\}$$

subject to (2.2.1), (2.2.2), (2.2.3), *and* (2.2.4).

As in the benchmark case, the objective function clearly shows that society does not care about how consumption is going to be distributed among individuals. Consequently, the above problem is one of constrained productive efficiency. This implies there can be many constrained efficient allocations, all of which have the same investment allocation and hence the same total production and welfare, but different consumption allocations. Nonetheless, it should also be noted that incentive compatibility arising from private information does put some discipline on the distribution of consumption across agents compared to the full information benchmark.

[19] Throughout the paper, I refer to efficient allocation under the informational problems as constrained-efficient allocation so as to distinguish it from the full information efficient allocation.

2.3.1 Characterizing the Constrained-Efficient Allocation

First make the following observation, which simplifies the analysis. If transfer levels depend on period two announcements of agents, then any agent will report the type that brings the highest level of transfers in period two. Therefore, any transfer mechanism in which a transfer level depends on a period two shock cannot be incentive-compatible. Consequently, without loss of generality, the rest of the paper restricts attention to allocations in which transfers are functions of period one announcements only, $\delta_1, \delta_2 : W \times I \rightarrow \Re$.

Now I make the second assumption comparing p and k^{fi}.

Assumption 6 $\frac{k^{fi}-p}{\beta} > \theta_l {k^{fi}}^{\alpha}$.

To understand this assumption, suppose it does not hold. Observe that in order to invest at the full information efficient level, the poor agent with an idea needs at least $k^{fi} - p$ additional resources in period one. Also observe that the most this agent can pay back in period two in low-return state is $\theta_l {k^{fi}}^{\alpha}$. When Assumption 6 does not hold, even in the worst contingency in period two, the poor agent with an idea would be able to pay back the amount she borrowed in period one to finance the full information level of investment. Obviously, in this case, the fact that entrepreneurial returns are private information would not have an effect. Therefore, no subsidy would be necessary for constrained efficiency. The society can implement the full information outcome by just making sure that simple, not state-contingent debt contracts are perfectly enforced (by punishing very harshly anyone who does not pay back the amount she borrowed in period one). Agents, then, sign these contracts that an set interest rate of $1/\beta$ and optimal investment would be attained. However, that even in the worst case an entrepreneur can pay back her debt is highly unrealistic, especially for businesses that are newly forming.[20]

The proposition below formally shows that when Assumption 6 does not hold, the full information allocation is trivially reached without any net present value (NPV) of transfers between agents.

Before getting to the proposition, define $\Delta(w,i) = \delta_1(w,i) + \beta\delta_2(w,i)$ as the NPV of transfers an agent gets under a given allocation. An agent (w,i) is said to be *subsidized* by the society under allocation (c,k,s,δ) if $\Delta(w,i) > 0$.

Proposition 8 *Suppose that* $\frac{k^{fi}-p}{\beta} \leq \theta_l {k^{fi}}^{\alpha}$. *Then, in the constrained-efficient allocation:*[21]

1. $k_1^*(w,1) = k^{fi}$ *and* $k_1(w,0) = 0$, *for all* $w \in W$;

2. $s_1^*(w,i) = 0$, *for all* $(w,i) \in W \times I$;

3. $\delta_1^*(p,1) = k^{fi} - p$ *and* $\delta_2^*(p,1) = \frac{-(k^{fi}-p)}{\beta}$;

4. $(\delta_1^*(w,i), \delta_2^*(w,i))_{(w,i)\neq(p,1)}$ *satisfy* $\Delta^*(w,i) = 0$, *aggregate feasibility, and individual feasibility with non-negative consumption for all.*

In words, no agent gets subsidized in the constrained-efficient allocation.

Proof.

Showing that the allocation described in Proposition 8 is in the constraint set of the

[20] That the lowest return is sufficiently dire is a standard assumption in financial contracting literature. Among others, see Diamond (1984), Gale and Hellwig (1985), Bolton and Scharfstein (1990), and DeMarzo and Fishman (2007).

[21] In fact, there is a set of constrained-efficient allocations that are unique up to the distribution of consumption. Since the current paper is not concerned with consumption distribution, I refer to this whole set as "the constrained-efficient allocation."

planner's problem is sufficient since under this allocation total output is equal to the full information total output level.

Choose $\delta_1^*(w,i)$ for $(w,i) \neq (p,1)$ such that:

$$\sum_{(w,i)\neq(p,1)} \zeta_w \eta_i \delta_1^*(w,i) = -(k^{fi} - p)\zeta_p \eta_1, \tag{2.3.1}$$

$$\delta_1^*(r,1) \geq k^{fi} - r, \tag{2.3.2}$$

$$0 \geq \delta_1^*(w,0) \geq -w, \tag{2.3.3}$$

and

$$\delta_2^*(w,i) = -\frac{\delta_1^*(w,i)}{\beta}. \tag{2.3.4}$$

By Assumption 4, such a δ^* exists. Observe that conditions (2.3.1) and (2.3.4) guarantee that transfers sum to zero in periods one and two, respectively. Thus, aggregate feasibility is satisfied.

Next, one has to show that non-negative consumption is feasible for each agent under the proposed allocation. Observe that the NPV of transfers of any agent is equal to zero in this allocation. In period one, a poor agent with an idea faces the budget $c_1 + k_1 + s_1 \leq k^{fi}$ and chooses $k_1^*(p,1) = k^{fi}$. In period two in the low-return state, her consumption is $c_2^*(p,1,\theta_l) = \theta_l {k^{fi}}^{\alpha} - \frac{k^{fi}-p}{\beta} \geq 0$ by assumption. This clearly implies that $c_2(p,1,\theta_h) \geq 0$, too. Condition (2.3.2) guarantees that $(r,1)$ agent can choose investment equal to k^{fi} and still consume a non-negative amount in period one. The consumption levels of agents without ideas are non-negative in both periods by (2.3.3).

The only thing left is to check that given δ^* agents will tell the truth about their types, but this is straightforward given that the NPV of transfers of any type is equal to zero. ■

From now on, the paper analyzes the more interesting case in which Assumption 6 holds: the lowest return to an idea, θ_l, is sufficiently low.

Remember that, under the ex ante welfare criterion, the only reason why the planner intervenes in this economy ($\delta \neq 0$) is because a poor agent with an idea invests at a very low level, p, on her own. In order to make her invest at the full information level, the planner has to set $\delta_1(p,1) \geq k^{fi} - p$. Since returns to business start-ups, θ, are private information, period two transfers cannot depend on the returns. Therefore, an agent who is poor and has an idea can pay back to the society an amount that is at most equal to the output she produces in the low-return state, $\delta_2(p,1) \geq -\theta_l {k^{fi}}^\alpha$. This implies that in order to attain full information efficiency, the NPV of transfers going to the poor agent with an idea should at least be $\Delta^{fi} \equiv k^{fi} - p - \beta\theta_l {k^{fi}}^\alpha$, which is strictly positive by Assumption 6.

In what follows, without loss of generality, I restrict attention to constrained-efficient allocations in which $\Delta^*(p,1) \leq \Delta^{fi}$. This is without loss of generality for the following reason. The discussion in the above paragraph shows that if $\Delta^*(p,1) = \Delta^{fi}$, then $k_1^*(p,1) = k^{fi}$, meaning full information efficiency is attained. Thus, in any allocation in which NPV of transfers going to $(p,1)$ is higher than Δ^{fi}, the value of the objective function in the social planner's problem under informational frictions is equal to the full information level. Thus, increasing the NPV of transfers going to $(p,1)$ above Δ^{fi} does not change social objective but only changes the distribution of consumption across agents.

Lemma 5 below implies that if the society wants to increase the investment level for poor agents with ideas, it has to increase the NPV of transfers going to these agents.

Lemma 5 *In the constrained-efficient allocation,* $\delta_1^*(p,1) = k_1^*(p,1) - p$ *and* $\delta_2^*(p,1) = -\theta_l k_1^*(p,1)^\alpha$.

Proof.

Observe that $\Delta^*(p,1) \leq \Delta^{fi}$ implies that $\delta_1^*(p,1) \leq k^{fi} - p$. If not, then $c_2^*(p,1,\theta_l) < 0$, a contradiction. This implies automatically that $c_1^*(p,1) = s_1^*(p,1) = 0$ and hence $\delta_1^*(p,1) = k_1^*(p,1) - p$. First, suppose for contradiction that $c_1^*(p,1) > 0$. Then, the agent can decrease her consumption by a small amount and increase her investment by the same amount. This would increase her welfare strictly since her investment level is strictly below k^{fi} when $c_1^*(p,1) > 0$, a contradiction. By the same logic, $s_1^*(p,1) = 0$ as well.

Now, suppose that $\delta_2^*(p,1) > -\theta_l k_1^*(p,1)^\alpha$. Define a new allocation with transfers $\tilde{\delta}_2(p,1) = \delta_2^*(p,1) - \epsilon$ and $\tilde{\delta}_1(p,1) = \delta_1^*(p,1) + \beta\epsilon$. The resulting allocation is incentive-feasible since the NPV of transfers is unchanged for all agents. For agents $(w,i) \neq (p,1)$, welfare is unchanged. The change in $(p,1)$'s welfare will be

$$\underbrace{\beta\epsilon\alpha k_1(p,1)^{\alpha-1}\sum_\theta \mu_\theta\theta}_{\text{Gain in pd. 2}} - \underbrace{\epsilon}_{\text{Loss in pd. 2}} > 0,$$

which means the new allocation improves over the constrained-efficient one, a contradiction. ■

Lemma 5 implies the following has to hold in the constrained-efficient allocation:

$$\Delta^*(p,1) = k_1^*(p,1) - p - \beta\theta_l k_1^*(p,1)^\alpha.$$

Taking the derivative of both sides with respect to $k_1^*(p,1)$ gives

$$\frac{d\Delta^*(p,1)}{dk_1^*(p,1)} = 1 - \beta\theta_l k_1^*(p,1)^{\alpha-1}\alpha.$$

Now observe that for $\Delta^*(p,1) \geq 0$, $1-\beta\theta_l k_1^*(p,1)^{\alpha-1} \geq \frac{p}{k_1^*(p,1)} > 0$. Therefore, $\Delta^*(p,1)$ is strictly increasing in $k_1^*(p,1)$, for $\Delta^*(p,1) \geq 0$. This implies $\Delta^*(p,1)$ is a one-to-one function of $k_1^*(p,1)$, as long as $\Delta^*(p,1) \geq 0$. As a result, the converse is also true: in order to increase $k_1^*(p,1)$, the planner needs to increase $\Delta^*(p,1)$. Therefore, Lemma 5 implies that society has to increase the NPV of transfers going to poor agents with ideas so as to bring these agents' investment levels closer to the full information level and thus so as to bring social welfare closer to the full information level.

Proposition 9 below is the main result of this section and shows that poor agents with ideas are subsidized in the constrained-efficient allocation.[22] Since rich agents with ideas can pretend to be poor, they have to get transfers with the same NPV. As a result, all agents with ideas, potential start-ups, receive the same subsidy in the constrained-efficient allocation. Proposition 9 makes it clear that for the subsidy result to hold, the storage technology has to be wasteful. It also proves that there is a

[22] Under the restriction that $\Delta^*(p,1) \leq \Delta^{fi}$, the constrained-efficient level of consumption, investment, saving, and transfers assigned to poor agents with ideas is unique. However, this is not true for other types agents. As a result, there is a set of constrained-efficient allocations all of which have the same allocation for poor agents with ideas. So, the term "the constrained-efficient allocation" that I use in the text actually refers to this set.

threshold level of the return to storage technology, $\bar{A}$, such that it is incentive-feasible to make transfers to $(p,1)$ with NPV equal to Δ^{fi} if and only if the return to storage is less than or equal to $\bar{A}$. Furthermore, Proposition 9 provides an exact calculation of the constrained-efficient allocation in either case.[23]

I have made Assumption 7 below. This assumption is not substantial in the sense that it is not necessary for the subsidy result. It is assumed merely for expositional purposes.[24]

Assumption 7 ***a.*** $\eta_1[k^{fi} - p - \beta\theta_l {k^{fi}}^{\alpha}] \leq \eta_0 p$.

b. $\eta_1[k^{fi} - p] \leq \eta_0 \sum_w \zeta_w w$.

Proposition 9 *Suppose Assumptions* 6 *and* 7 *hold. Then there exists a unique* $\bar{A} \equiv \frac{\eta_0\theta_l {k^{fi}}^{\alpha}}{k^{fi} - p - \eta_1\beta\theta_l {k^{fi}}^{\alpha}} \in (0, \beta^{-1})$ *such that in the constrained-efficient allocation:*

1. $k_1^*(r,1) = k^{fi}$ *and* $k_1^*(w,0) = 0$, *for all* $w \in W$;

2. $k_1^*(p,1) = k^{fi}$, *if* $A \leq \bar{A}$,

   $k_1^*(p,1) < k^{fi}$ *is given by the unique solution to* $A = \frac{\eta_0\theta_l k_1^*(p,1)^{\alpha}}{k_1^*(p,1) - p - \eta_1\beta\theta_l k_1^*(p,1)^{\alpha}}$, *if* $A > \bar{A}$;

3. $s_1^*(w,i) = 0$ *for all* $(w,i) \in W \times I$;

[23]Proposition 9 lays out a particular subset of the set of constrained-efficient allocations in which transfer levels are independent of wealth for agents with ideas. This is done only for the sake of expositional purposes.

[24]There are two parts to Assumption 7 and both parts are about relative fractions of agents with and without ideas. Assumptions 7*a* says that the fraction of agents without ideas is large enough relative to the fraction of agents with ideas that it is feasible for agents without ideas to finance the total amount of subsidies going to agents with ideas. Assumption 7*b* simply allows me to make the transfers of agents with ideas independent of their wealth types.

4. *$\delta_1^*(w,1) = k_1^*(p,1) - p$ and $\delta_2^*(w,1) = -\theta_l k_1^*(p,1)^\alpha$ for all $w \in W$;*

5. *$(\delta_1^*(w,0), \delta_2^*(w,0))_{(w\in W)}$ satisfy: $\Delta^*(w,0) = -\frac{\eta_1}{\eta_0}\Delta^*(w,1)$ for all $w \in W$, and individual feasibility with non-negative consumption for all.*

More importantly, $\Delta^*(w,1) > 0$, i.e., in the constrained-efficient allocation society transfers a strictly positive NPV of resources from agents without ideas to the ones with ideas.

Proof.

Lemma 5 already showed that the higher is the NPV of transfers that the poor agents with ideas receive the higher their level of investment and hence the closer the society is to full information efficiency. As a result, the planner's problem is equivalent to maximizing the NPV of transfers going to poor agents with ideas subject to non-negativity of consumption and incentive-feasibility.

The proof proceeds as follows. First, I show that storage technology is not used by anyone in the constrained-efficient allocation. Second, I show that the NPV of transfers going to agents with ideas is independent of wealth. Then, I claim, without actually proving, that the same is true for agents without ideas. These imply that the only incentive-compatibility condition one needs to check is the one regarding the deviation in which the agents without ideas pretend to have ideas. Analyzing that condition gives us the result.

First, suppose for a contradiction that $s_1^*(w^o, i^o) > 0$, for some $(w^o, i^o) \in W \times I$. Now define a new allocation $(\tilde{c}, \tilde{k}, \tilde{s}, \tilde{\delta})$ that is identical to the constrained-efficient allocation, except for $\tilde{s}_1(w^o, i^o) = s_1^*(w^o, i^o) - \epsilon$, $\tilde{k}_1(p,1) = k_1^*(p,1) + \epsilon\frac{\zeta_{w^o}\eta_{i^o}}{\zeta_p\eta_1}$, and

$\tilde{c}_2(w,i,\theta) = c_2^*(w,i,\theta) + \epsilon\frac{\zeta_{wo}\eta_{io}}{\zeta_p\eta_1}[1/\beta - A]$. The new allocation satisfies aggregate feasibility in period one by construction. Observe that $k_1^*(p,1) \leq k^{fi}$ in the constrained-efficient allocation. Thus, marginal returns to investing additional ϵ units in poor agents with ideas is at least $1/\beta$ units, for ϵ small. Thus, period two aggregate feasibility condition is also satisfied. Since period one consumption levels are unchanged from the original allocation for all agents and period two consumption levels increased by the same amount, if the original allocation is incentive-compatible, the new allocation has to be incentive-compatible as well. As a result, the new allocation is in the constraint set of the planner and provides all agents with higher welfare which means the original allocation cannot be constrained-efficient, a contradiction.

Now, I show that $\Delta^*(p,1) = \Delta^*(r,1)$. If $\Delta(p,1)^* > \Delta(r,1)^*$, then $(r,1)$ lies to be $(p,1)$ and gets the transfers with higher NPV; therefore, this cannot be true. On the other hand, if $\Delta(p,1)^* < \Delta(r,1)^*$, then one can propose a new allocation with a transfer system $\tilde{\delta}$ that is the same as δ^*, except for $\tilde{\delta}_1(p,1) = \delta_1^*(p,1) + \epsilon$ and $\tilde{\delta}_1(r,1) = \delta_1^*(r,1) - \frac{\zeta_p}{\zeta_r}\epsilon$. Clearly, this transfer mechanism is a part of a feasible allocation. This new allocation is also incentive-compatible: $(r,1)$ does not lie to be $(p,1)$ since $\epsilon > 0$ is small and agents without ideas do not lie to be $(p,1)$ since with original transfers they were not lying to be $(r,1)$ and the NPV of transfers of $(p,1)$ is still lower than that of $(r,1)$ under the original transfer mechanism. But the allocation that is attained by this transfer mechanism has strictly greater aggregate utility. The reason is $(p,1)$ agent's utility increases strictly more than ϵ with the new allocation since she was investment-constrained under δ^*. Then, Assumption 7b allows the planner to set $\delta_1^*(r,1) = k_1^*(p,1) - p$ and $\delta_2^*(r,1) = -\theta_l k_1^*(p,1)^\alpha$, without loss of generality.

One can similarly show that in the constrained-efficient allocation $\Delta^*(p,0) = \Delta^*(r,0)$.

Aggregate feasibility then implies

$$\Delta^*(w,0) = -\frac{\eta_1}{\eta_0}\Delta^*(p,1). \tag{2.3.5}$$

So far it has been shown that in order to increase ex ante welfare, the planner has to increase $\Delta^*(w,1)$, and this has to be financed by taking resources from agents without ideas.

The only incentive-compatibility condition that one needs to check is the one regarding deviations in which an agent without an idea lies to have an idea. The allocation has to satisfy for all w, w'

$$w+\Delta^*(w,0) \geq \begin{cases} w + \Delta^*(w',1) + \frac{\delta_1^*(w',1)-\Delta^*(w',1)}{\beta A}[-1+\beta A], & \text{if } A(\delta_1^*(w',1)+w) \geq -\delta_2^*(w',1); \\ -\infty, & \text{if else.} \end{cases} \tag{2.3.6}$$

Here, the left-hand side of the equation is the utility of truth-telling, whereas right-hand side is the utility of lying to be $(w,1)$. The left-hand side already takes into account the fact that transfers are such that when agents without ideas tell the truth, they do not have to use risk-free technology. Hence, they do not use it. The right-hand side already has the fact that $(w,0)$ has to set $s_1' \geq \frac{-\delta_2^*(w',1)}{A} = \frac{\delta_1^*(w',1)-\Delta^*(w,1)}{\beta A}$ to have non-negative consumption in period two. When $A(\delta_1^*(w',1)+w) \geq -\delta_2^*(w',1)$ this is possible. Otherwise, lying to be $(w,1)$ implies they have to consume a negative amount in either period one or period two, meaning they get negative infinity utility.

Plugging in $\delta_1^*(w,1) = k_1^*(p,1) - p$ and $\delta_2^*(w,1) = -\theta_l k_1^*(p,1)^\alpha$, (2.3.6) implies that $k_1^*(p,1)$ units of investment for the poor agent with an idea is attained in the constrained-efficient allocation if and only if it satisfies, for any $w \in W$:

$$A < \frac{\theta_l k_1^*(p,1)^\alpha}{k_1^*(p,1) - p + w}, \text{ or } A \leq \frac{\theta_l k_1^*(p,1)^\alpha}{k_1^*(p,1) - p - \Delta^*(w,0)}.$$

It follows from Assumption 7a and aggregate feasibility that $\frac{\theta_l k_1^*(p,1)^\alpha}{k_1^*(p,1)-p+w} < \frac{\theta_l k_1^*(p,1)^\alpha}{k_1^*(p,1)-p-\Delta^*(w,0)}$. Hence, $k_1^*(p,1)$ is attained at the constrained-efficient allocation if and only if $A \leq \frac{\theta_l k_1^*(p,1)^\alpha}{k_1^*(p,1)-p-\Delta^*(w,0)}$. Using equation (2.3.5), it follows that $k_1^*(p,1)$ is constrained-efficient if and only if

$$A \leq \frac{\eta_0 \theta_l k_1^*(p,1)^\alpha}{k_1^*(p,1) - p - \eta_1 \beta \theta_l k_1^*(p,1)^\alpha}. \tag{2.3.7}$$

Hence, for $A \leq \bar{A}$, as it is defined in Proposition 9, $k_1^*(p,1) = k^{fi}$ is incentive-compatible.

To see that for $A > \bar{A}$, $k_1^*(p,1)$ is given by $A = \frac{\eta_0 \theta_l k_1^*(p,1)^\alpha}{k_1^*(p,1)-p-\eta_1\beta\theta_l k_1^*(p,1)^\alpha}$, first observe that the expression on the right-hand side of equation (2.3.7) is strictly decreasing in k_1. Now suppose for a contradiction that $A < \frac{\eta_0 \theta_l k_1^*(p,1)^\alpha}{k_1^*(p,1)-p-\eta_1\beta\theta_l k_1^*(p,1)^\alpha}$. Then, define a new transfer system $\tilde{\delta}$ which is identical to the constrained-efficient one, δ^*, except for $\tilde{\delta}_1(w,1) = \delta_1^*(w,1) + \epsilon$ and $\tilde{\delta}_1(w,0) = \delta_1^*(w,0) - \frac{\eta_1}{\eta_0}\epsilon$, $\epsilon > 0$. By Lemma 5, this means the investment level for the poor agents with ideas is $\tilde{k_1} = k_1^*(p,1) + \epsilon$. This decreases the right-hand side of equation (2.3.7). However, for ϵ small, equation (2.3.7) still holds under the new allocation. Thus, this new allocation is incentive-compatible. It is clearly feasible. Finally, it strictly increases total welfare since it increases $(p,1)$ agents' investment increases. Then, δ^* cannot be constrained-efficient, a contradiction. ■

The intuition for why the NPV of transfers going to poor agents with ideas is related to the returns to storage is simple. Someone has to finance the subsidy going to agents with ideas. Consequently, individuals without ideas end up getting strictly negative NPV of resources. But is it incentive-compatible to transfer resources from agents without ideas to ones with them? Or do agents with no ideas lie to have an idea and get the subsidy? The answer depends on the returns to the storage technology, A. The reason is that period two transfers of agents with ideas is strictly negative, and hence if agents without ideas want to pretend to have ideas, they have to pay back to the society in period two. For agents without ideas, the only way to carry resources into period two is via the storage technology.

If $A = 0$, then it is impossible for agents without ideas to carry resources to period two. In that case, they cannot pretend to have ideas; therefore, planner can transfer Δ^{fi} to agents with ideas and attain full information efficiency. As A increases, storing resources instead of consuming in period one becomes less wasteful. There is a threshold level of the return to storage technology, $\bar{A}$, such that above this level, the benefit of lying to have an idea (not financing but enjoying the subsidy) exceeds the cost of doing so. As a result, when $A > \bar{A}$, agents with ideas cannot be subsidized Δ^{fi}, and hence poor agents with ideas cannot invest at the full information level, k^{fi}.

Nonetheless, as long as $A < \beta^{-1}$, some subsidy is still incentive-compatible since $A < \beta^{-1}$ implies that it is costly to store resources and hence lie for agents without ideas to have ideas. In this case, the amount of subsidy going to agents with ideas is determined by equating the benefit and cost to the agents without ideas of lying to have ideas.

Proposition 9 shows that under some parameters, the society attains the full information efficient allocation, even under the informational constraints. This result is an artifact of risk neutrality and hence will vanish if more general utility functions are assumed. On the other hand, the main result of Proposition 9, that due to informational problems productive efficiency requires transferring resources from agents without ideas to ones with them, holds with risk-averse preferences as well.

Proposition 9 also points to an interesting property of the model economy: the distribution of wealth affects the constrained-efficient distribution of productive activity in the economy. To see this, remember that Proposition 9 tells that when $A > \bar{A}$, the constrained-efficient level of investment for a poor agent with an idea depends on her wealth level, p. Now, consider another wealth distribution with ζ_p fraction of agents having initial wealth $p + \epsilon$ and ζ_r fraction having $r - \epsilon\frac{\zeta_p}{\zeta_r}$, where $\epsilon > 0$ and small. This new wealth distribution is a perturbation of the old one in a way that preserves the mean. By Proposition 9, in the economy with the perturbed wealth distribution, $k_1^*(r,1) = k^{fi}$ and investment level for a poor agent with an idea is given by $A = \frac{\eta_0\theta_l k_1^*(p,1)^\alpha}{k_1^*(p,1)-(p+\epsilon)-\eta_1\beta\theta_l k_1^*(p,1)^\alpha}$.[25] This means that in the current model, when Assumptions 6 and 7 hold and $A > \bar{A}$, the constrained-efficient distribution of productive activity depends on how initial wealth is distributed across agents. This result is summarized in the following corollary.

Corollary 2 *Suppose Assumptions* 6 *and* 7 *hold, and* $A > \bar{A}$. *Distribution of productive activity in the constrained-efficient allocation depends on the distribution of wealth.*

[25]Here, I abuse the notation, hoping this does not cause any confusion. In the original economy, p denotes two things: poor agents and their wealth level. In the perturbed economy, p denotes poor agents, whereas $p + \epsilon$ denotes their wealth level. The same is true for r.

It is important to note that this result crucially depends on private information assumptions. In the full information efficient allocation, investment levels for both agents with ideas is k^{fi}, independent of how a total of $\sum_w \zeta_w w$ units is distributed across agents. The intuition for why the constrained-efficient level of productive activity depends on the distribution of wealth is as follows. Under private information, the marginal social cost of investment is not only equal to its resource cost. For a given distribution of initial wealth, increasing the investment level of poor agents with ideas tightens some incentive-compatibility conditions. Thus, there is an incentive cost of increasing investment in addition to the resource cost. Changing the wealth distribution changes this incentive cost of investment while leaving the resource cost untouched. Consequently, between two otherwise identical economies with different distributions of wealth, the resource cost of investment is the same, which implies that the full information efficient allocation is the same. However, the incentive costs in these two economies are potentially different, making the social marginal costs of investment different, which results in different distributions of constrained-efficient productive activity.

2.3.2 Discussion of Assumptions

This section discusses the role of informational assumptions on the subsidy result. The assumption that θ, the returns to a start-up, is unobservable is the sole cause of the subsidy result. To see this, consider a version of the model in which, for each agent, θ is realized publicly. Assume that initial type, (w, i), and actions are still private information. In that case, the planner can attain full information efficiency without subsidizing any agent, even under Assumption 6. This result is shown in

Proposition 10 below.

Proposition 10 *Suppose that θ is observable for each agent. Then, in the constrained-efficient allocation:*

1. $k_1^*(w,1) = k^{fi}$ *and* $k_1(w,0) = 0$, *for all* $w \in W$;

2. $s_1^*(w,i) = 0$, *for all* $(w,i) \in W \times I$;

3. $\delta_1^*(p,1) = k^{fi} - p$ *and* $\delta_2(p,1,\theta_l) = -\theta_l {k_1^{fi}}^{\alpha}$, *and* $\delta_2(p,1,\theta_h)$ *such that* $\Delta^*(p,1) = 0$;

4. $(\delta_1^*(w,i), \delta_2^*(w,i))_{(w,i)\neq(p,1)}$ *satisfy* $\Delta^*(w,i) = 0$, *aggregate feasibility, and individual feasibility with non-negative consumption for all.*

In words, no agent gets subsidized in the constrained-efficient allocation.

Proof.

Since the allocation described attains productive efficiency, we only need to check that it is incentive-compatible and satisfies aggregate and individual feasibility conditions with no agent consuming a negative amount. Incentive compatibility directly follows from the fact that the NPV of transfers is zero for each agent. Aggregate and individual feasibility is by construction. That each agent consumes a non-negative amount in any period is obvious except for $(p,1)$ agent in θ_h state. So, we need to show that $\delta_2(p,1,\theta_h) \leq \theta_h {k^{fi}}^{\alpha}$.

For a contradiction, suppose that $\delta_2(p,1,\theta_h) > \theta_h {k^{fi}}^{\alpha}$. By construction, $\delta_2(p,1,\theta_h)$ satisfies

$$k^{fi} - p - \beta[\mu_l \theta_l {k^{fi}}^{\alpha} + \mu_h \delta_2(p,1,\theta_h)] = 0.$$

Therefore, we get:

$$k^{fi} - p > \beta \sum_{\theta} \mu_\theta \theta {k^{fi}}^{\alpha}. \tag{2.3.8}$$

Remember that the first-order condition that gives k^{fi} is:

$$1 = \beta \alpha {k^{fi}}^{\alpha - 1} \sum_{\theta} \mu_\theta \theta.$$

Multiplying both sides by k^{fi} and then subtracting p from both sides gives

$$k^{fi} - p = \beta \alpha {k^{fi}}^{\alpha} \sum_{\theta} \mu_\theta \theta - p,$$

which contradicts with (2.3.8). ■

The intuition is simple. When θ is observable, the planner can make period two transfers depend on the realization of θ. Therefore, even if the low state return, θ_l, is very low (Assumption 6), the agent can still pay back to the society the future value of resources transferred to her in period one, $k^{fi} - p$, by paying a sufficiently high amount in the high state. This proposition precisely establishes that the only reason in the model why the society has to subsidize agents with ideas is because start-up returns are private information.

The assumptions that initial type and actions are observable imply that the planner has to respect incentive-compatibility conditions when subsidizing poor agents with ideas. Consider, for instance, a model that is identical to the baseline model introduced in section 2, except that initial type, (w, i), is publicly known at no cost. As long as θ is unobservable, the society still has to make Δ^{fi} units of transfers to poor agents with ideas. However, now it is trivial to make this transfer since the planner knows exactly the agents who have ideas but lack resources to invest in them.

Similarly, if investment is assumed to be observable, keeping the rest of the model the same as the baseline model, subsidizing agents with ideas would be trivial. It is not beneficial for an agent without an idea to lie to have one and get the subsidy since she has to invest it, and hence cannot consume it. The exercise in which everything else is kept the same but storage is assumed to be observable is the same as assuming there is no storage technology, or $A = 0$. From Proposition 9, it follows that in this case, the planner can make transfers with NPV that is sufficient to attain full information efficiency.

2.4 Implementation

This purpose of this section is to provide an implementation of the constrained-efficient allocation via a program like the U.S. SBA's Business Loan Program. The SBA is the major government institution in the United States assisting business start-ups in particular and small businesses in general. The total amount of outstanding small business loans that was subsidized by the SBA's loan program was $75.5 billion as of the end of fiscal year 2007. This amounts to 30% of all small business borrowing.[26] One can consider the paper's implementation as providing a justification for the subsidies that the SBA's Business Loan Program hands out to start-up firms.

I first show that laissez-faire markets cannot carry out the required subsidy and hence cannot implement the constrained-efficient allocation. Then, I introduce the paper's implementation, and finally I compare it to the SBA's Business Loan Program.

[26] This and further information can be found in the SBA report *Performance and Financial Highlights Fiscal Year 2007*.

The physical and informational environment is the same as described in section 2. The main difference is that there is an incomplete markets structure that allows agents to competitively trade risk-free bonds in period one. Bonds pay back a gross return R in period two that is determined in equilibrium. Individual trades in the bonds market are public information and there is full enforcement, meaning that no one can die without paying back their debt.

There are two institutions: a government and an institution of the government that aids start-up businesses. The government taxes all individuals in the society lump-sum, by an amount T, and transfers these funds to its institution. Any individual can apply to this institution for a subsidy. The institution asks the agent to report her wealth, business idea, and investment plan, w', i', and k_1', respectively. Then, after observing the amount borrowed (or lent) and the reports, the institution decides whether or not to provide the subsidy, $\tau(b_1, w', i', k_1')$.

Taking the tax-subsidy system (T, τ) and the interest rate R as given, an agent (w, i) who decided to apply to the institution for a subsidy solves the following problem:

Agent's problem.

$$\max_{c_1, c_2, k_1, s_1, b_1, w', i', k_1'} u(c_1) + \beta \sum_{\theta} \mu_\theta u(c_{2\theta}) \tag{2.4.1}$$

s.t.

$$c_1 + k_1 + s_1 + b_1 \le w - T(b_1) + \tau(b_1, w', i', k_1'),$$
$$c_{2\theta} \le i\theta k_1^\alpha + Rb_1,$$
$$k_1, s_1 \ge 0.$$

An agent who does not apply for a subsidy ($a = 0$ agent) would solve a very similar problem. The only difference is there would be no $\tau(b_1, w', i', k_1')$ in that agent's problem, and hence there would not be any w' and k_1' choice. However, since in the current setup there is no cost of applying for a subsidy, without loss of generality, assume that all agents apply.

Below is the definition of incomplete markets equilibrium with a tax-subsidy system.

Definition 7 *Given (T, τ), an incomplete markets (IM) equilibrium is individual choices*

$(c_1(w,i), c_2(w,i,\theta), k_1(w,i), s_1(w,i), b_1(w,i), w'(w,i), i'(w,i), k_1'(w,i))_{w \in W, i \in I, \theta \in \Theta}$ *and interest rate R s.t.*

1. *Given R, for each agent (w,i),*

   $(c_1(w,i), c_2(w,i,\theta), k_1(w,i), s_1(w,i), b_1(w,i), w'(w,i), i'(w,i), k_1'(w,i))$ *solves* (2.4.1),

2. *Bond market clears:* $\sum_w \sum_i \zeta_w \eta_i b_1(w,i) = 0$,

3. *The institution's budget balances:* $\sum_w \sum_i \zeta_w \eta_i \tau(b_1(w,i), w'(w,i), i'(w,i), k_1'(w,i)) = T$.

An allocation (c, k, s, δ) is *implementable in the market with a tax-subsidy system* (T, τ) if, given (T, τ), (c, k, s, b, w', i', k') with some interest rate R constitute an IM equilibrium.

2.4.1 Incomplete Markets under Laissez-Faire

Before providing an actual tax-subsidy system that implements constrained-efficient allocation, this subsection first analyzes what happens under no government intervention, i.e., $(T, \tau) = 0$.

Proposition 11 *Suppose Assumptions* 6 *and* 7 *hold. Then, constrained-efficient allocation cannot be attained in the equilibrium of IM under laissez-faire.*

Proof.

Suppose for contradiction that the constrained-efficient allocation can be achieved. Then, $b_1(p, 1) \leq -(k^{fi} - p)$. Due to linearity of preferences, $R = 1/\beta$. Thus, $c_2(p, 1, \theta_l) \leq \theta_l {k^{fi}}^{\alpha} - \frac{k^{fi} - p}{\beta} < 0$, by Assumption 5. But this cannot be an optimal choice for the agent since the agent could do better just by setting $b_1(p, 1) = 0$. Thus, we have a contradiction. ■

Proposition 9 already proved that the constrained-efficient allocation involves transferring strictly positive NPV of resources from agents without ideas to those with ideas. Proposition 11 then follows since markets cannot make such transfers on their own.[27] A separate entity, like a government, should intervene and make the necessary transfers between agents.[28]

[27]That I restrict attention to incomplete markets from the outset is without loss of generality. It is easy to show that under given informational assumptions and the assumption that agents cannot write contracts ex ante (before (w, i) is realized), agents cannot reach an allocation with higher total output than the incomplete markets equilibrium.

[28]It is important to note that laissez-faire IM equilibrium is ex post Pareto efficient. However, it is not output maximizing (ex ante Pareto efficient), which is the focus of this paper.

2.4.2 Optimal Tax-Subsidy System

This subsection provides an actual tax-subsidy system that implements the constrained-efficient allocation.

In that regard, define

$$T = \tfrac{\eta_1}{\eta_0}\Delta^*(w,1),$$

$$\tau(b_1, w', i', k_1') = \begin{cases} \tfrac{1}{\eta_0}\Delta^*(w,1), & \text{if } b_1 \leq -\beta\theta_l k_1^*(p,1)^\alpha; \\ 0, & \text{otherwise.} \end{cases} \tag{2.4.2}$$

Proposition 12 *Suppose Assumptions* 6 *and* 7 *hold. Then, the incomplete markets equilibrium with the tax-subsidy system defined in* (2.4.2) *implements the constrained-efficient allocation.*[29]

Proof.

Now I construct an IM equilibrium where $R = 1/\beta$ and agents with ideas invest at their corresponding constrained-efficient investment level.

Under the specified taxes, an agent faces the following problem:

Agent's problem with taxes.

$$\max_{c,k,b,s} u(c_1) + \beta \sum_{\theta} \mu_\theta u(c_{2\theta})$$

[29] Observe that the subsidy function does not depend on any of the agent's reports. This is due to the fact that the current model abstracts away from any form of monitoring. Hence, these reports are just cheap talk.

s.t.

$$c_1 + k_1 + b_1 + s_1 \leq \begin{cases} w + \Delta^*(w,1), & \text{if } b_1 \leq -\beta\theta_l k_1^*(p,1)^\alpha; \\ w - \frac{\eta_1}{\eta_0}\Delta^*(w,1), & \text{if else,} \end{cases}$$

$$c_{2\theta} \leq i\theta k_1^\alpha + Rb_1 + As_1,$$
$$s_1, k_1 \geq 0.$$

First, consider an agent who has an idea in period one. If a poor agent with an idea chooses $b_1 \leq -\beta\theta_l k_1^*(p,1)^\alpha$, she chooses $k_1 = k_1^*(p,1)$ and $b_1 = p - k_1^*(p,1) + \Delta^*(w,1) = -\beta\theta_l k_1^*(p,1)^\alpha$. Suppose for contradiction that this is not true. Then, there is (k_1', b_1', s_1'), where $(k_1', s_1') \neq (k_1^*(p,1), 0)$, which gives strictly greater utility to the agent. $s_1' = 0$ follows immediately from the fact that the return to bonds is strictly greater than the risk-free return; hence, it has to be that $k_1' \neq k_1^*(p,1)$. Then, define a new allocation with transfers $\delta_1'(w,1) = k_1' - p$, $\delta_2'(w,1) = -\frac{k_1'-p-\Delta^*(w,1)}{\beta}$, and $(\delta_1'(w,0), \delta_2'(w,0))_{w\in W}$ such that $\Delta'(w,0) = -\eta_1/\eta_0\Delta'(w,1)$ and individual consumption levels are non-negative. $(p,1)$ chooses (k_1', b_1') in the market, implying that she chooses k_1' in the planner's problem when she faces δ'. The only thing left to check is incentive compatibility. That holds because of the way in which the new allocation is constructed, $\Delta'(w,1) = \Delta^*(w,1)$. Therefore, this new allocation is incentive-feasible, keeps the welfare of $(w,i) \neq (p,1)$ unchanged compared to the constrained-efficient allocation, and provides strictly greater welfare than the constrained-efficient level for poor agents with ideas. This means that the new allocation is an improvement over the constrained-efficient allocation, a contradiction.

Similarly, one can show that when $b_1(r,1) \leq -\beta\theta_l k_1^*(p,1)^\alpha$, $(r,1)$ agent chooses to invest at the constrained-efficient level, k^{fi}.

Now we need to show that agents with ideas choose $b_1 \leq -\beta\theta_l k_1^*(p,1)^\alpha$. The utility of $(w,1)$ type when she chooses $b_1 \leq -\beta\theta_l k_1^*(p,1)^\alpha$, is:

$$w + \Delta^*(w,1) - k_1^*(w,1) - b_1(w,1) + \beta \sum_\theta \mu_\theta[\theta k_1^*(w,1)^\alpha + b_1(w,1)/\beta]$$
$$= w + \Delta^*(w,1) - k_1^*(w,1) + \beta \sum_\theta \mu_\theta \theta k_1^*(w,1)^\alpha. \quad (2.4.3)$$

On the other hand, if an agent with an idea chooses $b_1 > -\beta\theta_l k_1^*(p,1)^\alpha$, then, letting her optimal choices be $\tilde{k}_1(w,1), \tilde{b}_1(w,1)$, her utility would be:

$$w - \frac{\eta_1}{\eta_0}\Delta^*(w,1) - \tilde{k}_1(w,1) - \tilde{b}_1(w,1) + \beta \sum_\theta \mu_\theta[\theta \tilde{k}_1(w,1)^\alpha + \tilde{b}_1(w,1)/\beta]$$
$$= w - \frac{\eta_1}{\eta_0}\Delta^*(w,1) - \tilde{k}_1(w,1) + \beta \sum_\theta \mu_\theta \theta \tilde{k}_1(w,1)^\alpha. \quad (2.4.4)$$

The difference between the maximized values of $(w,1)$ agents' problem in the market under $b_1 \leq -\beta\theta_l k_1^*(p,1)^\alpha$ and under $b_1 > -\beta\theta_l k_1^*(p,1)^\alpha$ then is given by subtracting (2.4.4) from (2.4.3) :

$$\frac{\Delta^*(w,1)}{\eta_0} + [-k_1^*(w,1) + \beta \sum_\theta \mu_\theta \theta k_1^*(w,1)^\alpha] - [-\tilde{k}_1(w,1) + \beta \sum_\theta \mu_\theta \theta \tilde{k}_1(w,1)^\alpha]. \quad (2.4.5)$$

Given that k^{fi} maximizes the function $-k + \beta \sum_\theta \mu_\theta \theta k^\alpha$, it is obvious that this difference is strictly positive for $(r,1)$. For $(p,1)$, under $b_1 > -\beta\theta_l k_1^*(p,1)^\alpha$, $\tilde{k}_1(w,1) < k_1^*(p,1) \leq k^{fi}$. This, combined with the fact that the function $-k + \beta \sum_\theta \mu_\theta \theta k^\alpha$ is strictly increasing in k, for $k \leq k^{fi}$, implies that the expression in (2.4.5) is also strictly positive. Hence, we showed that agents with ideas act according to the constrained-efficient allocation in the market.

Now consider agents who do not have an idea in period one. If they choose $b_1 \leq -\beta\theta_l k_1^*(p,1)^\alpha$, then $c_2(w,0) \leq -\theta_l k_1^*(p,1)^\alpha + As_1(w,0)$. To keep consumption non-negative, $s_1(w,0) \geq \frac{\theta_l k_1^*(p,1)^\alpha}{A}$. Since $A < \beta^{-1}$, these agents will invest as little as possible in risk-free technology. This implies they choose $b_1 = -\beta\theta_l k_1^*(p,1)^\alpha$ and $s_1(w,0) = \frac{\theta_l k_1^*(p,1)^\alpha}{A}$. The utility then is $w + \Delta^*(w,1) + \beta\theta_l k_1^*(p,1)^\alpha - \frac{\theta_l k_1^*(p,1)^\alpha}{A}$.

When an agent with no ideas chooses $b_1 > -\beta\theta_l k_1^*(p,1)^\alpha$, she sets $s_1 = 0$ and chooses the constrained-efficient allocation. The utility she gets is $w - \frac{\eta_1}{\eta_0}\Delta^*(w,1)$.

We need to show that, for agents without ideas, utility under $b_1 > -\beta\theta_l k_1^*(p,1)^\alpha$ is greater than utility under $b_1 \leq -\beta\theta_l k_1^*(p,1)^\alpha$. The difference is equal to

$$\begin{aligned} w - \frac{\eta_1}{\eta_0}\Delta^*(w,1) - [w + \Delta^*(w,1) + \beta\theta_l k_1^*(p,1)^\alpha - \frac{\theta_l k_1^*(p,1)^\alpha}{A}] \\ = \frac{-\Delta^*(w,1)}{\eta_0} - \theta_l k_1^*(p,1)^\alpha(\beta - 1/A) \\ = 0, \end{aligned}$$

where the last inequality follows from $A = \frac{\eta_0\theta_l k_1^*(p,1)^\alpha}{k_1^*(p,1) - p - \eta_1\beta\theta_l k_1^*(p,1)^\alpha}$.

Market clearing and government budget balance conditions are immediate from the fact that the constrained-efficient allocation satisfies aggregate feasibility and has non-negative consumption for all agents. ■

The way the implementation works is as follows. An agent who borrows above the threshold gets a net subsidy of $-\frac{\eta_1}{\eta_0}\Delta^*(p,1) + \Delta^*(p,1)/\eta_0 = \Delta^*(p,1)$. Remember that this is exactly the amount of the NPV of transfers agents with ideas get in the planner's problem. Therefore, agents with ideas borrow at the threshold level, get the subsidy, and invest at the constrained-efficient level. Agents without ideas would like to do the same; however, for them, the only way to pay back in period two is to save through the storage technology, s_1, which is costly since s_1 is wasteful. The threshold amount of borrowing required to get the subsidy is chosen such that this cost is weakly higher than the benefit of getting the subsidy. Therefore, only agents with ideas get the subsidy, and hence the budget of the agency balances.

2.4.3 Comparing the Model's Implementation to the SBA's Business Loan Program

A comparison between the implementation provided above and the actual Business Loan Program is in order. The above implementation is similar to the actual system in the United States in the sense that in both, the government taxes all citizens and transfers some of its tax revenue to the agency that deals with start-ups, with the intention of subsidizing potential start-ups that are financially constrained.[30] Another similarity is that, in the model, the government agency uses borrowing and lending activities of agents as a device for screening agents with ideas and subsidizes only those agents who borrow above a threshold. The loan program of the SBA follows a similar strategy: the SBA subsidizes only those who get loans from commercial banks.

However, since the model is very simple, there are significant differences between the paper's implementation and the actual system in the United States. Here, I stress two of those discrepancies and what causes them.

First, there is no default in the model economy; agents only sign non-state-

[30] The fact that in the paper's implementation all the tax revenue goes to the agency dealing with start-ups is immaterial. One can add exogenous government to the model, G, and then only $T - G$ units would be transferred to the agency. As long as this G is also subtracted from the right-hand side of the aggregate feasibility condition in the planner's problem, all the analysis goes unchanged.

contingent that they have to honor by assumption. This creates a discrepancy between the model and the actual program because the actual loan program does not give out direct subsidies but rather provides loan guarantees to qualified borrowers. These guarantees ensure the lenders that in case of default the SBA will pay back a certain percentage of the loan. This, in turn, causes the interest rate on SBA backed loan to be lower relative to other loans, thereby effectively subsidizing borrowers.

Second, in real life it is possible to monitor some features of start-ups at some cost, while the model abstracts away from any sort of monitoring. As a result, the actual Business Loan Program takes people's reports about, say, their ideas more seriously and spends some resources (labor) to determine whether or not the ideas are worth subsidizing. In the model, once an agent sends a report to the agency saying she has an idea, there is no way to check whether she is lying or not. Therefore, in the model's implementation, the function τ, which determines who gets subsidized and potentially depends on agents' reports, does not actually depend on agents' reports.

2.5 Generalizations

The purpose of this section is to convince the reader that the results of the paper are in fact quite general. I try to do this by changing the model in various dimensions and explaining how the results still hold in the resulting new environments. For each alteration of the model, I keep everything else the same so as to focus on the specific feature being altered. Also, for simplicity I am going to set $A = 0$, so there is no storage technology.[31]

[31] Formal analysis of the first extension is omitted for the sake of brevity; however, it is available from the author upon request.

2.5.1 All Agents Have Ideas

Instead of having some agents have ideas and some not, suppose that all agents have ideas, but fraction η_g have good ideas ($i = g$) and fraction η_b have bad ideas ($i = b$).

An agent of type (w, i) operates the following production technology:

$$y = \theta k^{\alpha}, \alpha \in (0, 1),$$

where, as before, k is the amount invested in period one, $\theta \in \{\theta_l, \theta_h\}$ is the idiosyncratic random return, and y is the output. The probability that an agent (w, i) receives return θ is $\mu_i(\theta)$, where

$$\mu_g(\theta_h) > \mu_b(\theta_h).$$

So, both agents with good and bad ideas produce, but the former is more likely to get a high return.

Define $k^{i,fi}$ as the full information efficient level of investment for agents with idea type i. Obviously, $k^{g,fi} > k^{b,fi}$. An analog of Assumption 6 here is that $\frac{k^{g,fi}-p}{\beta} > \theta_l k^{g,fi\alpha}$. Also, suppose for simplicity that $k^{b,fi} < p$. So, even the poor agents with bad ideas can invest at the full information level on their own. It is possible to show that θ being unobservable, together with this version of Assumption 6, implies that constrained efficiency requires that agents with good ideas receive a subsidy. That agents with bad ideas can pretend to have good ideas limits the amount of the subsidy going to agents who have good ideas. Nevertheless, this does not completely eliminate the subsidy since it is costly for agents with bad ideas to lie to have good ideas. The same implementation also works in this environment with some simple adjustments to the tax-subsidy system.

2.5.2 Allowing for Negative Consumption

Suppose utility function is of the form

$$u(c) = \begin{cases} c, & \text{if } c \geq 0; \\ \lambda c, & c < 0, \end{cases} \tag{2.5.1}$$

where $\lambda > 1$ is a constant.

The purpose of this extension is to show that our main subsidy result does not crucially hinge upon non-negativity restriction on consumption. However, we need $\lambda > 1$, meaning marginal disutility of decreasing consumption when $c \leq 0$ is strictly greater than marginal disutility of decreasing consumption when $c > 0$.[32] As long as this holds, the constrained-efficient allocation involves NPV of transfers from unproductive to productive agents.

Full information efficiency is the same as it is in the benchmark case: both agents with ideas invest at the socially efficient level, k^{fi}, and consumption distribution is such that aggregate feasibility holds with equality and no agent consumes a negative amount in any period and any state.

Now consider constrained efficiency. Since the only change in the physical en-

[32] The reason why $\lambda > 1$ is crucial is as follows. The only reason why there is any exchange between agents in this model is to finance investment. If $\lambda = 1$, then any agent can finance her own investment by consuming a negative amount, and hence even under autarky full information efficiency would be achieved.

vironment compared to the baseline model is the utility function, the definition of constrained efficiency remains the same as in section 3. Exact calculation of the constrained-efficient allocation for any $\lambda > 1$ is very lengthy and tedious due to numerous cases and hence is omitted.[33] Instead, the paper first provides a proposition that shows if λ is greater than or equal to a threshold level $\underline{\lambda}$, then the amount of NPV of transfers that poor agents with ideas have to receive in order to attain the full information solution is incentive-compatible. Hence, the constrained-efficient allocation coincides with a full information efficient allocation.[34] Then, the paper goes on to prove that any constrained-efficient allocation features subsidizing agents who have ideas, as long as $\lambda > 1$.

Proposition 13 *Suppose Assumptions* 6 *and* 7 *hold. Then, there exists a unique* $\underline{\lambda} = 1 + \frac{k^{fi} - p - \beta\theta_l {k^{fi}}^{\alpha}}{\eta_0 \beta \theta_l {k^{fi}}^{\alpha}} > 1$ *such that if* $\lambda > \underline{\lambda}$, *then in the constrained-efficient allocation:*

1. $k_1^*(w, 1) = k^{fi}$ *and* $k_1^*(w, 0) = 0$, *for all* $w \in W$;

2. $\delta_1^*(w, 1) = k^{fi} - p$ *and* $\delta_2^*(w, 1) = -\theta_l {k^{fi}}^{\alpha}$, *for all* $w \in W$;

3. $(\delta_1^*(w, 0), \delta_2^*(w, 0))_{(w \in W)}$ *satisfy:* $\Delta^*(w, 0) = -\frac{\eta_1}{\eta_0}\Delta^*(w, 1)$ *and individual feasibility with non-negative consumption for all.*

More importantly, in the constrained-efficient allocation society transfers strictly positive NPV of resources from agents without ideas to those with ideas.

[33] This calculation is available from the author upon request.

[34] Remember that the full information efficient allocation is indeterminate in terms of individual consumption as long as no one consumes a negative amount.

Proof.

The allocation described in Proposition 13 attains full information efficiency, provided that it is in the constraint set of the planner. Therefore, all one needs to do is to prove that this allocation is incentive-feasible. Individual and aggregate feasibility conditions and non-negativity of consumption for all agents at all times and states hold by construction.

Given that $\Delta^*(w,1)$ is independent of w, an agent with an idea does not lie to be a different agent with an idea. The same is true for agents without ideas. Agents with ideas do not lie to be agents without ideas, since that brings transfers with strictly smaller NPV. Therefore, one only needs to check that agents without ideas do not lie to have ideas under the proposed allocation.

Given that no one gets to consume a negative amount in the constrained-efficient allocation, the payoff to those without ideas from telling the truth is $w + \delta^*(w,0)$. The value from lying to be $(w',1)$ is $w + \delta_1^*(w',1) + \beta\lambda\delta_2^*(w',1)$. Thus, this allocation is incentive-compatible if

$$\begin{gathered} -\tfrac{\eta_1}{\eta_0}\Delta^*(w,1) \geq \Delta^*(w,1) + \beta(\lambda-1)\delta_2^*(w',1) \\ \Rightarrow \lambda \geq 1 - \tfrac{\Delta^*(w,1)}{\eta_0\beta\delta_2^*(w,1)}. \end{gathered}$$

Rearranging (2.5.2) and plugging in $\delta_1^*(w,1)$ and $\Delta^*(w,1)$, we get that the proposed allocation is incentive-compatible if $\lambda \geq 1 + \frac{k^{fi} - p - \beta\theta_l {k^{fi}}^\alpha}{\eta_0\beta\theta_l {k^{fi}}^\alpha}$, which is the assumption made in the proposition. ■

Consequently, if λ is large enough, the amount of subsidy needed to achieve the full information investment level, Δ^{fi}, can be reached even in the case with private information. The intuition is as follows. From the perspective of agents without ideas, the benefit of lying to be an agent with an idea is getting transfers with NPV equal to $(1+\frac{\eta_1}{\eta_0})\Delta^{fi}$, whereas the cost comes from consuming a negative amount in period two. When λ is sufficiently high, the cost outweighs the benefit and hence deters those without an idea from reporting having one.

The rest of this subsection considers the constrained-efficient allocation when $\lambda < \underline{\lambda}$. The social planner still transfers strictly positive NPV of resources to agents with an idea; however, now the amount is smaller than Δ^{fi} due to incentive compatibility.

The social planner's goal is still to make poor agents with an idea invest as close to the full information level as possible and do this without making her consume a huge negative amount. This pushes for a subsidy from other agents to poor agents with ideas, and incentive-compatibility constraints push in the reverse direction. The constrained-efficient allocation arises from this trade-off. The following corollary, which follows directly from Proposition 9 and the discussion above, states formally that in the constrained-efficient allocation there is a transfer of resources from agents with no ideas to those with ideas, even when $\lambda \in (1, \underline{\lambda})$.

Corollary 3 *Suppose Assumptions* 6 *and* 7 *hold, and* $\lambda \in (1, \underline{\lambda})$. *Then, in the constrained efficient allocation,* $\Delta^*(w, 1) > 0$, *for all* $w \in W$.

2.6 Conclusion

This paper provides a novel rationale for governments to subsidize agents who have ideas (potential start-ups) but do not have enough resources to invest in them. If we accept that returns to start-up firms are privately observed by the owners of the firms, then constrained efficiency calls for subsidizing poor agents with ideas. If society knew who has ideas but lacks resources to invest in them, then it is simple to implement the subsidy. However, I assume here that people's wealth levels, whether they have ideas or not, and how they use their resources are unobservable to others. These additional private information assumptions imply that the subsidy going to poor agents with ideas is limited by incentive compatibility.

The paper also provides an implementation of the constrained-efficient allocation similar to the U.S. SBA's Business Loan Program. Even though the main idea behind both the implementation in the model and the actual Business Loan Program are the same, to subsidize financially constrained individuals with productive ideas, there are still significant discrepancies between the model's implementation and the actual program. This is due mainly to the fact that the model economy is very simple. Introducing default and/or a costly monitoring technology may bring the model close enough to reality that the implementation of the model may allow us to analyze the efficiency of the details of the SBA's actual loan program and similar government programs in the rest of the world. This may be an interesting direction for future work.

Bibliography

Admati, A. R., and P. C. Peiderer (1994): "Robust Financial Contracting and the Role of Venture Capitalists," *Journal of Finance*, 49(2), 488–500.

Aghion, P., and P. Bolton (1997): "A Theory of Trickle-Down Growth and Development," *Review of Economic Studies*, 64(2), 151–172.

Aghion, P., E. Caroli, and C. Garcia-Penalosa (1999): "Inequality and Economic Growth: The Perspective of the New Growth Theories," *Journal of Economic Literature*, 37(4), 1615–1660.

Aiyagari, S. R., J. Greenwood, and A. Seshadri (2002): "Efficient Investment in Children," *Journal of Economic Theory*, 102(2), 290–321.

Atkeson, A., and R. E. Lucas (1992): "On Efficient Distribution with Private Information," *Review of Economic Studies*, 59(3), 427–453.

——— (1995): "Efficiency and Equality in a Simple Model of Efficient Unemployment Insurance," *Journal of Economic Theory*, 66(1), 64–88.

Banerjee, A. V., and A. F. Newman (1991): "Risk-Bearing and the Theory of Income Distribution," *Review of Economic Studies*, 58(2), 211–235.

Becker, G. S., and N. Tomes (1979): "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility," *Journal of Political Economy*, 87(6), 1153–1189.

——— (1986): "Human Capital and the Rise and Fall of Families," *Journal of Labor Economics*, 4(3), S1–S39.

Bernheim, B. D. (1989): "Intergenerational Altruism, Dynastic Equilibria and Social Welfare," *Review of Economic Studies*, 56(1), 119–128.

Bolton, P., and D. S. Scharfstein (1990): "A Theory of Predation Based on Agency Problems in Financial Contracting," *American Economic Review*, 80(1), 93–106.

Cunha, F., and J. Heckman (2006): "Formulating, Identifying, and Estimating the Technology of Cognitive and Noncognitive Skill Formation," *University of Chocago*, Unpublished paper.

Currie, J. (2001): "Early Childhood Education Programs," *Journal of Economic Perspectives*, 15(2), 213–238.

DeMarzo, P. M., and M. J. Fishman (2007): "Agency and Optimal Investment Dynamics," *Review of Financial Studies*, 20(1), 151–188.

Diamond, D. W. (1984): "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies*, 51(3), 393–414.

Farhi, E., and I. Werning (2007): "Inequality and Social Discounting," *Journal of Political Economy*, 115(3), 365–402.

——— (2008): "Progressive Estate Taxation," Unpublished Paper.

Gale, D., and M. Hellwig (1985): "Incentive-Compatible Debt Contracts: The One-Period Problem," *Review of Economic Studies*, 52(4), 647–663.

Galor, O., and J. Zeira (1993): "Income Distribution and Macroeconomics," *Review of Economic Studies*, 60(1), 35–52.

Golosov, M., and A. Tsyvinski (2007): "Optimal Taxation with Endogenous Insurance Markets," *Quarterly Journal of Economics*, 122(2), 487–534.

Gompers, P. A. (1995): "Optimal Investment, Monitoring, and the Staging of Venture Capital," *Journal of Finance*, 50(5), 1461–1490.

Green, E. J. (1987): "Lending and the Smoothing of Uninsurable Income," *In: Prescott, E.C., Wallace, N. (Eds.), Minnesota Studies in Macroeconomics, Vol. I: Contractual Arrangements for Intertemporal Trade, University of Minnesota*, pp. 3–25.

Hosseini, R., L. E. Jones, and A. Shourideh (2009): "Risk Sharing, Inequality, and Fertility," Unpublished manuscript.

Hubbard, G. R. (1998): "Capital-Market Imperfections and Investment," *Journal of Economic Literature*, 36(1), 193–225.

Jovanovic, B., and B. Szentes (2007): "On the Return to Venture Capital," Working Paper 12874, National Bureau of Economic Research.

Keane, M. P., and K. I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105(3), 473–522.

LEIBOWITZ, A. (1974): "Home Investments in Children," *Journal of Political Economy*, 82(2), S111–S131, Unpublished paper.

LOURY, G. C. (1981): "Intergenerational Transfers and the Distribution of Earnings," *Econometrica*, 49(4), 843–867.

MAZUMDER, B. (2001): "Earnings Mobility in the US: A New Look at Intergenerational Inequality," *Federal Reserve Bank of Chicago*, Working Paper 2001-18.

MYERSON, R. B. (1982): "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems," *Journal of Mathematical Economics*, 10(1), 67–81.

NEAL, D. A., AND W. R. JOHNSON (1996): "The Role of Premarket Factors in Black-White Wage Differentials," *Journal of Political Economy*, 104(5), 869–895.

PHELAN, C. J. (2006): "Opportunity and Social Mobility," *Review of Economic Studies*, 73(2), 487–504.

PRESCOTT, E. C., AND R. TOWNSEND (1984): "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard," *Econometrica*, 52(1), 21–45.

SESHADRI, A., AND K. YUKI (2004): "Equity and Efficiency Effects of Redistributive Policies," *Journal of Monetary Economics*, 51(7), 1415–1447.

SLEET, C., AND S. YELTEKIN (2006): "Credibility and Endogenous Societal Discounting," *Review of Economic Dynamics*, 9(3), 410–437.

SOLON, G. (1992): "Intergenerational Income Mobility in the Uninted States," *American Economic Review*, 82(3), 393–408.

STIGLITZ, J. (1987): "Pareto Efficient Income Taxation and the New New Welfare Economics," *Handbook of Public Economics*, 2, 991–1042.

STOKEY, N. L. (1996): "Shirtsleeves to Shirtsleeves: The Economics of Social Mobility," *Nancy L. Schwarz Lecture*, Delivered May 8, 1996 at the J.L. Kellogg School of Management Northwestern University.

THOMAS, J., AND T. WORRALL (1990): "Income fluctuation and asymmetric information: An example of a repeated principal-agent problem," *Journal of Economic Theory*, 51(2), 367–390.

TODD, P. E., AND K. I. WOLPIN (2006): "The Production of Cogntive Achievement in Children: Home, School and Racial Test Score Gaps," *University of Pennsylvania*, Unpublished paper.

ZIMMERMAN, D. J. (1992): "Regression Toward Mediocrity in Economic Stature," *American Economic Review*, 82(3), 409–429.