Ashoka University Economics
Discussion Paper 72

Perceptions, Biases, and Inequality

November 2021

Dyotona Dasgupta, O.P. Jindal Global University
Anuradha Saha, Ashoka University

https://ashoka.edu.in/economics-discussionpapers

Plot No. 2, Rajiv Gandhi Education City,
National Capital Region,
P.O.Rai, Sonepat, Haryana 131029

Perceptions, Biases, and Inequality*

Dyotona Dasgupta† Anuradha Saha‡

November 30, 2021

Abstract

In a novel framework, this paper captures the effects of *perceived self-efficacy beliefs*, built on the basis of the socio-economic background, on human capital investments and skill distribution. *Ex ante* children are homogeneous, but depending on parental education and job status, parents form different beliefs on the returns to their children's education. An unskilled (poor) parent underestimates the probability of her child getting a skilled job upon getting education and overestimates the corresponding income. The skilled (rich) parents do the opposite. We find that the steady-state mass of educated adults, skilled workers, and income inequality depend on the degree of bias and the degree of affinity for the well-being of the child. In economies with low child affinity, irrespective of the degree of bias, there is always a poverty trap. For moderate child affinity, behavioral biases may give rise to multiple equilibria as well as lower the steady-state inequality. For huge child affinity, even a small bias induces poor adults to invest with a higher probability than the rich.

Keywords: Behavioral Inequality, Human Capital Investment, Behavioral Bias, Behavioral Trap
JEL Codes: J62, D91, E2

*An earlier version of this paper was titled "The Glasses are Tinted: Self-Confidence and Poverty Trap". We acknowledge useful comments from Mausumi Das, Sayantan Ghosal, Parikshit Ghosh, Ajit Mishra, Dilip Mookherjee, Abhiroop Mukhopadhyay, Tridip Ray, Prabal Roy Chowdhury, Sudipta Sarangi, and Arunava Sen and participants of Delhi Theory Workshop at Indian Statistical Institute and seminar participants at Delhi School of Economics and Indian Statistical Institute, Delhi. Dyotona Dasgupta acknowledges postdoctoral fellowship funding jointly provided by the Centre for Development Economics, Delhi School of Economics and the Institute of Economic Growth during 2019-20. All errors are our own.

†Corresponding author. Jindal Global Business School, O.P. Jindal Global University, Narela Road, Sonepat, Haryana-131001, India. E-mail: dasgupta.dyotona@gmail.com

‡Ashoka University, Plot 2 Rajiv Gandhi Education City, Sonepat, Haryana 131029. India. E-mail: anuradha.saha@ashoka.edu.in.

1 Introduction

Socio-economic background plays an important role in investment decisions. Income, race, gender, etc. not only affect a person's access to opportunities, but also how they *perceive* benefits from different opportunities. For example, a high-school educated parent who owns a small business may invest less in their child's schooling due to their limited financial capabilities or because they *do not think* education would significantly help their child's employment prospects. Beliefs and biases matter. Attitudes of youngsters and their parents play a key role in explaining the rich-poor gap in secondary education attainment (Chowdry et al. (2011)).

In this paper, we construct a theoretical model to capture the fact that socio-economic background affects beliefs about self-efficacy and we study their long-run effects.[1] We find, extreme pessimism unambiguously (weakly) increases income inequality and the more affluent parents always invest in their children's education with a higher probability. More interestingly, in the presence of both over and under-confident parents, we find that such behavioral biases may reduce income inequality. It may even induce some less affluent parents to invest in their children's education with a higher probability relative to the richer parents. Further, the economy may witness multiple steady states as well as multiple paths to the steady states.

Do behavioral biases manifest in economics outcomes? In an empirical paper Chetty et al. (2019), establish that black-white inter-generational income gap for children, at the same level of parental income, is substantial for men *but not for women.*[2] Even after controlling for family or children's characteristics, the authors are not able to explain why black-white inter-generational gap is so small for females. In an environment where biases against black men are more widely prevalent compared to those against black women (Opportunity Insights), our theory can explain the differential racial gaps across gender. Biases, against black men in this case, not only limit the opportunities available to such individuals but also contribute to their perceptions on the relative benefits of various opportunities.

While research on the effects of behavioral anomalies is limited, it is increasingly becoming imperative to study behavioral inequality. We should do so for three reasons. First, in reality individuals are not as starkly sorted as is assumed in existing theories. People could be fortuitous in one dimension but not in others, for example a rich-uneducated carpenter versus a poor-college teacher. There is a need to formulate a generalized theory which captures such multi-dimensional people and depicts how such individuals access and evaluate their choices. We build such a theory where it is fundamentally *beliefs* which influence how people evaluate different opportunities. The theory can explain why people may take different decisions in spite of access to *similar* opportunities. Second, in the existing discussions on inequality (based on differences in income, gender, geography or ethnicity) there exists a superior group which always receives better outcomes compared to other groups. Again, such stark ranking of groups may not be true. In our model, there are under confident

[1] In psychology, there is a huge literature on self-efficacy, anchoring, and social identification (Cervone and Peake (1986), Gramzow et al. (2001), Van Veelen et al. (2016) to name a few). Ross (2019) shows that parental education and economic activity influence their and their children's beliefs about future job prospects.

[2] See Figure V in Chetty et al. (2019).

and overconfident groups, but each group can achieve relatively better outcomes in different economic conditions.[3] By including this element of realism, the paper can comment on conditions which lead to extreme inequality. Third, inequality stemming from external constraints limits the opportunities available to a person while behavioral inequality limits how individuals respond to opportunities. Behavioral public policies help change culture, which is usually amiss in other poverty alleviation programmes, and require fewer resources to implement.

In an overlapping generations model, we build a theoretical framework where a decision maker's assessment of returns to investment maybe biased. Adults differ in terms of their education – educated or uneducated, and jobs – skilled or unskilled. The income of a skilled worker is higher than that of an unskilled worker. This, however, requires a schooling cost as education is necessary, though not sufficient, for getting a skilled job. There is no intrinsic difference among educated children, i.e. the exogenous probability of an educated child getting a skilled job is the same for all. Adults derive utility from their own consumption and the perceived expected income of their children, and thus, they decide whether to invest in her child's education.

A biased adult forms beliefs about the probability of their educated child getting a skilled job based on their own education and job. When two adults share fewer common attributes, the "degree of association" is weaker and they believe that their own child is less likely to grow up to work in a job like that of the other adult. Due to higher dissimilarity with the skilled workers, an uneducated-unskilled is severely under confident about her child getting a skilled job, an educated-unskilled worker is under confident while an educated-skilled worker is overconfident.[4] All agents are non-Bayesian and there is no learning[5] or convergence of beliefs. However, *given their beliefs*, each parent correctly calculates the equilibrium mass of future skilled workers and their income. Unsuccessful people overestimate the reward in case of success and successful people do the converse. These two opposing forces – confidence and estimated rewards – determine the investment decision of any individual. While all agents have economic resources to invest, some adults do not invest as they perceive the returns not to be sufficiently high, thus capturing the effect of cognitive rather than economic limitations on investment.

We find that the relative weight a parent places on the utility from her child's perceived expected income plays a key role in the characterization of dynamic and steady state properties. We call this the *warm glow parameter* (or intergenerational altruism as in Ghatak (2015)). The parameter is non-negative,[6] time-independent, and common for all parents in an economy. In our characterization, the economic outcomes have distinct properties as per four ranges of the warm glow parameter, namely, low, moderate, high, and huge.

[3]In the confidence scale, extreme positions are unwelcome and being at the centre is most desirable. However, people do not sustain this disposition and tend to be under or overconfident.

[4]Deshpande and Newman (2007) find that graduating students from reserved (backward) categories have significantly lower occupational *expectations* than their non-reserved counterparts. Barber and Odean (2001) find men are overconfident and invest more compared to women.

[5]The absence of learning is prominently shown through experiments in Tversky and Kaheneman (2004).

[6]As noted in Boca et al. (2014) children may be valued more or less than parents' own consumption, correspondingly the degree of warm glow may be a fraction or greater than unity. (See Browning et al. (2014) pp. 106-120 for further discussion).

We highlight four of our main findings. First, biases can lower long-run income inequality. As uneducated parents are pessimistic, they do not invest in their children's education. This lowers the supply of the skilled workers and hence increases their income. Second, behavioral anomalies may influence educated-unskilled parents to invest more than the skilled parents. In an environment of relatively high warm glow and biases-driven expected future skilled incomes, the educated-unskilled parents may invest much more than the skilled parents because the perceived gains are larger for them. Third, when behavioral bias is high and degree of warm glow is moderate, there could be multiple equilibria. In a dynamic setting, this would cause aggregate fluctuations in investment and income, or lead to *behavior-driven business cycles*. Multiple equilibrium paths stem from the fact that the perceived expected benefit from an educational investment for each type of parent is different and can not be unambiguously ranked. Finally, the biases may persuade high-income households to invest at a rate lower than what they would have invested in the absence of biases.

In terms of policy making, our paper encapsulates that the manifestation of behavioral biases on economic outcomes depends on the environment. Two key examples of systematic behavioral inequality are women in STEM careers (Sterling et al. (2020)) and black-white investment in education in the United States (Klugman and Xu (2008)). The confidence gaps in female STEM or black students remain uncorrected because biases are difficult to change. Here, *behavioral policies* are required and we advocate for creating policies appropriate for local conditions. Our paper provides a theoretical foundation of behavioral policy-making, which has been advocated by Chetty et al. (2019), BRAC, student groups such as Harvard Law Review and corporate professionals (Forbes).

The plan of the paper is as follows. Next, we discuss the related literature. Section 3 studies the benchmark case where there is no behavioral anomaly. Section 4 addresses two types of behavioral anomalies – Section 4.1 analyzes the case where only uneducated workers are under behavioral trap whereas Section 4.2 addresses the case where all types of workers are biased. Section 5 compares these cases and formally shows the implications of the behavioral anomaly. In section 6 we discuss the effects of the behavioral anomaly. The final section concludes with some policy implications. Main proofs are collected in an Appendix. The proofs of the results marked (S) are available in a Supplementary Appendix (available online).

1.1 Related Literature

This paper is related to various strands of literature. There is a large literature addressing the physical constraints such as credit market imperfection, non-convex technology (Banerjee and Newman (1993), Galor and Zeira (1993), Mookherjee and Ray (2002)) to explain persistent inequality and poverty trap. We contribute to the emerging literature on behavioral inequality.[7] We depict behavioral biases differently from the existing literature. Behavioral theory on inequality and poverty, to the best of our knowledge, has relied on non-standard utility functions – time inconsistency (captured through quasi-hyperbolic discounting as in Bernheim et al. (2015)), or temptation (as in Banerjee and Mullainathan (2010)), or aspira-

[7] A theoretical comparison between external and internal constraints based explanations for poverty trap has been done in Ghatak (2015) and an empirical comparison is in Balboni et al. (2020).

tion (captured through 'milestone' utility in Ray (2006), Dalton et al. (2016) and subsequent papers). In contrast, the agents, in our model, maximize a standard lifetime utility function. The anomaly comes from the fact that their judgment is prejudiced by their own life experiences.

This paper brings the role of self-confidence in inter-generational investments. Traditionally, it was thought that "the most robust finding in the psychology of judgment is that people are overconfident" (De Bondt and Thaler, 1995). Overconfidence has been studied in housing market (Case and Shiller, 2003), among CEOs (Malmendier and Tate, 2005), in financial investment (Biais et al., 2005), among others. It has been argued that self-confidence enhances motivation (Bénabou and Tirole, 2002), improves self-control (Bénabou and Tirole, 2004). Of late, it is found that overconfidence is not as a robust characteristic as it was thought to be (Clark and Friesen, 2009). Moore (2007) finds that people are under (over) confident when the task in consideration is difficult (easy). In our paper, the *perception* about the difficulty of a task (getting a skilled job) is assumed to be group-identity specific.

As discussed above, there is empirical evidence that social background, gender, wealth etc. can influence individual's self-confidence. Filippin and Paccagnella (2012) develop a theoretical model where small differences in self-confidence lead to gaps in human capital. Unlike us, they consider a lack of self-confidence as an information problem while we look at it as a behavioral problem. In our model, people have self-deception; they believe that the probability of change in economic status is low; and there is no learning.

2 Model

2.1 The Firms

In a discrete-time framework, we consider a single good economy. The good can be produced in either a skilled sector or an unskilled sector, both differ in their technologies and the kind of labor employed. The mass of labor is normalized to one. Labor is of two types – skilled (L_{st}) and unskilled (L_{ut}).[8] The production function of the skilled sector is AL_{st}^{ϕ}, where $0 < \phi < 1$, and $A \geq 1$ – the production function is strictly increasing and strictly concave. The production function of the unskilled sector is L_{ut}. At any period t, the profit functions of the representative firms of the skilled and unskilled sectors, π_{st} and π_{ut}, are:

$$\pi_{st} = AL_{st}^{\phi} - w_{st}L_{st}, \qquad \text{and} \qquad \pi_{ut} = L_{ut} - w_{ut}L_{ut}$$

where w_{jt} denotes the wage rate of a worker of type j, $j = \{s, u\}$. Solving the profit maximization problems we get

$$w_{st} = A\phi L_{st}^{-(1-\phi)}, \quad \pi_{st} = (1-\phi)AL_{st}^{\phi}, \qquad \text{and } w_{ut} = 1. \tag{1}$$

The profit of the skilled sector is divided among skilled workers. So, the income of a skilled worker is $m_{st} \equiv w_{st} + \pi_{st}/L_{st} = AL_{st}^{-(1-\phi)}$ and that of an unskilled worker is $m_{ut} = 1$.

[8] In all notations, subscript t denotes time, and subscripts s and u designate skilled and unskilled workers, except where otherwise mentioned.

Observation 1 (S)**.** *A skilled worker earns (weakly)*[9] *more than an unskilled worker.*

All results with (S) notation are proved in the Supplementary Appendix.

2.2 The Households

We consider an overlapping generations model with no population growth. An individual lives for two periods: first as a child and later as an adult. In each generation, there is a continuum of individuals of size 1. Adults share a common degree of child affinity, $\delta(> 0)$. Each household consists of an adult and a child. The adult works, earns income, consumes, and decides whether to invest in her child's education.[10] Required investment in education is fixed at $\bar{s}$, where $\bar{s} \in (0, 1)$. Education is necessary but not sufficient for becoming a skilled worker – an educated individual, denoted e, becomes a skilled worker with probability β whereas an uneducated person, denoted n, becomes an unskilled worker with certainty: $Pr(L_t = L_{st}|e) = \beta$ and $Pr(L_t = L_{st}|n) = 0$.

An adult derives utility from her own consumption and from her child's *perceived* expected income earned in the next period. The utility of an adult of type ij where i denotes her education $i \in \{e, n\}$ and j denotes her skill $j \in \{s, u\}$ is

$$U_t^{ij}\left(c_t^{ij}, E\omega_{t+1}^{ij}\right) = \frac{\left(c_t^{ij}\right)^{\sigma}}{\sigma} + \delta \frac{\left(E\omega_{t+1}^{ij}\right)^{\sigma}}{\sigma}, \quad \sigma < 0$$

c_t^{ij} and $E\omega_{t+1}^{ij}$ denote her consumption and the *perceived* expected income of her child respectively. Observe, the utility function is strictly increasing and strictly concave.

The investment decision on a child's education is made on the basis of the *perceived* expected income of a child. It depends on the probability of her becoming a skilled worker upon getting the education, and the income she earns as a skilled worker. A parent forms beliefs about this probability, and based on that belief, she calculates the mass of skilled workers and their income in the next period. There is no inherent difference in the probability of getting a skilled job across educated children of different parent types (education and income). So, any type-dependent belief captures the agent's cognitive limitation. This is the only behavioral anomaly we focus on. The agent is, otherwise, rational. *Given her belief about the probability of her child becoming a skilled worker*, she accurately calculates the mass of skilled workers in the next period and makes the investment decision accordingly.

Let p_{t+1}^{ij} be the probability with which a parent of type ij believes that her educated child will become a skilled worker at $t+1$. Given her belief, a parent of type ij *conjectures* that the mass of skilled workers would be L_{st+1}^{ij}, and their income would be ω_{st+1}^{ij}. Thus, the *perceived* expected income of her educated child would be $E\omega_{t+1}^{ij} = p_{t+1}^{ij} \cdot \omega_{st+1}^{ij} + (1 - p_{t+1}^{ij}) \cdot \omega_{ut+1}^{ij} = p_{t+1}^{ij}\omega_{st+1}^{ij} + 1 - p_{t+1}^{ij}$. The expected income of an uneducated child is $E\omega_{t+1}^{ij} = 1$.

At any period t, a parent compares *perceived* expected utility from investing in her child's

[9]The income of a skilled and an unskilled worker could be equal only at $t = 0$, when the economy starts with all skilled workers and $A = 1$.

[10]For simplicity, we assume that an individual consumes only in her adulthood.

education with that from not investing and invests only when the former is (weakly) higher

$$
\begin{aligned}
& U_t^{ij}(\text{from investing in child's education}) \geq U_t^{ij}(\text{from not investing in child's education}) \\
\Rightarrow\quad & \frac{(m_{it}-\bar{s})^\sigma}{\sigma} + \delta \frac{[p_{t+1}^{ij}\cdot\omega_{st+1}^{ij} + (1-p_{t+1}^{ij})\cdot\omega_{ut+1}^{ij}]^\sigma}{\sigma} \geq \frac{m_{it}^\sigma}{\sigma} + \delta\frac{\omega_{ut+1}^{ij}}{\sigma} \\
\Rightarrow\quad & \delta\left[\frac{[p_{t+1}^{ij}\omega_{st+1}^{ij} + (1-p_{t+1}^{ij})]^\sigma}{\sigma} - \frac{1}{\sigma}\right] \geq \frac{m_{it}^\sigma}{\sigma} - \frac{(m_{it}-\bar{s})^\sigma}{\sigma} \quad \text{as } \omega_{ut+1}^{ij} = 1. \qquad (2)
\end{aligned}
$$

The L.H.S. of the above inequality is the *perceived* expected net benefit from investing in a child's education whereas the R.H.S. is the utility cost of making that investment.[11]

An equilibrium, in our model, has two features:

(i) Parents calculate the expected return from investment which must be consistent with their beliefs.

(ii) No parent has an incentive to deviate unilaterally.

3 Benchmark Case

All types of workers believe the probability of an educated child becoming a skilled worker is β. Thus, their optimal decisions differ only due to differences in their incomes.

Let, at any period t, the probability with which a worker of type j invests in her child's education be λ_{jt}. So, at period $t+1$, the mass of skilled workers and their income would be

$$L_{st+1} = \beta[\lambda_{st}L_{st} + \lambda_{ut}L_{ut}], \quad \text{and } m_{st+1} = A\big[\beta[\lambda_{st}L_{st} + \lambda_{ut}L_{ut}]\big]^{-(1-\phi)}.$$

At t, a worker of type j invests in her child's education with probability λ_{jt} if and only if

$$\delta\left[\frac{\big[\beta^\phi A[\lambda_{st}L_{st} + \lambda_{ut}L_{ut}]^{-(1-\phi)} + 1 - \beta\big]^\sigma}{\sigma} - \frac{1}{\sigma}\right] \geq \frac{m_{jt}^\sigma}{\sigma} - \frac{(m_{jt}-\bar{s})^\sigma}{\sigma}. \qquad (3)$$

where $L_{ut} = 1 - L_{st}$ and the inequality binds for j^{th} type when $\lambda_{jt} \in (0,1)$.

An equilibrium is denoted by $\langle \lambda_{ut}, \lambda_{st}\rangle$ which satisfies the features described in Section 2.2. Observe, here the equilibrium concept is Nash Equilibrium. Comparing the investment decisions of the skilled and unskilled workers, we find:

Lemma 1 (S). *Consider any equilibrium* $\langle \lambda_{ut}, \lambda_{st}\rangle$

1. *if an unskilled worker invests in her child's education with a positive probability (*$\lambda_{ut} > 0$*), then a skilled worker invests in her child's education with certainty (*$\lambda_{st} = 1$*),*

2. *at any period* t*, the probabilities of investment of both types of workers (weakly) increase with an increase in income of a skilled worker at that period.*

[11] If $\bar{s}$ were zero then all types of parents would have invested. If $\bar{s}$ were greater than 1, then no unskilled worker could have afforded the investment. The assumption $\bar{s} \in (0,1)$ rules out these uninteresting cases.

Intuitively, Part 1. is an immediate implication of our assumption of a concave utility function. It implies that the utility cost of investment decreases with the income of a worker. The benefit of investment is the same for all the parents. Hence, whenever an unskilled worker invests, a skilled worker with a higher income (see Observation 1) invests with certainty. For Part 2., we refer to equation (3). When the income of skilled workers increases, (i) the utility cost of investment for the skilled workers decreases whereas that of unskilled workers remains the same, and (ii) the benefit from investment increases. The reason for the former is, again, the concavity of the utility function. The intuition behind the latter is that the benefit from investment at any period t, increases with the probability of becoming a skilled worker (β) and the next period's income of a skilled worker (m_{st+1}). We show that the income of skilled workers of two consecutive periods is positively (non-negatively) related, keeping investment decisions the same. Therefore, the benefit from investment, at any period t, increases with the income of the skilled workers of that period. Note this lemma implies that the income of a skilled worker at any period is the state variable of that period.

The degree of child affinity[12] plays an important role in the parent's investment decision. Next, we define three thresholds of child affinity which will be useful in further analyses.

Definition 1. *The degree of child affinity is 'high' when* $\delta \geq \bar{\delta}$*, where* $\bar{\delta} \equiv \dfrac{(1-\bar{s})^{\sigma}-1}{1-(A\beta^{\phi}+1-\beta)^{\sigma}}$*, 'moderate' when* $\delta \in [\underline{\delta}, \bar{\delta})$*, where* $\underline{\delta} \equiv (1-\bar{s})^{\sigma}-1$*, and 'low' when* $\delta < \underline{\delta}$*.*

Observation 2 (S)**.** $0 < \underline{\delta} < \bar{\delta}$.

Consider any equilibrium $\langle \lambda_{ut}, \lambda_{st} \rangle$. Given Lemma 1, when $\lambda_{ut} > 0$, then $\lambda_{st} = 1$. Based on this, for a given degree of child affinity, we define three thresholds of the state variable.

Definition 2. *Let* $\langle \lambda_{ut}, \lambda_{st} \rangle$ *be an equilibrium at state variable* m_{st}*. For a given child affinity*

- $\underline{b}_s(\delta)$ *is the maximum value of the state variable, at which the skilled workers do not invest, i.e.* $\lambda_{st} = 0$ *if and only if* $m_{st} \leq \underline{b}_s(\delta)$*.*
- $\bar{b}_s(\delta)$ *is the minimum value of the state variable, at which skilled workers invest with certainty, i.e.* $\lambda_{st} = 1$ *if and only if* $m_{st} \geq \bar{b}_s(\delta)$*.*
- $\underline{b}_u(\delta)$ *is the maximum value of the state variable, at which unskilled workers do not invest, i.e.* $\lambda_{ut} = 0$ *if and only if* $m_{st} \leq \underline{b}_u(\delta)$*.*

The formal expressions of these thresholds of the state variable can be found in Appendix 8.1.

Next, we cumulate the ranking and other features of these thresholds in the following lemma.

Lemma 2 (S)**. Properties of the thresholds of the state variable**

1. $\underline{b}_s(\delta) < \bar{b}_s(\delta) < \underline{b}_u(\delta)$*, and all the thresholds are decreasing in* δ*.*
2. *Suppose, child affinity is (i) moderate, then* $\underline{b}_s(\delta) \leq 1 < \bar{b}_s(\delta)$*, and (ii) low, then* $\underline{b}_u(\delta) = \infty$ *and* $1 < \underline{b}_s(\delta)$*.*

[12]For brevity, we use child affinity and degree of child affinity interchangeably.

The intuition for part 1. is straightforward – (i) the benefit from investment increases with child affinity without changing the cost of investment, and (ii) at the thresholds, parents must be indifferent. To make them that, the thresholds must adjust. Hence, we find the thresholds decrease with an increase in child affinity. The ranking of the thresholds directly follows from Lemma 1.

Given the parameters $\delta, \sigma, \bar{s}, \beta$, and the state variable m_{st} of an economy, we, next, characterize the equilibria of this benchmark case.

Proposition 1. Characterization of the Equilibria

1. *For any parameter values and at any state variable, the equilibrium is unique.*

2. *Suppose child affinity is high. At the unique equilibrium, all parents invest with prob. 1.*

3. *Suppose the degree of child affinity is moderate. The unique equilibrium is such that if*

   a. $m_{st} > \underline{b}_u(\delta)$*: unskilled workers invest with a probability such that* (3) *binds and skilled workers invest with probability 1,*

   b. $m_{st} \in [\bar{b}_s(\delta), \underline{b}_u(\delta)]$*: unskilled workers do not invest and skilled workers invest with probability 1,*

   c. $m_{st} < \bar{b}_s(\delta)$*: unskilled workers do not invest and skilled workers invest with a probability such that* (3) *binds.*

4. *Suppose the degree of child affinity is low. The unique equilibrium is such that if*

   a. $m_{st} \geq \bar{b}_s(\delta)$*: unskilled workers do not invest and skilled workers invest with prob. 1,*

   b. $m_{st} \in (\underline{b}_s(\delta), \bar{b}_s(\delta))$*: unskilled workers do not invest and skilled workers invest with a probability s.t.* (3) *binds,*

   c. $m_{st} \leq \underline{b}_s(\delta)$*: no worker invests.*

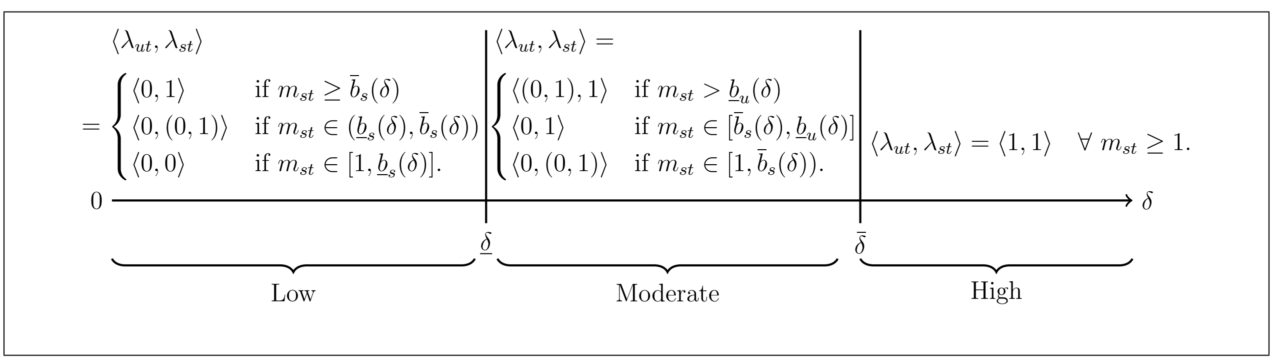


Figure 1: Characterization of the Equilibria in the Benchmark Case

We prove this in Appendix 8.2 and depict the equilibria in Figure 1.

The intuition behind this proposition is, now, immediate. The uniqueness follows from Lemma 1 – given that the benefits for both types of workers are equal and the utility cost of investment for the skilled workers is strictly lower, at any parametric condition, the probabilities of investment for both types of workers are unique.

When the degree of child affinity is high, the parents care for their children so much that they invest in the entire range of the state variable. When the degree of child affinity is moderate, the unskilled workers no longer invest with certainty and the probability of investment decreases with a decrease in the state variable. If the state variable falls below $\underline{b}_u(\delta)$, then the unskilled workers do not invest at all. As discussed above, $\underline{b}_u(\delta)$ is negatively related to the parent's degree of child affinity. It becomes infinite when the degree of child affinity is low – an unskilled worker with low child affinity never invests. The corresponding intuition for a skilled worker is similar. Only the thresholds are different as the income of a skilled worker is higher which makes her utility cost of investment lower.

Next, we analyze the dynamics and steady state of an economy. We say there is a *poverty trap* if there exists a positive mass of families that never become rich, which in our model corresponds to adult working as skilled workers. Alternatively, there is no poverty trap if at any period, the probability with which a family becomes rich is positive.

Proposition 2. Dynamics and the Steady States

1. *When the degree of child affinity is not low, there is no poverty trap in the economy.*

   a. *When child affinity is high, the economy immediately reaches the steady state – all parents invest, the mass of skilled workers is β and the income of a skilled worker is $A\beta^{-(1-\phi)}$. At any period, the probability with which a family becomes rich is β.*

   b. *When child affinity is moderate, the economy converges to a unique steady state. At the steady state, the unskilled workers randomize and the skilled workers invest with certainty. At any period, the probability with which a family becomes rich is lower than β and it decreases with a decrease in child affinity. The steady-state mass of skilled worker is $\beta\left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}}$ and their income is $\beta^{-(1-\phi)}\underline{b}_u(\delta)$.*

   c. *The income inequality at the steady state (weakly) increases with a decrease in child affinity – it remains constant for high child affinity and strictly increases with δ for moderate values of child affinity.*

2. *When child affinity is low, if the state variable is*

   a. *higher than $\underline{b}_s(\delta)$, then the mass of skilled workers decreases over time and tends to zero, correspondingly their income tends to infinity,*

   b. *no higher than $\underline{b}_s(\delta)$, then the economy immediately reaches the steady state where all workers are unskilled and no parent invests. At the steady state, all families are in a poverty trap and there is no inequality.*

We prove this in Appendix 8.3.
When the degree of child affinity is high, all types of workers invest with certainty. Thus, the economy immediately reaches the steady state where all children are educated. At any period, a family becomes rich with a positive probability, so there is no poverty trap. Since all parents invest with certainty at any $\delta \geq \bar{\delta}$, the inequality at the steady state – the

difference between the income of a skilled worker and that of an unskilled worker – remains constant in this range of child affinity.

When child affinity is moderate, the unskilled workers no longer invest with certainty. Recall, the income of skilled workers of consecutive periods are positively related. When the initial skilled income is no higher than $\bar{b}_s(\delta)$, then the expected benefit from investment is small such that only the skilled workers invest in their children's education. As incomes rise and exceed $\underline{b}_u(\delta)$, the future income becomes lucrative enough to make the unskilled workers invest with a positive probability. For moderate child affinity, there is a unique steady-state skilled income, $\beta^{-(1-\phi)}\underline{b}_u(\delta)$ at which unskilled workers invest with a constant probability (dependent on child affinity). Any deviation from the steady state would bring back the economy to the steady state.

Here at the steady state, the probability with which a family becomes rich is positive. However, that probability is less than β because unskilled workers invest with probability less than 1 and at any period, the probability that the adult of a family works as an unskilled worker is positive. The steady-state probability with which an unskilled worker invests decreases with a decrease in the degree of child affinity. So, the probability with which a certain family becomes rich at a particular period decreases with child affinity. The intuition behind the increase in inequality at the steady state with a decrease in child affinity is quite obvious. Lower child affinity implies a smaller probability of investment and that increases the steady-state skilled income and hence the inequality.

When child affinity is low, unskilled workers never invest. The skilled workers invest but only a β fraction of their children become skilled workers – the mass of skilled workers asymptotes to zero. In the steady state, everyone is unskilled and there is no inequality.
Next, we address the main focus of this paper – the case where the parents are biased.

4 Behavioral Anomaly

Parents underestimate the probability of intergenerational mobility. Each parent identifies herself with a group represented by a set of features or attributes namely education and job. The similarity between groups increases with the addition of common features (following Tversky (2004) pp. 10-11). An individual feels less connected with more dissimilar groups. We capture this through a *"degree of association"*. The degree of association between two individuals belonging to the same group is normalized to 1. Let the degree of association between two individuals belonging to two groups which differ by one attribute be θ and that when they differ by two attributes be η, so $\eta < \theta \in [0, 1)$.[13] Thus, the degree of association of an educated-unskilled[14] worker with a skilled worker is θ and that of an uneducated worker with a skilled worker is η as education is necessary to become a skilled worker.

While forming the beliefs about the probability of her educated child becoming a skilled

[13] Observe, in the benchmark case, $\eta = \theta = 1$.

[14] A word about notation: workers can be of three types – uneducated-unskilled, educated-unskilled, and educated-skilled. Here, we need to denote unskilled workers – uneducated versus educated – differently, as they choose differently. For brevity, in further analysis, we will denote the former as uneducated because without education it is not possible to get a skilled job and the latter as educated-unskilled. Similarly, as education is necessary for a skilled job, educated-skilled workers are denoted by skilled workers.

worker, a parent looks through her group identity. She discounts the possibility of her child becoming a worker of a different type than herself via the degree of association. Recall, the true probability with which an educated child becomes a skilled worker is independent of her parent's group identity. So, this captures the bias in our model.

4.1 Via Education: Behavioral Trap

We start our analysis with the case where only uneducated parents are biased. Lack of education imprisons them in a *behavioral trap* – they believe that an educated child from their group would never get a skilled job. A parent invests only when that provides her (weakly) higher utility. So, the immediate implication of η being zero is

Observation 3. *In presence of a behavioral trap, uneducated workers never invest.*

Educated parents take this into account and invest accordingly. Let the probability with which an educated-unskilled worker invests be ρ_{ut}[15] and that for a skilled worker be ρ_{st}.
At period t, a worker of type j invests in child's education with probability ρ_{jt} if and only if

$$\delta\left[\frac{\left[\beta^{\phi}A[\rho_{ut}\cdot(1-\beta)N_{et}+\rho_{st}\cdot\beta N_{et}]^{-(1-\phi)}+1-\beta\right]^{\sigma}}{\sigma}-\frac{1}{\sigma}\right]\geq\frac{m_{jt}^{\sigma}}{\sigma}-\frac{(m_{jt}-\bar{s})^{\sigma}}{\sigma}. \tag{4}$$

recall N_{et} is the mass of educated workers, $(1-\beta)N_{et}$ is that of educated-unskilled workers and βN_{et} is that of skilled workers. The inequality binds for j^{th} type when $\rho_{jt}\in(0,1)$.
An equilibrium $\langle\rho_{ut},\rho_{st}\rangle$ satisfies the features stated in Section 2.2. Like the benchmark case,

Observation 4. *In any equilibrium $\langle\rho_{ut},\rho_{st}\rangle$, if educated-unskilled workers invest with a positive probability ($\rho_{ut}>0$), then all skilled workers invest with certainty ($\rho_{st}=1$).*

The proof is very similar to that of Lemma 1, so we skip it here.
Due to the behavioral trap, there does not exist any degree of child affinity where all parents invest. We define the following new threshold of the state variable.

Definition 3. *Let $\langle\rho_{ut},\rho_{st}\rangle$ be an equilibrium at the state variable m_{st}. For a given degree of child affinity, $\bar{b}_u(\delta)$ is the minimum value of the state variable (m_{st}) at which educated-unskilled workers invest with certainty, i.e. $\rho_{ut}=1$ if and only if $m_{st}\geq\bar{b}_u(\delta)$.*

We provide the formal expression of this threshold of the state variable in Appendix 8.4.
The thresholds stated in Definition 2 continue to be relevant here. As in the benchmark case, unskilled workers invest only when the state variable is higher than $\underline{b}_u(\delta)$. So the effect of the behavioral trap, via non-investment of uneducated workers, does not change these thresholds. The following observation documents some features of the new threshold:

Observation 5 (S). *(i) For $\delta\in[\underline{\delta},\bar{\delta})$, $\bar{b}_u(\delta)$ is decreasing in δ, $\bar{b}_u(\delta)=\beta^{-(1-\phi)}\underline{b}_u(\delta)$ and $\bar{b}_u(\bar{\delta})=A\beta^{-(1-\phi)}$. (ii) For $\delta\in(0,\bar{\delta})$, $\bar{b}_u(\delta)=\infty$.*

[15]Here, unlike the benchmark case, subscript u denotes educated-unskilled. Uneducated workers never invest, so this is for the brevity of notation.

Given the parameters $\delta, \sigma, \bar{s}, \beta, \eta$ and state variable m_{st}, we next characterize the equilibria.

Proposition 3. Characterization of the Equilibria

1. *For any parameter values and at any state variable, the equilibrium is unique.*
2. *The uneducated workers never invest.*
3. *If child affinity is high, at the unique equilibrium all educated workers invest with prob. 1.*
4. *Suppose the degree of child affinity is moderate. The unique equilibrium is such that if*
   a. $m_{st} \geq \bar{b}_u(\delta)$*: all educated workers invest with probability 1,*
   b. $m_{st} \in (\underline{b}_u(\delta), \bar{b}_u(\delta))$*: educated-unskilled workers invest with a probability such that* (4) *binds, and skilled workers invest with probability 1,*
   c. $m_{st} \in [\bar{b}_s(\delta), \underline{b}_u(\delta)]$*: educated-unskilled workers do not invest, and skilled workers invest with probability 1,*
   d. $m_{st} < \bar{b}_s(\delta)$*: educated-unskilled workers do not invest and the skilled workers invest with a probability such that* (4) *binds.*
5. *Suppose child affinity is low, unique equilibrium is such that if*
   a. $m_{st} \geq \bar{b}_s(\delta)$*: educated-unskilled workers do not invest and skilled workers invest with probability 1,*
   b. $m_{st} \in (\underline{b}_s(\delta), \bar{b}_s(\delta))$*: educated-unskilled workers do not invest, and skilled workers invest with a probability such that* (4) *binds,*
   c. $m_{st} \leq \underline{b}_s(\delta)$*: no worker invests.*

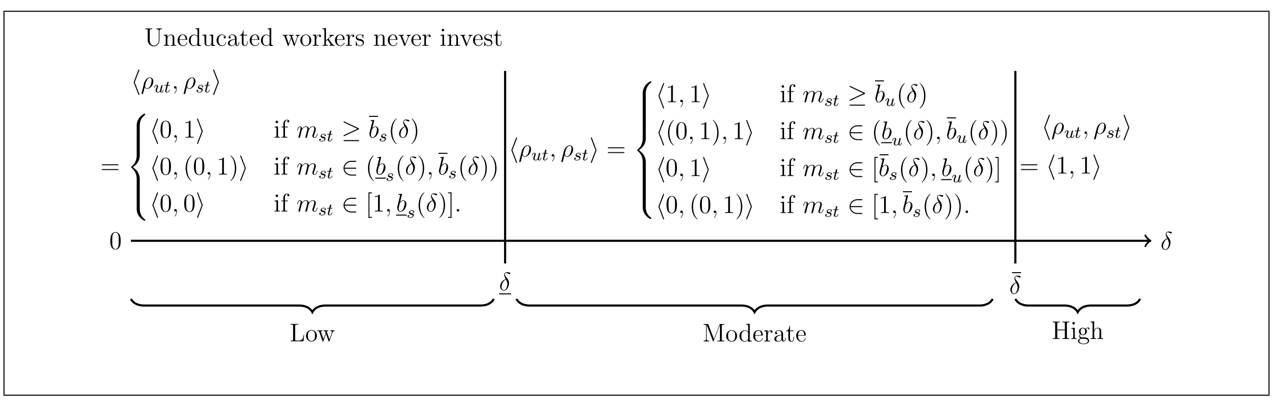


Figure 2: Characterization of the Equilibria with a Behavioral Trap

We prove this in Appendix 8.5 and depict the equilibria in Figure 2.

Let us highlight an interesting implication of the behavioral trap. Compared to the benchmark, the educated workers always invest with a weakly higher probability. While the probability of investment is unchanged for the skilled workers, the educated-unskilled workers invest with a strictly higher probability at a certain parametric condition, which entails that in the benchmark case, the probability of investment for the unskilled workers is a strict fraction and the mass of uneducated workers is positive. This follows from comparing the utility benefit of the educated workers (as the costs of investment are the same) in the two set-ups:

$$\lambda_{ut}L_{ut} + L_{st} = \gamma_{ut}(1-\beta)N_{et} + \beta N_{et} \qquad \Rightarrow \qquad \lambda_{ut}\underbrace{[1-N_{et}]}_{\text{uneducated}} = (\gamma_{ut} - \lambda_{ut})\ \underbrace{(1-\beta)N_{et}}_{\text{educated-unskilled}}.$$

Thus, when the mass of uneducated workers is positive, $1 - N_{et} > 0$, their non-investment is compensated by over investment of educated-unskilled workers.

The next proposition depicts the dynamics and steady states of an economy with the behavioral trap. Behavioral trap gives rise to multiple steady states. We define 'least unequal steady state' as the steady state with the lowest inequality between skilled and unskilled workers. When child affinity is high, it is the steady state where $m_s^* = \bar{b}_u(\bar{\delta}) = A\beta^{-(1-\phi)}$. And, when child affinity is moderate, it is the steady state where $m_s^* = \bar{b}_u(\delta)$.

Proposition 4. Dynamics and the Steady States

1. *There is almost always a poverty trap in an economy.*
2. *Dynamics – When the degree of child affinity is*
   - a. *high: any $m_s \geq 1$ is a steady state where all educated workers invest with probability 1. The steady-state income of a skilled worker is the initial income, $m_s^* = m_{st}$,*
   - b. *moderate and $m_{st} \geq \bar{b}_u(\delta)$: economy immediately reaches a steady state where educated workers invest with prob. 1. The steady-state income of a skilled worker is $m_s^* = m_{st}$,*
   - c. *moderate and $m_{st} < \bar{b}_u(\delta)$: unskilled workers invest with probability less than 1. The mass of educated individuals and the mass of skilled workers decrease over time. The income of a skilled worker increases over time and converges to some $m_s^* \geq \bar{b}_u(\delta)$,*
   - d. *low and $m_{st} > \underline{b}_s(\delta)$: the mass of skilled workers decreases over time and tends to zero. The income of a skilled worker tends to infinity,*
   - e. *low and $m_{st} \leq \underline{b}_s(\delta)$: unique steady state is immediately reached where no one invests.*
3. *Properties of the Steady States – When the degree of child affinity is*
   - a. *not low: there are multiple steady states ranked on the basis of inequality. The inequality at the least unequal steady state (weakly) increases with a decrease in child affinity – remains constant when child affinity is high, and strictly increases when it is moderate,*
   - b. *low: at the unique steady state all workers are unskilled.*

We prove this in Appendix 8.6.

Apart from the multiple steady states, there is another interesting implication of behavioral trap – a mixed strategy being played at the steady state is not possible, as that would decrease the mass of educated, hence skilled workers over time. Thus, an individual's dynasty's education status does not change in the long run. Only the job status of educated workers can change across generations. The inter-generational job mobility among educated workers is bi-directional – skilled to unskilled and the other way around.

Here, it is important to emphasize that the existence of behavioral trap affects society at large, even though only the uneducated are imprisoned in the behavioral trap. With a smaller mass of workers investing in education, the return to education is higher in comparison to the benchmark case. The inequality is weakly higher in the society.

Next, we discuss the case where the educated parents are also biased.

4.2 Via Education & Job: Behavioral Trap & Behavioral Bias

All parents are biased. An educated-unskilled worker believes that an educated child from her group becomes a skilled worker with probability $\theta\beta$. And, a skilled worker believes that with probability $\theta(1-\beta)$[16] an educated child from her group becomes an unskilled worker. So, she believes the probability that such a child becomes a skilled worker is $1-\theta(1-\beta)$. As $\theta < 1$, educated-unskilled workers are under confident and skilled workers are overconfident.[17] Like before, uneducated workers are imprisoned in a behavioral trap, so they do not invest. We assume that while calculating the probability with which an educated child from a different group becomes a skilled worker, an individual can see clearly.

Since a parent's belief about the probability of success – becoming a skilled worker – of an educated child from her own group is type dependent, the 'conjectured' mass of skilled workers and their income would also be type dependent. To characterize the investment decisions, we discuss the conjectured expected benefit from investment. Suppose, at period t, a worker of type j, where $j \in \{u, s\}$,[18] invests with probability γ_{jt}. Then an educated-unskilled worker conjectures that the mass of skilled workers and their income would be

$$L^u_{st+1} = \theta\beta \cdot \gamma_{ut}(1-\beta)N_{et} + \beta \cdot \gamma_{st}\beta N_{et}, \quad \text{and } \omega^u_{st+1} = AL^{u^{-(1-\phi)}}_{st+1}.$$

recall, N_{et} is the mass of educated workers, $(1-\beta)N_{et}$ is the mass of educated-unskilled who invest with probability γ_{ut} and βN_{et} is the mass of skilled workers who invest with γ_{st}. Thus, the conjectured benefit from investment of an educated-unskilled worker is

$$\theta\beta \cdot \left[A(\beta N_{et})^{-1-\phi}[\theta(1-\beta)\gamma_{ut}+\beta\gamma_{st}]^{-1-\phi}\right] + 1-\theta\beta = \theta\beta[\theta(1-\beta)\gamma_{ut}+\beta\gamma_{st}]^{-(1-\phi)}m_{st} + 1-\theta\beta.$$

A skilled worker conjectures that the mass of skilled workers and their income would be

$$L^s_{st+1} = \quad \beta \cdot \gamma_{ut}(1-\beta)N_{et} + [1-\theta(1-\beta)] \cdot \gamma_{st}\beta N_{et} \quad \text{and } \omega^s_{st+1} = AL^{s^{-(1-\phi)}}_{st+1}.$$

Therefore, the conjectured benefit from investment of a skilled worker is

$$[1-\theta(1-\beta)] \cdot \left[(1-\beta)\gamma_{ut} + [1-\theta(1-\beta)]\gamma_{st}\right]^{-(1-\phi)} m_{st} + \theta(1-\beta).$$

[16] Recall, $1-\beta$ is the probability with which an educated individual becomes an unskilled worker.

[17] $\theta\beta < \beta \leq 1-\theta(1-\beta)$.

[18] Here also, we use subscript u for educated-unskilled workers and subscript s for the skilled workers.

Observe, here the conjectured benefits of the two types of workers cannot be ranked. This is because the under (over) confident educated-unskilled (skilled) workers under (over) estimate the mass of future skilled workers and hence, over (under) estimate their income.
At any period t, an educated-unskilled worker invests with probability γ_{ut} if and only if

$$\frac{(1-\bar{s})^{\sigma}}{\sigma} + \delta \frac{\left[\theta\beta[\theta(1-\beta)\gamma_{ut} + \beta\gamma_{st}]^{-(1-\phi)}m_{st} + 1 - \theta\beta\right]^{\sigma}}{\sigma} \geq \frac{1}{\sigma} + \frac{\delta}{\sigma}$$
$$\Rightarrow \quad \delta\left[\frac{\left[\theta\beta[\theta(1-\beta)\gamma_{ut} + \beta\gamma_{st}]^{-(1-\phi)}m_{st} + 1 - \theta\beta\right]^{\sigma}}{\sigma} - \frac{1}{\sigma}\right] \geq \frac{1}{\sigma} - \frac{(1-\bar{s})^{\sigma}}{\sigma}. \tag{5}$$

The L.H.S. is the *conjectured* net benefit and the R.H.S. is the net utility cost from investment. Similarly, at any t, a skilled worker invests with probability γ_{st} if and only if

$$\delta\left[\frac{\left[[1-\theta(1-\beta)]\cdot\left[(1-\beta)\gamma_{ut} + [1-\theta(1-\beta)]\gamma_{st}\right]^{-(1-\phi)}m_{st} + \theta(1-\beta)\right]^{\sigma}}{\sigma} - \frac{1}{\sigma}\right]$$
$$\geq \frac{m_{st}^{\sigma}}{\sigma} - \frac{(m_{st}-\bar{s})^{\sigma}}{\sigma}. \tag{6}$$

The utility cost of investment is lower for a skilled worker. But, the *conjectured* net benefits from investment cannot be ranked, so Part 1. of Lemma 1 is no longer true. However,

Lemma 3 (S). *At any equilibrium, if educated-unskilled workers invest, then skilled workers invest with positive probability: suppose $\langle\gamma_{ut}, \gamma_{st}\rangle$ is an equilibrium, and $\gamma_{ut} > 0$ then $\gamma_{st} > 0$.*

Intuitively, when no skilled workers invest and educated-unskilled workers invest then the conjectured benefit of a skilled worker is higher than that of an educated-unskilled worker. We have already observed that the utility cost of a skilled worker is lower. Hence, the lemma.
We, now, define an additional threshold of degree of child affinity for further analyses.

Definition 4. *Child affinity is huge when* $\delta \geq \delta_a \equiv \dfrac{(1-\bar{s})^{\sigma} - 1}{1 - [\theta\beta(\theta(1-\beta)+\beta)^{-(1-\phi)} + 1 - \theta\beta]^{\sigma}}$.

δ_a along with $\underline{\delta}$, as in Definition 1, characterize the equilibria. The ranking is as follows.

Observation 6 (S). $0 < \underline{\delta} < \delta_a$.

The degree of child affinity is *moderately high* when $\underline{\delta} \leq \delta < \delta_a$ and recall *low* when $\delta < \underline{\delta}$.

The next observation follows directly from the optimal investment decisions of educated-unskilled workers and skilled workers as stated in equations (5) and (6) respectively.

Observation 7. *Suppose at any m_{st}, when workers of type k invest with probability γ_{kt}, the workers of type j optimally invest with probability γ_{jt}, where $k, j = \{u, s\}$ and $k \neq j$. Then at any $\tilde{m}_{st} > m_{st}$, when workers of type k invest with probability no higher than γ_{kt}, the workers of type j optimally invest with probability no less than γ_{jt}.*

Now, we introduce various thresholds of the state variable m_{st}. The first threshold addresses an equilibrium. The rest of the thresholds address optimal decisions – second, third and the fourth (or the last three) thresholds relate to the optimal decisions of the skilled (or educated-unskilled) workers *if* they believe that the educated-unskilled (or skilled) workers choose the mentioned γ_{ut} (or γ_{st}). Note at such a threshold, the mentioned $\langle \gamma_{ut}, \gamma_{st} \rangle$ may not be an equilibrium or there may exist another equilibrium.

Definition 5. *For a given degree of child affinity,*

- *suppose $\langle \gamma_{ut}, \gamma_{st} \rangle$ is an equilibrium. $\underline{a}_s(\delta)$ is the maximum value of the state variable at which the skilled workers do not invest,*
- *suppose the educated-unskilled workers do not invest, then $a_6(\delta)$ is the minimum value of the state variable at which skilled workers invest with certainty,*
- *suppose the educated-unskilled workers invest with probability 1, then $a_4(\delta)$ is the maximum value of the state variable at which skilled workers do not invest,*
- *suppose the educated-unskilled workers invest with probability 1, then $a_2(\delta)$ is the minimum value of the state variable at which skilled workers invest with certainty,*
- *suppose the skilled workers invest with probability 1, then $a_5(\delta)$ is the maximum value of the state variable at which educated-unskilled workers do not invest,*
- *suppose the skilled workers do not invest, then $a_3(\delta)$ is the minimum value of the state variable at which educated-unskilled workers invest with certainty,*
- *suppose the skilled workers invest with probability 1, then $a_1(\delta)$ is the minimum value of the state variable at which educated-unskilled workers invest with certainty.*

The formal expressions for these thresholds are given in Appendix 8.7.
For further analyses, in the following lemma, we collect important features of the thresholds.

Lemma 4 (S)**. Properties of the thresholds of the state variable**

1. *All thresholds of the state variable are decreasing in δ.*
2. *The thresholds related to the skilled workers' investment decisions are such that: $\forall\ \delta > 0$, we have (i) $1 < a_2(\delta)$, (ii) $\underline{a}_s(\delta) < a_4(\delta) < a_2(\delta)$, (iii) $\underline{a}_s(\delta) < a_6(\delta) < a_2(\delta)$, and (iv) if and only if $\theta(1-\beta) > \beta$, $a_4(\delta) > a_6(\delta)$.*
3. *The thresholds related to the educated-unskilled workers' decisions are such that:*
   a. *If and only if $\delta > \underline{\delta}$, $\quad a_1(\delta)$, $a_3(\delta)$ and $a_5(\delta)$ are finite.*
   b. *If and only if $\delta < \delta_a$, $\quad 1 < a_1(\delta)$.*
   c. *$\forall\ \delta > \underline{\delta}$, $\quad a_3(\delta) > a_5(\delta)$ if and only if $\theta(1-\beta) > \beta$, and $\max\{a_5(\delta), a_3(\delta)\} < a_1(\delta)$.*
4. *$\forall\ \delta \leq \delta_a$, we have $a_4(\delta) \leq a_1(\delta)$ and $\forall\ \delta > \delta_a$, we have $a_4(\delta) < 1$.*
5. *Cut-offs relative to the benchmark case: (i) $\underline{b}_s(\delta) = \underline{a}_s(\delta)$, and (ii) $\underline{b}_u(\delta) < a_5(\delta)$.*

Now, we provide the intuition of some of the properties of the thresholds depicted in Lemma 4. Observe, on the one hand, the benefit from the investment of any type of worker is decreasing in the conjectured mass of skilled workers and is increasing in the state variable. On the other hand, the cost of investment is non-increasing in the state variable – it is decreasing for the skilled workers and constant for the educated-unskilled workers. Hence, the ranking of the thresholds depend on the conjectured mass of skilled workers at the premises of the definitions – higher is that mass higher is the threshold. For example, at the premise of $\underline{a}_s(\delta)$ the conjectured mass of skilled worker is zero whereas that at $a_6(\delta)$ is positive – at any given state variable, the benefit from investment at the premise of $\underline{a}_s(\delta)$ is higher than at the premise of $a_6(\delta)$. Thus, $a_6(\delta)$ must be higher than $\underline{a}_s(\delta)$. The other rankings depicted in Part 2. (ii), (iii), and those in Part 3. c. follow from similar reasoning. The ranking between $a_4(\delta)$ and $a_6(\delta)$ follows from the additional fact that when $\theta(1-\beta) > \beta$ then the educated-unskilled workers' conjectured mass of skilled workers in the next period coming from their group is higher than the skilled workers' conjectured mass of skilled workers in the next period coming from their group. The same goes for the ranking between $a_3(\delta)$ and $a_5(\delta)$. Observe the state variable, by assumption, cannot be less than 1. So, if any threshold of the state variable is less than 1, and the premise of the definition is satisfied, then the optimal strategy described in the definition would always be true. For example, we show that $a_1(\delta) < 1$ when $\delta > \delta_a$. This implies when child affinity is huge and all skilled workers invest with certainty, then irrespective of the value of the state variable, the educated-unskilled workers optimally invest with certainty. Here, it further implies when child affinity is huge, the educated-unskilled workers invest with certainty, irrespective of the investment decision of the skilled workers. Finally, the intuition behind the ranking between $a_4(\delta)$ and $a_1(\delta)$ follows directly from Lemma 3.

Next, we provide boundary conditions on equilibrium strategies. The first two conditions provide lower and upper bounds, respectively, on the equilibrium strategy of the skilled workers and the last two provide the same of the educated-unskilled workers.

Boundary Conditions for Equilibrium Probabilities. *Consider any equilibrium $\langle\gamma_{ut}, \gamma_{st}\rangle$, then γ_{st} satisfies Condition $\underline{\Gamma}_s$ and Condition $\bar{\Gamma}_s$, and γ_{ut} satisfies Condition $\underline{\Gamma}_u$ and Condition $\bar{\Gamma}_u$ where the conditions are as follows:*

Condition $\underline{\boldsymbol{\Gamma}}_{\mathbf{s}}$ *for any m_{st}, γ_{st} is bounded below by $\underline{\gamma}_s(m_{st})$,*

Condition $\bar{\boldsymbol{\Gamma}}_{\mathbf{s}}$ *for any m_{st}, γ_{st} is bounded above by $\bar{\gamma}_s(m_{st})$,*

Condition $\underline{\boldsymbol{\Gamma}}_{\mathbf{u}}$ *for any m_{st}, γ_{ut} is bounded below by $\underline{\gamma}_u(m_{st})$,*

Condition $\bar{\boldsymbol{\Gamma}}_{\mathbf{u}}$ *for any m_{st}, γ_{ut} is bounded above by $\bar{\gamma}_u(m_{st})$.*

The formal expressions are given in Appendix 8.8. There we also show $\underline{\gamma}_s(m_{st})$ is strictly increasing $\forall\ m_{st} \in [a_4(\delta), a_2(\delta))$, $\bar{\gamma}_s(m_{st})$ is strictly increasing $\forall\ m_{st} \in [\underline{a}_s(\delta), a_6(\delta))$, and

$$\underline{\gamma}_s(m_{st}) \begin{cases} = 0 & \forall\ m_{st} \leq a_4(\delta), \\ \in (0,1) & \forall\ m_{st} \in (a_4(\delta), a_2(\delta)), \\ = 1 & \forall\ m_{st} \geq a_2(\delta), \end{cases} \text{ and } \bar{\gamma}_s(m_{st}) \begin{cases} = 0 & \forall\ m_{st} \leq \underline{a}_s(\delta), \\ \in (0,1) & \forall\ m_{st} \in (\underline{a}_s(\delta), a_6(\delta)), \\ = 1 & \forall\ m_{st} \geq a_6(\delta). \end{cases}$$

Similarly, $\underline{\gamma}_u(m_{st})$ is strictly increasing $\forall m_{st} \in [a_5(\delta), a_1(\delta))$, and $\bar{\gamma}_u(m_{st})$ is non-decreasing $\forall\ m_{st} < a_3(\delta)$, and

$$\underline{\gamma}_u(m_{st}) \begin{cases} = 0 & \forall\ m_{st} \leq a_5(\delta), \\ \in (0,1) & \forall\ m_{st} \in (a_5(\delta), a_1(\delta)), \\ = 1 & \forall\ m_{st} \geq a_1(\delta), \end{cases} \text{ and } \bar{\gamma}_u(m_{st}) \begin{cases} = 1 & \forall\ m_{st} \geq a_3(\delta), \\ < 1 & \text{at } m_{st} < a_3(\delta). \end{cases}$$

A word about how we get these boundary conditions. Let us consider **Condition $\underline{\Gamma}_s$** and **Condition $\bar{\Gamma}_s$**. $\underline{\gamma}_s(m_{st})$ captures the optimal response of a skilled worker when $\gamma_{ut} = 1$, and $\bar{\gamma}_s(m_{st})$ captures the optimal response of a skilled worker when $\gamma_{ut} = 0$. Since, the probability of investment of educated-unskilled workers can at most be one, and is at least zero, at any equilibrium γ_{st} is bounded below by $\underline{\gamma}_s(m_{st})$, and above by $\bar{\gamma}_s(m_{st})$. From the definition of $a_4(\delta)$, we can see that at $a_4(\delta)$, $\underline{\gamma}_s(m_{st})$ is equal to zero. From Observation 7, we can see that for any $m_{st} < a_4(\delta)$, $\underline{\gamma}_s(m_{st})$ is zero, and for any $m_{st} > a_4(\delta)$ it is positive and increasing. From the definition of $a_2(\delta)$, we can see that at $a_2(\delta)$, $\underline{\gamma}_s(m_{st})$ is one and again from Observation 7, we note that for $m_{st} > a_2(\delta)$, $\underline{\gamma}_s(m_{st})$ continues to be one. Similar intuition follows for **Condition $\underline{\Gamma}_u$**, and **Condition $\bar{\Gamma}_u$**.

Next, given parameters $\delta, \sigma, \bar{s}, \beta, \eta, \theta$, and state variable m_{st}, we characterize the equilibria.

Proposition 5. Characterization of the Equilibria

1. *The uneducated workers never invest.*

2. *Suppose child affinity is huge. At any $m_{st} \geq 1$, there exists a unique equilibrium $\langle \gamma_{ut}, \gamma_{st} \rangle$: educated-unskilled workers invest with prob. 1, skilled workers invest with $\underline{\gamma}_s(m_{st})$ as in* **Condition $\underline{\Gamma}_s$**.

3. *Suppose the degree of child affinity is moderately high, i.e. $\delta \in (\underline{\delta}, \delta_a]$*

   a. *at any $m_{st} \geq \min\{a_1(\delta), a_2(\delta)\}$, there exists a unique equilibrium $\langle \gamma_{ut}, \gamma_{st} \rangle$ where at least one type of workers invest with probability 1, and the other type, say type j, invests with probability $\underline{\gamma}_j(m_{st})$:*

      (i) if $a_1(\delta) \leq a_2(\delta)$ then $\gamma_{ut} = 1$ and $\gamma_{st} = \underline{\gamma}_s(m_{st})$ $\forall\ m_{st} \geq a_1(\delta)$,

      (ii) if $a_1(\delta) > a_2(\delta)$ then $\gamma_{ut} = \underline{\gamma}_u(m_{st})$ and $\gamma_{st} = 1$ $\forall\ m_{st} \geq a_2(\delta)$,

      where $\underline{\gamma}_s(m_{st})$ and $\underline{\gamma}_u(m_{st})$ as in **Condition $\underline{\Gamma}_s$** *and* **Condition $\underline{\Gamma}_u$**,

   b. *at any $m_{st} \in [1, \min\{a_1(\delta), a_2(\delta)\})$, there could be multiple equilibria only when $\beta \geq \theta(1-\beta)$, otherwise, there is a unique equilibrium. At any such equilibrium, at least one type of parents invest with a positive probability and* **Condition $\underline{\Gamma}_s$**, **Condition $\bar{\Gamma}_s$**, **Condition $\underline{\Gamma}_u$**, *and* **Condition $\bar{\Gamma}_u$** *are satisfied. Further, if $\beta < \theta(1-\beta)$ and $a_1 < a_2$, then $\gamma_{st} < \gamma_{ut}$. And, if $\beta > \theta(1-\beta)$ and at any $m_{st} \geq 1$ there are multiple equilibria, then at most in one such equilibrium both types of workers play mixed strategies.*

4. *Suppose the degree of child affinity is low, the equilibrium $\langle \gamma_{ut}, \gamma_{st} \rangle$ is unique, such that at any m_{st}: educated-unskilled workers do not invest and skilled invest with prob. $\bar{\gamma}_s(m_{st})$.*

Uneducated workers never invest

$\langle \gamma_{ut}, \gamma_{st} \rangle = \begin{cases} \langle 0, 1 \rangle & \text{if } m_{st} \geq a_6(\delta) \\ \langle 0, (0,1) \rangle & \text{if } m_{st} \in (\underline{a}_s(\delta), a_6(\delta)) \\ \langle 0, 0 \rangle & \text{otherwise.} \end{cases}$

$\langle \gamma_{ut}, \gamma_{st} \rangle$ is such that

- Unique equilibrium whenever $\beta < \theta(1-\beta)$,
- $max\{\gamma_{ut}, \gamma_{st}\} > 0$,
- Condition $\underline{\mathbf{\Gamma}}_{\mathbf{s}}$, Condition $\bar{\mathbf{\Gamma}}_{\mathbf{s}}$, Condition $\underline{\mathbf{\Gamma}}_{\mathbf{u}}$, and Condition $\bar{\mathbf{\Gamma}}_{\mathbf{u}}$ are satisfied.

$\langle \gamma_{ut}, \gamma_{st} \rangle = \langle 1, \underline{\gamma}_s(m_{st}) \rangle \quad \forall\ m_{st} \geq 1.$

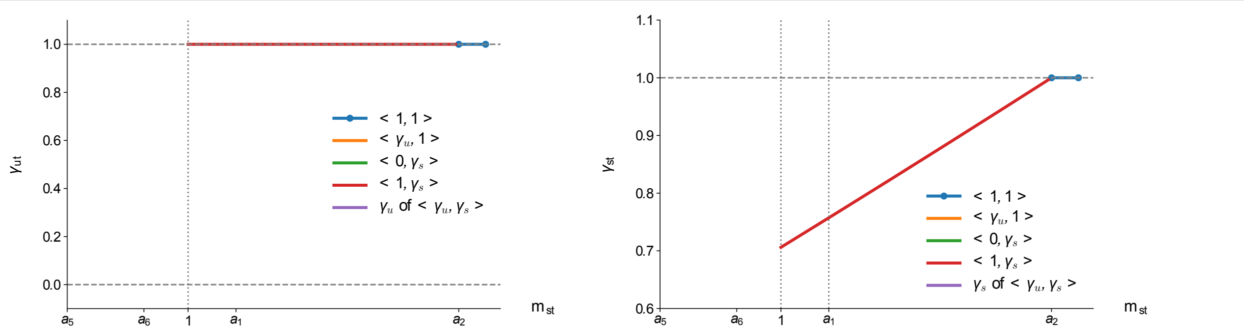
$0 \longrightarrow \underline{\delta} \longrightarrow \delta_a \longrightarrow \delta$

$\underbrace{0 \quad \cdots \quad \underline{\delta}}_{\text{Low}} \quad \underbrace{\underline{\delta} \quad \cdots \quad \delta_a}_{\text{Moderately High}} \quad \underbrace{\delta_a \quad \cdots}_{\text{Huge}}$

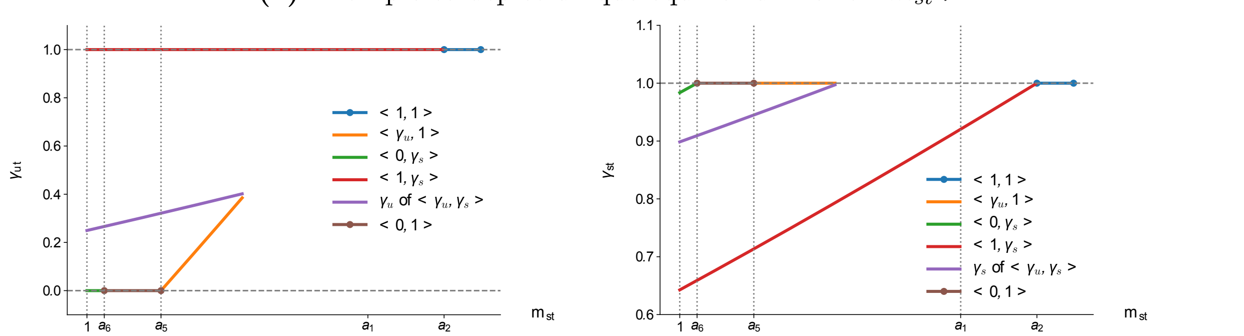
Figure 3: Characterization of the Equilibria with Behavioral Trap and Behavioral Bias

We prove this in Appendix 8.9 and Figure 3 depicts the equilibria at a glance.

Next, we consider numerical examples to show that when $\beta \geq \theta(1-\beta)$, depending on the parametric conditions, there can be unique or multiple equilibria.[19] Figure 4a shows unique equilibrium at any m_{st} when $\theta = 0.4$ and $\beta = 0.7$, and Figure 4b provides an example of multiple equilibria for the same values of θ and β. In this example, the difference in the two plots stems from differences in the values of ϕ and δ.



(a) Example to depict unique equilibrium for all $m_{st} > 1$.



(b) Example to depict multiple equilibria for some $m_{st} > 1$.

Figure 4: Example: Unique and Multiple equilibria when $\beta > \theta(1-\beta)$. [Colored Graphs]

We, now, analyze the dynamics and steady states. Here also, as the uneducated workers

[19]A numerical example for $\beta < \theta(1-\beta)$ can be found in the Supplementary Appendix.

never invest, we have multiple steady states and they can be ranked in terms of inequality. The steady state where $m_s^* = \max\{a_1(\delta), a_2(\delta)\}$, we call that the 'least unequal steady state'.

Proposition 6. Dynamics and the Steady States

1. *There is almost always a poverty trap in an economy.*

2. *Dynamics and Steady States – When the degree of child affinity is*

   a. *not low: any $m_{st} \geq \max\{a_1(\delta), a_2(\delta)\}$ is a steady state where all educated workers invest with prob. 1. Steady-state income of a skilled worker is initial income $m_s^* = m_{st}$,*

   b. *not low and $m_{st} < \max\{a_1(\delta), a_2(\delta)\}$: at least one type of workers invest with probability less than 1. The mass of educated individuals and that of skilled workers decrease over time. Skilled income increases over time and converges to some $m_s^* \geq \max\{a_1(\delta), a_2(\delta)\}$,*

   c. *low and $m_{st} > \underline{a}_s(\delta)$: the mass of skilled workers decreases over time and tends to zero, and correspondingly the income of a skilled worker tends to infinity,*

   d. *low and $m_{st} \leq \underline{a}_s(\delta)$: unique steady state is immediately reached where no one invests.*

3. *Properties of the Steady States – When the degree of child affinity is*

   a. *not low: There are multiple steady states ranked on the basis of inequality. The inequality at the least unequal steady state (weakly) increases with a decrease in child affinity,*

   b. *low: At the unique steady state all workers are unskilled and there is no inequality.*

We prove in Appendix 8.10.
The intuitions behind multiple steady states and only pure strategies being played in any such steady state are very similar to those discussed in Section 4.1.
Next, we discuss the welfare implications of the behavioral anomalies.

5 Comparison: Implications of Behavioral Anomalies

We analyze the implications of behavioral anomaly, focusing on two aspects:

(i) *Distortions due to over and under investment*:[20] We discuss the parametric conditions in which behavioral anomalies induce over and under investment. Recall, in the benchmark case, at any equilibrium when the unskilled workers invest with a positive probability, the skilled workers invest with probability one. This feature holds for the economy with behavioral trap also. But, when both behavioral trap and behavioral bias are present this may not hold. Both types of educated workers may under or over invest. Further, skilled workers may under invest due to the overinvestment of educated-unskilled workers. We call this *crowding out* of the investment.

[20] Observe, a child always prefers to be educated. We analyze from the parent's point of view and do not take her child's point of view into account.

(ii) *Poverty Trap and Inequality at the steady states*: We compare the mass of families in the poverty trap. We also compare the inequality at the 'least unequal steady states' with that at the unique steady state of the benchmark case. We observe that even when the inequalities are equal, the equilibria can be ranked in terms of *opportunities*.

5.1 Implications of Behavioral Trap Only

We begin by comparing the benchmark case with the case where only uneducated workers are imprisoned in a behavioral trap.

(i) *Distortions*: We show in the next observation that no educated worker under invests and characterize the parametric condition where educated-skilled workers overinvest.

Observation 8. *The distortions in investment decisions are as follows:*

1. *When the degree of child affinity is not low,*
   - a. *at any $m_{st} \geq 1$ the skilled workers invest with the same probability in both the cases*
   - b. *under behavioral trap educated-unskilled workers invest with a strictly higher probability than that in the benchmark case when the degree of child affinity is moderate, the state variable is higher than $\underline{b}_u(\delta)$, and the mass of uneducated workers is positive; otherwise the probabilities are equal.*
2. *When the degree of child affinity is low, all educated workers invest with the same probability as in the benchmark case.*

We prove this in Appendix 8.11.

The intuition is as follows. The existence of a behavioral trap can affect the equilibrium strategy of the educated workers only when the unskilled workers invest with a positive (but strictly less than one) probability in the benchmark case, and the mass of uneducated workers (who invest in the benchmark case but not under the behavioral trap) is positive. This is because in that case, the non-investment of unskilled workers improves the expected benefit from investment for the educated workers and that makes the educated-unskilled workers invest with a strictly higher probability. Observe, at this parametric condition, the skilled workers invest with probability one even in the benchmark case, so this improvement in benefit does not affect their investment probability.

(ii) *Poverty Trap and Inequality at the steady states*: When child affinity is not low, there is no poverty trap in the steady state of the benchmark case, whereas with behavioral trap, whenever an economy starts with a positive mass of uneducated workers, there is a poverty trap in the steady state. Further, comparing the steady-state inequalities, we find

Observation 9. *When the degree of child affinity is not low, steady-state inequality is (weakly) higher under the behavioral trap.*

We prove this in Appendix 8.12.
The intuition is as follows. When child affinity is high, in the benchmark case, all workers invest whereas, in the case of behavioral trap, all educated workers invest. Therefore, when an economy starts with all educated adults, then the steady-state income of skilled workers and hence, the steady-state inequalities in both cases are the same. In all other cases, inequality under the behavioral trap is higher. When child affinity is moderate, in the benchmark case, the steady-state income of a skilled worker is $\beta^{-(1-\phi)}\underline{b}_u(\delta)$. Under the behavioral trap, the steady-state income of a skilled worker is the same amount only at the 'least unequal steady state', at all other steady states, it is strictly higher. Hence, the observation.

5.2 Implications of Behavioral Trap & Behavioral Bias

Let us now consider the economy where educated workers are also biased. For that first, we cumulate the thresholds of child affinity defined in Section 3 and Section 4.2.

Observation 10 (S)**.** $0 < \underline{\delta} < \bar{\delta} < \delta_a$.

Like before, we first compare distortions in investment. At any m_{st}, we say that the educated-unskilled workers under (or over) invest if γ_{ut} is lower (or higher) than ρ_{ut}. Similarly, skilled workers under (or over) invest as γ_{st} is lower (or higher) than ρ_{st}.

Observation 11. *The distortions in investment decisions are as follows:*

1. *When the degree of child affinity is huge,*

   a. *at any $m_{st} \geq 1$ the educated-unskilled workers with or without behavioral bias invest with the same probability,*

   b. *skilled workers with behavioral bias under invest when the state variable is lower than $a_2(\delta)$.*

2. *When the degree of child affinity is high and the state variable is less than $\max\{a_1(\delta), a_2(\delta)\}$, then both types of educated workers may under invest.*

3. *When child affinity is moderate, then both types of educated workers may over or under invest.*

4. *When child affinity is low, unskilled workers never invest. The skilled workers over or under invest depending on $a_6(\delta)$ is lower or greater than $\bar{b}_s(\delta)$ respectively.*

This observation follows from Propositions 3 and 5.
Behavioral trap induces uneducated workers to not invest under any parametric condition. So, in the observation, we mainly focus on the distortion in educated workers' investments. Observe, there is no crowding out in economies with high or huge child affinity – as, in the benchmark case, investing with certainty is a strictly dominating strategy for each type of worker. For moderate child affinity, crowding out is possible.

We compare steady-state inequalities in the next observation which follows from Propositions 4 and 6.

Observation 12. *In presence of both behavioral bias and behavioral trap:*

1. *When the degree of child affinity is high, the inequality at the 'least unequal steady state' is equal to that at the unique steady state of the benchmark case only if the following conditions hold – (i) the economy starts with all educated workers, and (ii)* $\max\{a_1(\delta), a_2(\delta)\} \leq A\beta^{-(1-\phi)}$. *Otherwise, the former is strictly higher than the latter.*

2. *When the degree of child affinity is moderate, the inequality at the 'least unequal steady state' may be lower, equal, or higher than that at the unique steady state of the benchmark case depending on* $\beta(\underline{b}_u(\delta)/A)^{-(1-\phi)} \gtreqless \max\{a_1(\delta), a_2(\delta)\}$ *respectively.*

3. *When child affinity is low as in the benchmark there is no inequality at the steady state.*

Behavioral anomalies may create lower inequality when the educated-unskilled workers overinvest. Such a parent regret *ex post*, however, this improves the probability with which such a family becomes rich. Now, uneducated workers when imprisoned in a behavioral trap never invest, and due to behavioral bias, educated workers may under invest. For these, we may observe the mass of families in the poverty trap to be higher.

6 Discussion

We analyze the implications of behavioral anomalies in terms of distortions in investments and steady-state inequality. Behavioral trap begets extreme pessimism in uneducated parents, while behavioral biases affect the confidence levels of educated parents.

In the benchmark economy where parents' warm glow factor is low, uneducated parents do not invest in education. The addition of a behavioral trap does not affect uneducated parents investment decisions, as they were investing with zero probability even in the benchmark case. Hence, when warm glow is low, the behavioral trap does not affect the economy. With further addition of behavioral biases, it may influence the skilled workers to over or under-invest (depending on the parameters of the model – whether the threshold $a_6(\delta)$ is lower or greater than the threshold $\bar{b}_s(\delta)$). The inclusion of behavioral anomalies does not affect the steady state, as eventually, everyone becomes uneducated and unskilled in this economy.

When the warm glow factor is not low, the introduction of the behavioral trap alone affects the investment decision of *only* the unskilled workers. For moderate warm glow and the income of the skilled parents higher than the threshold $\underline{b}_u(\delta)$, i.e. $m_{st} > \underline{b}_u(\delta)$, the educated-unskilled worker over invests and uneducated workers under invest. Thus, when uneducated workers are extremely pessimistic, the educated-unskilled parents may anticipate fewer future educated adults and, hence, invest more than the benchmark case. Skilled workers continue to invest more than educated-unskilled workers who invest more than an uneducated worker. While the behavioral trap does not bring change to the societal structure, it makes some groups invest at the expense of other groups. Hence, the steady-state inequality is weakly higher under the behavioral trap.

Finally, when the warm glow factor is not low and the economy witnesses behavioral trap *as well as behavioral biases*, we find some interesting results. For huge warm glow factor and the income of the skilled parents higher than the threshold $a_2(\delta)$, i.e. $m_{st} < a_2(\delta)$, skilled workers under-invest and also invest less than the educated-unskilled worker. This result is driven by the skilled workers' overconfidence. The steady-state inequality is higher, except in some extreme conditions[21]. For a moderate warm glow, educated workers may over or under invest. The steady-state inequality could also be higher or lower than the benchmark. Here, the strength of pessimism and optimism plays a role that makes it possible for the economy to have more educated workers and less income inequality relative to the benchmark levels.

7 Conclusion

Homo Sapiens, unlike *Homo Economicus*, does get affected by experiences of her own or that of individuals she perceives as similar. We provide a behavioral explanation of inequality where individuals form beliefs about the efficacy of their children based on their own experiences. Incorporating psychological beliefs in a dynamic macroeconomic framework is the modelling novelty of our paper. Except for their beliefs about their children's job prospects, the parents are otherwise rational. An overconfident parent underestimates the market-driven skilled incomes and the opposite is true for under confident parents. While all agents have economic resources to invest, some adults do not invest as they perceive the returns not to be sufficiently high, thus capturing the effect of cognitive rather than economic limitations on investment.

The paper highlights that a community's personal beliefs and biases can have societal effects. A behavioral trap would cause a poverty trap and limits intergenerational mobility. Abstinence of the uneducated persons from investments induces higher investment in other people. These cross-sectional effects of beliefs and biases are missing in homogeneous agent models and can be useful in the study of economic disparities. Further, while traditional theory can explain poverty trap through high fixed costs of education, it misses out behavioral macroeconomic effects such as multiple equilibria, multiple paths to steady state, etc. Thus, behavioral anomalies have distinct outcomes in the short run as well as the long run.

This paper highlights these effects and we hope it directs the focus of public policies towards behavioral impediments. Behavioral policies such as mentoring programs, improving social interaction, etc. have to become a part of a policy maker's toolbox. As we expand our understanding of the behavioral sciences, we see that the point of intervention is at the locality not at the national level and our paper aligns with this finding.

[21] When warm glow is high, the inequality at the 'least unequal steady state' is equal to that at the unique steady state of the benchmark case only if the following conditions hold *(i) the economy starts with all educated workers, and (ii)* $\max\{a_1(\delta), a_2(\delta)\} \leq A\beta^{-(1-\phi)}$. Otherwise, the former is strictly higher than the latter.

8 Appendix

8.1 Formal Expression for the Definition 2

At $\underline{b}_s(\delta)$ a *skilled* worker is indifferent between investing and not investing, when no other worker invests. Thus, $N_{et+1} = 0$, $L_{st+1} = 0$, and $m_{st+1} \to \infty$ and it must be that

$$\delta \left[\frac{[\beta m_{st+1} + (1-\beta)]^\sigma}{\sigma} - \frac{1}{\sigma} \right] = \frac{\underline{b}_s{}^\sigma}{\sigma} - \frac{(\underline{b}_s - \bar{s})^\sigma}{\sigma} \quad \Rightarrow \quad \underline{b}_s(\delta) : \quad \underline{b}_s^\sigma - (\underline{b}_s - \bar{s})^\sigma + \delta = 0.$$

The formal expression for $\bar{b}_s(\delta)$ and $\underline{b}_u(\delta)$ are also found similarly (see Online Appendix).

8.2 Proof of Proposition 1

1. The uniqueness follows directly from Lemma 1 and equation (3).

2. Consider any $\delta > \bar{\delta}$, from (3) it can be seen that $\gamma_{jt} = 1$ is the strictly dominating strategy for j^{th} type of worker, where $j \in \{u, s\}$. When $\delta = \bar{\delta}$, similarly, it can be seen that $\gamma_{st} = 1$ is the strictly dominating strategy for a skilled worker and $\gamma_{ut} = 1$ is weakly dominating strategy for an unskilled worker. Now, observe again from (3), if a positive mass of unskilled worker plays any strategy other than $\gamma_{ut} = 1$, then such an unskilled worker has an incentive to deviate and play $\gamma_{ut} = 1$. Therefore, $\langle 1, 1 \rangle$ is a unique equilibrium $\forall\ \delta \geq \bar{\delta}$.

3. Consider $\delta \in [\underline{\delta}, \bar{\delta})$. From Lemma 1, we have that in any equilibrium where $\lambda_{ut} > 0$, $\lambda_{st} = 1$. Now, from (3), it can bee seen that for any $m_{st} \geq 1$, at $\langle 1, 1 \rangle$, the benefit from investment of an unskilled worker is strictly lower than her cost of investment. So, she has an incentive to deviate. Hence, $\langle 1, 1 \rangle$ cannot be an equilibrium.

   3.a., 3.b., and 3.c. follow directly from the definitions of $\underline{b}_u(\delta)$, $\bar{b}_s(\delta)$ and from Lemma 2 that $\underline{b}_s(\delta) \leq 1$. Further, from the definition of mixed strategy, when $\lambda_{jt} \in (0, 1)$ then (3) must bind for the j^{th} type of workers.

4. Consider $\delta < \underline{\delta}$. It can be seen from (3) that at any $\langle \lambda_{ut}, 1 \rangle$ where $\lambda_{ut} > 0$, the benefit from investment of an unskilled worker is strictly lower than her cost of investment. Hence, there does not exist any equilibrium where $\lambda_{ut} > 0$.

   4.a., 4.b. and 4.c. follow from the definitions of $\bar{b}_s(\delta)$ and $\underline{b}_s(\delta)$. It is also obvious that when $\lambda_{st} \in (0, 1)$, (3) must bind. □

8.3 Proof of Proposition 2

1.a. In this case, from Proposition 1 (subpoint 2.), we have that all parents invest with certainty. So, the economy immediately reaches a steady state, the mass of skilled worker is $L_s^* = \beta \cdot 1$ and the income of a skilled worker is $A(L_s^*)^{-(1-\phi)} = A\beta^{-(1-\phi)}$.

At the steady state, the prob. that an adult works as a skilled worker

$$=\beta \cdot [\lambda_s^* \cdot \text{ prob that her parent was skilled} + \lambda_u^* \cdot \text{ prob that her parent unskilled}] = \beta$$

where the last equality is coming from the fact that $\lambda_s^* = \lambda_u^* = 1$.

1.b. Observe $\langle 0, 1\rangle$ or $\langle 0, (0, 1)\rangle$ cannot be the equilibrium strategy at any steady state because in those cases, the mass of skilled workers decreases over time. So, if an economy starts with a mass of skilled workers higher than $\beta \left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}}$, then only skilled workers invest (as m_{st} is lower than $\underline{b}_u(\delta)$). The mass of skilled workers decreases and their income increases over time and reaches $\underline{b}_u(\delta)$ when the unskilled workers start to invest.

We ask at what λ_u^* is the economy at the steady state? Consider the incentive constraint of an unskilled worker when all other unskilled workers invest with probability λ_{ut} and skilled workers invest with certainty. At the steady state, $L_{st+1} = L_{st} = L_s^*$ which implies

$$\beta(L_{st} + (1 - L_{st})\lambda_{ut}) = L_{st+1} = \left[\frac{1}{\beta A}\left[\left[\frac{1 + \delta - (1 - \bar{s})^\sigma}{\delta}\right]^{\frac{1}{\sigma}} - (1 - \beta)\right]\right]^{-\frac{1}{1-\phi}} \equiv \beta\left(\frac{\underline{b}_u(\delta)}{A}\right)^{-\frac{1}{1-\phi}}$$

$$\Rightarrow \quad \lambda_{ut} = \frac{\left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}} - L_{st}}{1 - L_{st}} = \frac{\left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}} - L_s^*}{1 - L_s^*} = \lambda_u^*$$

where the first equality is coming from the fact that the mass of educated individuals at $t + 1$ would be $L_{st} + \lambda_{ut}(1 - L_{st})$, and β fraction of them would work as a skilled worker, the second equality is coming from the investment decision of an unskilled worker (and getting the value of L_{st+1} from that):

$$\delta\left[\frac{[\beta \cdot AL_{st+1}^{-(1-\phi)} + (1 - \beta)]^\sigma}{\sigma} - \frac{1}{\sigma}\right] = \frac{1}{\sigma} - \frac{(1 - \bar{s})^\sigma}{\sigma}$$

At the steady state, mass of skilled worker $L_s^* \equiv \beta\left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}}$, wage of a skilled worker $m_s^* \equiv \beta^{-(1-\phi)}\underline{b}_u(\delta)$ and $\lambda_u^* \equiv \dfrac{\left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}} - L_s^*}{1 - L_s^*}$.

Observe $\lambda_u^* \in (0, 1)$: (i) $\lambda_u^* < 1$ as $\underline{b}_u(\delta) = A$ at $\delta = \bar{\delta}$, and $\underline{b}_u(\delta)$ is decreasing in δ, so for $\delta < \bar{\delta}$, $\underline{b}_u(\delta) > A$, which implies $\left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}} < 1$ (ii) $\lambda_u^* > 0$ as $\left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}} - L_s^* = (1 - \beta)\left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}} > 0$.

At the steady state, the prob. that an adult works as a skilled worker

$$=\beta \cdot [\lambda_s^* \cdot \text{ prob that her parent was skilled} + \lambda_u^* \cdot \text{ prob that her parent unskilled}] < \beta$$

where the last inequality comes from $\lambda_u^* < 1$. Also, observe differentiating λ_u^* with respect to $\underline{b}_u(\delta)$, we get that λ_u^* strictly increases with decrease in $\underline{b}_u(\delta)$, and $\underline{b}_u(\delta)$ strictly decreases with an increase in δ, i.e., as δ decreases λ_u^* strictly decreases and $\lambda_s^* = 1$. Hence the result.

1.c. When the degree of child affinity is high, the steady-state income of a skilled worker is $A\beta^{-(1-\phi)}$ and that of an unskilled worker is 1. So, the inequality is the same $\forall\ \delta \geq \bar{\delta}$.

When child affinity is moderate, the steady-state income of a skilled worker is $\beta^{-(1-\phi)}\underline{b}_u(\delta)$ and that of an unskilled worker is 1. Now, $\underline{b}_u(\delta)$ strictly decreases with increase in δ. So, the difference between the income of a skilled worker and that of an unskilled worker decreases with increase in δ. Hence, the result.

2.a. From Proposition 1 (subpoint 4. a. and b.), we have that when $\delta < \underline{\delta}$, then no unskilled workers invest at any m_{st}. Moreover, when $m_{st} > \underline{b}_s(\delta)$, then skilled workers invest with a positive probability. So, the mass of educated workers and hence, the mass of skilled workers decrease over time and converge to zero, whereas the income of a skilled worker increases over time and tends to infinity.

2.b. From Proposition 1 (subpoint 4. c.), for low child affinity and $m_{st} \leq \underline{b}_s(\delta)$, no parents invest. So, the economy is in a steady state where no parent invests and all workers are unskilled. □

8.4 Formal Statement of Definition 3

From defn. of $\bar{b}_u(\delta)$ and Lemma 1 (subpoint 1.), at $\bar{b}_u(\delta)$, an educated-unskilled worker is indifferent between investing and not investing, when all other educated workers invest with certainty. So, the mass of educated worker at $t+1$ remains N_{et}, $m_{st+1} = m_{st}$ and

$$\bar{b}_u(\delta): \quad \delta\left[\frac{[\beta\bar{b}_u(\delta)+1-\beta]^{\sigma}}{\sigma} - \frac{1}{\sigma}\right] = \frac{1}{\sigma} - \frac{(1-\bar{s})^{\sigma}}{\sigma}. \tag{A.1}$$

8.5 Proof of Proposition 3

1. The uniqueness follows directly from Observations 3, 4 and equation (4).

2. See Observation 3.

3. Consider (4), $\rho_{ut} = 1$ if and only if

$$\delta\left[\frac{\left[\beta^{\phi}A[(1-\beta)N_{et}+\beta N_{et}]^{-(1-\phi)}+1-\beta\right]^{\sigma}}{\sigma} - \frac{1}{\sigma}\right] \geq \frac{1}{\sigma} - \frac{(1-\bar{s})^{\sigma}}{\sigma}.$$

   The benefit, i.e. the L.H.S. increases with decrease in N_{et} whose max. value is 1. The L.H.S. increases with δ. So, to prove the claim, it is sufficient to show that L.H.S. is no less than R.H.S. at $N_{et} = 1$ and $\delta = \bar{\delta}$: $\bar{\delta}\left[\frac{\left[\beta^{\phi}A+1-\beta\right]^{\sigma}}{\sigma} - \frac{1}{\sigma}\right] \geq \frac{1}{\sigma} - \frac{(1-\bar{s})^{\sigma}}{\sigma}$.

   Now, at $\bar{\delta}$, observe L.H.S. is equal to R.H.S. Hence, for $\delta \geq \bar{\delta}$, we get $\rho_{ut} = 1 \;\forall\, N_{et} \in [0,1]$.

4. a., b., c. and d. follow from the definitions of $\bar{b}_u(\delta)$, $\underline{b}_u(\delta)$, $\bar{b}_s(\delta)$, $\underline{b}_s(\delta)$, and Lemma 2.

5. a., b. and c. also follow from the definitions of $\bar{b}_u(\delta)$, $\bar{b}_s(\delta)$, $\underline{b}_s(\delta)$, and Lemma 2. □

8.6 Proof of Proposition 4

1. Uneducated workers never invest, so if an economy starts with any positive mass of uneducated workers then there will always be a poverty trap.
   When $\delta < \bar{\delta}$, even if the economy starts with all educated workers, educated-unskilled workers invest with prob. less than 1. So, there is a poverty trap.
   Only when the previous two scenarios do not hold, i.e. the economy does not have any uneducated workers *and* the degree of child affinity is high, there is no poverty trap.

2.a. From Proposition 3., when $\delta \geq \bar{\delta}$, there is a unique steady state where all educated workers invest. If the mass of uneducated workers is zero then the steady-state income of a skilled worker would be $\bar{b}_u(\bar{\delta})$, otherwise it would be strictly higher than $\bar{b}_u(\bar{\delta})$. All educated workers invest always. Hence, the result.

2.b. From Proposition 3., when child affinity is moderate and $m_{st} \geq \bar{b}_u(\delta)$, all educated workers invest with certainty. So, the mass of educated workers, and hence the mass of skilled workers and their income remain constant over time. Therefore, any $m_{st} \geq \bar{b}_u(\delta)$ is a steady state.

2.c. From Proposition 3., when $m_{st} < \bar{b}_u(\delta)$, educated-unskilled workers invest with prob. less than 1. So, the mass of educated, and that of skilled workers, decrease over time. This implies the income of a skilled worker increases over time. This happens till $m_{st} \geq \bar{b}_u(\delta)$. Then, we are in the region described in Part 2.b., and reach a steady state, $m_s^* \geq \bar{b}_u(\delta)$.

2.d. The proof is similar to that of Proposition 2 (subpoint 2).

2.e. The proof is same as that for Proposition 2 (subpoint 2).

3.a. The multiplicity of steady states follows from the above. Steady-state inequality increases with increase in income of a skilled worker (as an unskilled worker's income is 1).
   For high child affinity, the least unequal steady state is at $m_s^* = \bar{b}_u(\bar{\delta}) = A\beta^{-(1-\phi)}$. For moderate affinity, it is $\bar{b}_u(\delta)$ which we noted in Observation 5, is decreasing in δ. Hence the claim.

3.b. For low child affinity, the unskilled workers do not invest, as in the benchmark. Hence, this proof is very similar to the proof of Proposition 2 (subpoint 2.). □

8.7 Formal Expressions for Definition 5

Given Lemma 3, when $\gamma_{st} = 0$, γ_{ut} is also zero. Hence, for a given degree of child affinity δ, at $\underline{a}_s(\delta)$ a skilled worker is indifferent between investing and not investing, when no other worker invests. So, from (6) we have

$$\underline{a}_s(\delta): \quad \underline{a}_s^{\sigma} - (\underline{a}_s - \bar{s})^{\sigma} + \delta = 0. \tag{A.2}$$

The formal expression for the other thresholds can be found similarly (see Online Appendix).

8.8 Boundary Conditions for Equilibrium Probabilities

Condition $\underline{\Gamma}_s$: $\underline{\gamma}_s(m_{st})$ captures the optimal response of a skilled worker when $\gamma_{ut} = 1$. Thus, for $m_{st} \geq a_4(\delta)$, the formal expression for $\underline{\gamma}_s(m_{st})$ is

$$\begin{aligned}
&\delta\left[\frac{\Big[[1-\theta(1-\beta)]\cdot\big[(1-\beta)+[1-\theta(1-\beta)]\underline{\gamma}_s(m_{st})\big]^{-(1-\phi)}m_{st}+\theta(1-\beta)\Big]^{\sigma}}{\sigma}-\frac{1}{\sigma}\right]\\
&\geq\frac{m_{st}^{\sigma}}{\sigma}-\frac{(m_{st}-\bar{s})^{\sigma}}{\sigma}
\end{aligned}$$

the inequality binds only when $m_{st} \in [a_4(\delta), a_2(\delta)]$. Clearly from the definition of $a_4(\delta)$ and Observation 7 we can find that $\underline{\gamma}_s(m_{st}) = 0$ for $m_{st} \leq a_4(\delta)$.
Similarly, including the definition of $a_2(\delta)$, we get

$$\underline{\gamma}_s(m_{st})\begin{cases} = 0 & \forall\ m_{st} \leq a_4(\delta)\\ \in (0,1) & \forall\ m_{st} \in (a_4(\delta), a_2(\delta))\\ = 1 & \forall\ m_{st} \geq a_2(\delta).\end{cases}$$

As γ_{ut} can atmost be one, now it is clear that at any equilibrium $\langle\gamma_{ut}, \gamma_{st}\rangle$, $\gamma_{st}(m_{st}) \geq \underline{\gamma}_s(m_{st})$.

Finally we show that if $m_{st} \in (a_4(\delta), a_2(\delta))$ then $\underline{\gamma}_s'(m_{st}) > 0$.
Suppose not. $a_4(\delta) < m_{st}^1 < m_{st}^2 < a_2(\delta)$ and $1 > \underline{\gamma}_s^1 \equiv \underline{\gamma}_s(m_{st}^1) \geq \underline{\gamma}_s(m_{st}^2) \equiv \underline{\gamma}_s^2 > 0$. Then, we must have

$$\begin{aligned}
&\frac{(m_{st}^2)^{\sigma}}{\sigma}-\frac{(m_{st}^2-\bar{s})^{\sigma}}{\sigma}\\
=&\delta\left[\frac{\Big[[1-\theta(1-\beta)]\cdot\big[(1-\beta)+[1-\theta(1-\beta)]\underline{\gamma}_s^2\big]^{-(1-\phi)}m_{st}^2+\theta(1-\beta)\Big]^{\sigma}}{\sigma}-\frac{1}{\sigma}\right]\\
>&\delta\left[\frac{\Big[[1-\theta(1-\beta)]\cdot\big[(1-\beta)+[1-\theta(1-\beta)]\underline{\gamma}_s^1\big]^{-(1-\phi)}m_{st}^1+\theta(1-\beta)\Big]^{\sigma}}{\sigma}-\frac{1}{\sigma}\right]\\
=&\frac{(m_{st}^1)^{\sigma}}{\sigma}-\frac{(m_{st}^1-\bar{s})^{\sigma}}{\sigma}
\end{aligned}$$

where the two equalities come from the formal expressions of $\underline{\gamma}_s^1$ and $\underline{\gamma}_s^2$, and the inequality is from $m_{st}^1 < m_{st}^2$ and $\underline{\gamma}_s^1 \geq \underline{\gamma}_s^2$.
But it is not possible as $\frac{(m_{st}^2)^{\sigma}}{\sigma}-\frac{(m_{st}^2-\bar{s})^{\sigma}}{\sigma} < \frac{(m_{st}^1)^{\sigma}}{\sigma}-\frac{(m_{st}^1-\bar{s})^{\sigma}}{\sigma}$.
Therefore, when $m_{st} \in (a_4(\delta), a_2(\delta))$ we have $\underline{\gamma}_s(m_{st})$ is increasing in m_{st}.

Condition $\bar{\Gamma}_s$: $\underline{\gamma}_s(m_{st})$ captures the optimal response of a skilled worker when $\gamma_{ut} = 0$.

Thus, for $m_{st} \geq \underline{a}_s(\delta)$, the formal expression for $\underline{\gamma}_s(m_{st})$ is

$$\delta \left[\frac{\left[[1-\theta(1-\beta)]^{\phi} [\underline{\gamma}_s(m_{st})]^{-(1-\phi)} m_{st} + \theta(1-\beta) \right]^{\sigma}}{\sigma} - \frac{1}{\sigma} \right] \geq \frac{m_{st}^{\sigma}}{\sigma} - \frac{(m_{st} - \bar{s})^{\sigma}}{\sigma}$$

where the above inequality binds when $m_{st} \in [\underline{a}_s(\delta), a_6(\delta))]$. Clearly from the definitions of $\underline{a}_s(\delta)$, $a_6(\delta)$ and Observation 7 we show that

$$\bar{\gamma}_s(m_{st}) \begin{cases} = 0 & \forall\ m_{st} \leq \underline{a}_s(\delta) \\ \in (0,1) & \forall\ m_{st} \in (\underline{a}_s(\delta), a_6(\delta)) \\ = 1 & \forall\ m_{st} \geq a_6(\delta). \end{cases}$$

Following similar arguments as above, it can be shown that at any equilibrium $\langle \gamma_{ut}, \gamma_{st} \rangle$, $\gamma_{st}(m_{st})$ is bounded above by $\bar{\gamma}_s(m_{st})$, and $\bar{\gamma}_s(m_{st})$ is strictly increasing in $m_{st} \in [\underline{a}_s(\delta), a_6(\delta))$. □

8.9 Characterization of Equilibria

We now introduce an observation and a lemma which would be used to prove Proposition 5 in Appendix C.9.1.

Observation 8.1. *Suppose at any $m_{st} \geq 1$, there are two equilibria $\langle \gamma_{ut}, \gamma_{st} \rangle$ and $\langle \tilde{\gamma}_{ut}, \tilde{\gamma}_{st} \rangle$. If $\gamma_{jt} < \tilde{\gamma}_{jt}$ then $\tilde{\gamma}_{kt} \leq \gamma_{kt}$ where $j, k \in \{u, s\}$ and $j \neq k$. The latter inequality binds only when $\tilde{\gamma}_{kt} = 1$.*

Proof. Immediate from investment decisions of both types of parents given by (5) and (6). □

Lemma 8.1. *Suppose $\delta \in (\underline{\delta}, \delta_a]$.*

1. *Suppose $\theta(1-\beta) \neq \beta$, then at any $m_{st} \in [1, min\{a_1, a_2\})$, there can be at most one equilibrium where both types of workers play mixed strategies.*

2. *At any $m_{st} \in [1, min\{a_1(\delta), a_2(\delta)\})$, there can be multiple equilibria only if $\beta \geq \theta(1-\beta)$.*

3. *Suppose $\beta < \theta(1-\beta)$. Let $\langle \gamma_{ut}, \gamma_{st} \rangle$ be an equilibrium at any $m_{st} \in [1, min\{a_1(\delta), a_2(\delta)\})$. If $\gamma_{st} \geq \gamma_{ut}$, then at all $\tilde{m}_{st} \in (m_{st}, max\{a_1(\delta), a_2(\delta)\})$ $\tilde{\gamma}_{st} > \tilde{\gamma}_{ut}$ where $\langle \tilde{\gamma}_{ut}, \tilde{\gamma}_{st} \rangle$ is an equilibrium at $\tilde{m}_{st}$.*

Proof. $\delta \in (\underline{\delta}, \delta_a]$. Then from Lemma 2 and Lemma 4, we have $\underline{a}_s(\delta) < 1$.

1. Suppose not. $\theta(1-\beta) \neq \beta$ and at some $m_{st} \in [1, min\{a_1, a_2\})$, there exist two equilibria $\langle \gamma_{ut}, \gamma_{st} \rangle$ and $\langle \tilde{\gamma}_{ut}, \tilde{\gamma}_{st} \rangle$ where both types of workers play mixed strategies, i.e. $0 < \gamma_{ut} \neq \tilde{\gamma}_{ut} < 1$ and $0 < \gamma_{st} \neq \tilde{\gamma}_{st} < 1$.

   Then from the decision of educated-unskilled workers, given by (5), we must have

$$\theta(1-\beta)\gamma_{ut} + \beta\gamma_{st} = \theta(1-\beta)\tilde{\gamma}_{ut} + \beta\tilde{\gamma}_{st}. \tag{A.3}$$

And, from the decision of skilled workers, given by (6), we must have

$$(1-\beta)\gamma_{ut} + [1-\theta(1-\beta)]\gamma_{st} = (1-\beta)\tilde{\gamma}_{ut} + [1-\theta(1-\beta)]\tilde{\gamma}_{st}. \tag{A.4}$$

As $\theta(1-\beta) \neq \beta$, from (A.3) and (A.4), we have $\gamma_{ut} = \tilde{\gamma}_{ut}$ and $\gamma_{st} = \tilde{\gamma}_{st}$ – a contradiction.

2. Given definitions of $a_1(\delta)$, $a_2(\delta)$, Lemma 3, and Observation 8.1, at any $m_{st} \in [1, min\{a_1(\delta), a_2(\delta)\})$, if there are two equilibria $\langle \gamma_{ut}, \gamma_{st} \rangle$ and $\langle \tilde{\gamma}_{ut}, \tilde{\gamma}_{st} \rangle$, then we must have

$$1 \geq \gamma_{ut} > \tilde{\gamma}_{ut} \geq 0 \quad \text{and } 1 \geq \tilde{\gamma}_{st} \geq \gamma_{st} > 0,$$

where at least one of the educated parents does not invest with certainty (as $m_{st} < min\{a_1(\delta), a_2(\delta)\}$), and skilled workers must invest with a positive probability (due to Lemma 3, and that $m_{st} \geq 1 > \underline{a}_s(\delta)$). Using (5) and (6), we get

$$\begin{aligned} & \theta(1-\beta)\tilde{\gamma}_{ut} + \beta\tilde{\gamma}_{st} \geq \theta(1-\beta)\gamma_{ut} + \beta\gamma_{st} \\ \Rightarrow \quad & \beta[\tilde{\gamma}_{st} - \gamma_{st}] \geq \theta(1-\beta)[\gamma_{ut} - \tilde{\gamma}_{ut}] \end{aligned} \tag{A.5}$$

$$\begin{aligned} & \text{and, } (1-\beta)\gamma_{ut} + [1-\theta(1-\beta)]\gamma_{st} \geq (1-\beta)\tilde{\gamma}_{ut} + [1-\theta(1-\beta)]\tilde{\gamma}_{st} \\ \Rightarrow \quad & (1-\beta)[\gamma_{ut} - \tilde{\gamma}_{ut}] \geq [1-\theta(1-\beta)][\tilde{\gamma}_{st} - \gamma_{st}]. \end{aligned} \tag{A.6}$$

Both conditions (A.5) and (A.6) hold, i.e. the necessary condition for the coexistence of $\langle \gamma_{ut}, \gamma_{st} \rangle$ and $\langle \tilde{\gamma}_{ut}, \tilde{\gamma}_{st} \rangle$ is $\beta \geq \theta[1-\theta(1-\beta)] \quad \Rightarrow \quad \beta \geq \theta(1-\beta)$.

3. Suppose not. $\gamma_{st} \geq \gamma_{ut}$ and $\exists\ \tilde{m}_{st} > m_{st}$ such that $\tilde{\gamma}_{ut} \geq \tilde{\gamma}_{st}$.

   First observe from the previous claim that for $\beta < \theta(1-\beta)$, at any m_{st} there will always be a unique equilibrium $\langle \gamma_{ut}, \gamma_{st} \rangle$. Second, m_{st} less than less than $\min\{a_1(\delta), a_2(\delta)\}$ and $\tilde{m}_{st}$ less than $\max\{a_1(\delta), a_2(\delta)\}$ imply $\gamma_{ut} < 1$ and $\tilde{\gamma}_{st} < 1$.

   Now from the investment decision of educated-unskilled workers, given by (5), we have

$$\frac{\theta\gamma_{ut}(1-\beta) + \beta\gamma_{st}}{\theta\tilde{\gamma}_{ut}(1-\beta) + \beta\tilde{\gamma}_{st}} \geq \left[\frac{\tilde{m}_{st}}{m_{st}}\right]^{-\frac{1}{1-\phi}}$$

   And from the investment decision of skilled workers, given by (6), we have

$$\frac{(1-\beta)\gamma_{ut} + [1-\theta(1-\beta)]\gamma_{st}}{(1-\beta)\tilde{\gamma}_{ut} + [1-\theta(1-\beta)]\tilde{\gamma}_{st}} < \left[\frac{\tilde{m}_{st}}{m_{st}}\right]^{-\frac{1}{1-\phi}}$$

   From these two conditions we get

$$\begin{aligned} & \frac{\theta\gamma_{ut}(1-\beta) + \beta\gamma_{st}}{\theta\tilde{\gamma}_{ut}(1-\beta) + \beta\tilde{\gamma}_{st}} > \frac{(1-\beta)\gamma_{ut} + [1-\theta(1-\beta)]\gamma_{st}}{(1-\beta)\tilde{\gamma}_{ut} + [1-\theta(1-\beta)]\tilde{\gamma}_{st}} \\ \Rightarrow \quad & [\tilde{\gamma}_{ut}\gamma_{st} - \gamma_{ut}\tilde{\gamma}_{st}][\beta - \theta[1-\theta(1-\beta)]] > 0 \\ \Rightarrow \quad & \beta - \theta[1-\theta(1-\beta)] > 0 \quad \Rightarrow \quad \beta > \theta(1-\beta). \end{aligned}$$

   the second last line follows from $\gamma_{st} > \gamma_{ut}$ and $\tilde{\gamma}_{ut} \geq \tilde{\gamma}_{st}$. A contradiction as $\beta < \theta(1-\beta)$.

□

C.9.1 Proof of Proposition 5

1. As $\eta = 0$, this is trivial.

2. $\delta > \delta_a$, so from Lemma 4. (subpoint 3.b.), we know $a_1(\delta) < 1$. So, by the definition of $a_1(\delta)$ $\gamma_{ut} = 1 \; \forall \; m_{st} \geq 1$.

   Now, from Lemma 4. (subpoint 4.), we know $a_4(\delta) < 1$. So, due to Condition $\underline{\Gamma}_s$, γ_{st} must be equal to $\underline{\gamma}_s(m_{st}) \; \forall \; m_{st} \geq 1$, and $\underline{\gamma}_s(m_{st}) > 0$ in this range.

   That the equilibrium $\langle 1, \underline{\gamma}_s(m_{st}) \rangle$ is unique is now trivial.

3.a. $m_{st} \geq min\{a_1(\delta), a_2(\delta)\}$.

   If $a_2(\delta) \geq a_1(\delta)$, by the definition $a_1(\delta)$, in this range of m_{st}, we have $\gamma_{ut} = 1 \; \forall \; \gamma_{st} \in [0, 1]$.

   Hence, due to Condition $\underline{\Gamma}_s$, γ_{st} must be equal to $\underline{\gamma}_s(m_{st})$ and that the equilibrium is unique is now immediate.

   If $a_2(\delta) < a_1(\delta)$, by the definition $a_2(\delta)$, in this range of m_{st}, we must have $\gamma_{st} = 1$.

   That the equilibrium is unique is now evident.

3.b. We have already stated Conditions $\underline{\Gamma}_s$, $\bar{\Gamma}_s$, $\underline{\Gamma}_u$ and $\bar{\Gamma}_u$ must be satisfied whenever possible.

   We now show if $\beta < \theta(1 - \beta)$ and $a_1 < a_2$, then $\gamma_{st} < \gamma_{ut}$. Suppose not and there exists $m_{st} \in [1, \min\{a_1(\delta), a_2(\delta)\})$ at which $\gamma_{st} \geq \gamma_{ut}$. It follows from Lemma C.9.1 (subpoint 3.) that at all $\tilde{m}_{st} \in (m_{st}, a_2(\delta))$ we will have $\tilde{\gamma}_{st} > \tilde{\gamma}_{ut}$.

   Now consider any $\tilde{m}_{st} \in (a_1(\delta), a_2(\delta))$, we know from the definitions of $a_1(\delta)$ and $a_2(\delta)$ that $\tilde{\gamma}_{ut} = 1$ and $\tilde{\gamma}_{st} < 1$, i.e. $\tilde{\gamma}_{st} < \tilde{\gamma}_{ut}$, which violates the above claim. Hence, we have proved by contradiction that for all $m_{st} \in [1, \min\{a_1(\delta), a_2(\delta)\})$ we have $\gamma_s < \gamma_u$.

   The rest directly follows from Lemma 8.1.

4. From Lemma 4 (subpoint 3.a.), we have $a_1(\delta)$, $a_3(\delta), a_5(\delta)$ tend to infinity for $\delta < \underline{\delta}$. Hence, the educated-unskilled workers do not invest: $\gamma_{ut} = 0 \; \forall \; \gamma_{st} \in [0, 1]$. Next, from Condition $\bar{\Gamma}_s$, we have for any m_{st}, $\gamma_{st} = \bar{\gamma}_s(m_{st})$. That the equilibrium is unique is immediate now. □

8.10 Proof of Proposition 6

1. Suppose an economy with degree of child affinity not low starts with all educated adults, then $m_{st} = A\beta^{-(1-\phi)}$. So, if $A\beta^{-(1-\phi)} \geq \max\{a_1(\delta), a_2(\delta)\}$, then all parents would invest at all t and there would be no poverty trap.

   Suppose an economy with degree of child affinity not low, but $A\beta^{-(1-\phi)} < \max\{a_1(\delta), a_2(\delta)\}$ or the economy starts with a positive mass of uneducated adults. Now, there will be a positive mass of uneducated workers from $t = 1$ onwards. We have seen that uneducated workers never invest and education is necessary for getting a skilled job. Hence, the mass of families which never become rich is positive.

When the degree of child affinity is low, then no unskilled worker invests, so there would be a poverty trap in the economy.

Thus, in any economy, there exist a poverty trap almost always.

3.a. We show that the inequality at the least unequal steady state (weakly) increases with a decrease in child affinity. Note that both $a_1(\delta)$ and $a_2(\delta)$ are decreasing in δ, so the $\max\{a_1(\delta), a_2(\delta)\}$ is also decreasing in δ. We argue that there exists a $\hat{\delta}$ such that $\max\{a_1(\hat{\delta}), a_2(\hat{\delta})\} = A\beta^{-(1-\phi)}$. As the maximum value of the state variable is $A\beta^{-(1-\phi)}$, so for $\delta \geq \hat{\delta}$, all educated workers invest with probability one – the steady-state skilled income is $A\beta^{-(1-\phi)}$. So, the inequality at the 'least unequal steady state' remains constant for all $\delta \geq \hat{\delta}$. And, for $\delta < \hat{\delta}$, the inequality at the 'least unequal steady state' strictly increases with a decrease in child affinity.

Such a $\hat{\delta}$ exists $\forall\ \theta > 0$ as the benefit from investment for both types of parents increase with δ and δ is not bounded above.

The rest of the proof is very similar to the proof of Proposition 4, so we skip it. □

8.11 Proof of Observation 8

1.a. It is obvious as $\bar{b}_s(\delta)$ and $\underline{b}_s(\delta)$ are the same in the benchmark case and in the case with behavioral trap. Moreover, at any m_{st}, when $\rho_{ut} > 0$, then $\rho_{st} = \lambda_{st} = 1$.

1.b. We first show when $\delta \in (\underline{\delta}, \bar{\delta})$, the mass of uneducated workers is positive and $m_{st} > \underline{b}_u(\delta)$ then $\rho_{ut} > \lambda_{ut} > 0$.

Propositions 1 and 5 imply for $m_{st} \in (\underline{b}_u(\delta), \bar{b}_u(\delta))$, we have $\rho_{st} = \lambda_{st} = 1$ and $\rho_{ut}, \lambda_{ut} \in (0, 1)$. Using these probabilities of investment, from (3) and (4) we can derive

$$\rho_{ut}(1-\beta)N_{et} = \lambda_{ut}[(1-\beta)N_{et} + (1-N_{et})] > \lambda_{ut}(1-\beta)N_{et} \Rightarrow \quad \rho_{ut} > \lambda_{ut}.$$

where we have used the fact that $1 - N_{et} > 0$ and $\lambda_{ut} > 0$.

Again from Propositions 1 and 5 for $m_{st} \geq \bar{b}_u(\delta)$: $\rho_{st} = \lambda_{st} = 1$ and $\rho_{ut} = 1 > \lambda_{ut}$.

This is now immediate that when $\delta \in (\underline{\delta}, \bar{\delta})$, the mass of uneducated workers is zero and $m_{st} > \underline{b}_u(\delta)$ then $\rho_{ut} = \lambda_{ut} > 0$.

When $\delta \geq \bar{\delta}$, then from Propositions 1 and 3, we have $\lambda_{ut} = \rho_{ut} = 1$.

Finally, from the definition of $\underline{b}_u(\delta)$, when $\delta \in (\underline{\delta}, \bar{\delta})$, and $m_{st} \leq \underline{b}_u(\delta)$ then $\rho_{ut} = \lambda_{ut} = 0$.

2. This is immediate as at $\delta < \underline{\delta}$, $\lambda_{ut} = \rho_{ut} = 0$ and $\lambda_{st} = \rho_{st}$.

8.12 Proof of Observation 9

From Proposition 2 and Proposition 4 we see that when the degree of child affinity is not low, then the inequality at the (unique) steady state of the benchmark case is equal to the inequality at the least unequal steady state of the case with behavioral trap. At any other steady state the inequality is higher. Hence, the result.

References

Balboni, C., Bandiera, O., Burgess, R., Ghatak, M., and Heil, A. (2020). Why do People Stay Poor? *CEPR Discussion Paper No. DP14534*. 3

Banerjee, A. V. and Mullainathan, S. (2010). The Shape of Temptation: Implications for the Economic Lives of the Poor. *Working Paper*, National Bureau of Economic Research. 3

Banerjee, A. V. and Newman, A. F. (1993). Occupational Choice and the Process of Development. *Journal of Political Economy*, 101(2):274–298. 3

Barber, B. M. and Odean, T. (2001). Boys will be Boys: Gender, Overconfidence, and Common Stock Investment. *The Quarterly Journal of Economics*, 116(1):261–292. 2

Bénabou, R. and Tirole, J. (2002). Self-confidence and Personal Motivation. *The Quarterly Journal of Economics*, 117(3):871–915. 4

Bénabou, R. and Tirole, J. (2004). Willpower and Personal Rules. *Journal of Political Economy*, 112(4):848–886. 4

Bernheim, B., Ray, D., and Yeltekin, S. (2015). Poverty and Self-Control. *Econometrica*, 83(5):1877–1911. 3

Biais, B., Hilton, D., Mazurier, K., and Pouget, S. (2005). Judgemental overconfidence, self-monitoring, and trading performance in an experimental financial market. *The Review of Economic Studies*, 72(2):287–312. 4

Boca, D. D., Flinn, C., and Wiswall, M. (2014). Household Choices and Child Development. *The Review of Economic Studies*, 81(1 (286)):137–185. 2

Browning, M., Chiappori, P.-A., and Weiss, Y. (2014). *Economics of the Family*. Cambridge University Press. 2

Case, K. E. and Shiller, R. J. (2003). Is there a Bubble in the Housing Market? *Brookings Papers on Economic Activity*, 2003(2):299–362. 4

Cervone, D. and Peake, P. K. (1986). Anchoring, Efficacy, and Action: The Influence of Judgmental Heuristics on Self-efficacy Judgments and Behavior. *Journal of Personality and Social Psychology*, 50(3):492. 1

Chetty, R., Hendren, N., Jones, M. R., and Porter, S. R. (2019). Race and Economic Opportunity in the United States: an Intergenerational Perspective*. *The Quarterly Journal of Economics*, 135(2):711–783. 1, 3

Chowdry, H., Crawford, C., and Goodman, A. (2011). The Role of Attitudes and Behaviours in Explaining Socio-economic Differences in Attainment at Age 16. *Longitudinal and Life Course Studies*, 2(1):59–76. 1

Clark, J. and Friesen, L. (2009). Overconfidence in Forecasts of Own Performance: An Experimental Study. *The Economic Journal*, 119(534):229–251. 4

Dalton, P. S., Ghosal, S., and Mani, A. (2016). Poverty and Aspirations Failure. *The Economic Journal*, 126(590):165–188. 4

De Bondt, W. F. and Thaler, R. H. (1995). Financial Decision-making in Markets and Firms: A Behavioral Perspective. *Handbooks in Operations Research and Management Science*, 9:385–410. 4

Deshpande, A. and Newman, K. (2007). Where the Path Leads: The Role of Caste in Post-University Employment Expectations. *Economic and Political Weekly*, pages 4133–4140. 2

Filippin, A. and Paccagnella, M. (2012). Family Background, Self-Confidence and Economic Outcomes. *Economics of Education Review*, 31(5):824 – 834. 4

Galor, O. and Zeira, J. (1993). Income Distribution and Macroeconomics. *Review of Economic Studies*, 60:35–52. 3

Ghatak, M. (2015). Theories of Poverty traps and Anti-Poverty Policies. *The World Bank Economic Review*, 29(Suppl_1):S77–S105. 2, 3

Gramzow, R. H., Gaertner, L., and Sedikides, C. (2001). Memory for in-group and out-group information in a minimal group context: The self as an informational base. *Journal of Personality and Social Psychology*, 80(2):188. 1

Klugman, J. and Xu, J. (2008). Racial differences in public confidence in education: 1974–2002*. *Social Science Quarterly*, 89(1):155–176. 3

Malmendier, U. and Tate, G. (2005). CEO Overconfidence and Corporate Investment. *The Journal of Finance*, 60(6):2661–2700. 4

Mookherjee, D. and Ray, D. (2002). Contractual Structure and Wealth Accumulation. *American Economic Review*, 92(4):818–849. 3

Moore, D. A. (2007). Not So Above Average After All: When People Believe They are Worse than Average and its Implications for Theories of Bias in Social Comparison. *Organizational Behavior and Human Decision Processes*, 102(1):42–58. 4

Ray, D. (2006). Aspirations, Poverty, and Economic Change. In *Understanding Poverty*. Oxford University Press. 4

Ross, P. H. (2019). Occupation Aspirations, Education Investment, and Cognitive Outcomes: Evidence from Indian Adolescents. *World Development*, 123:104613. 1

Sterling, A. D., Thompson, M. E., Wang, S., Kusimo, A., Gilmartin, S., and Sheppard, S. (2020). The confidence gap predicts the gender pay gap among stem graduates. *Proceedings of the National Academy of Sciences*, 117(48):30303–30308. 3

Tversky, A. (2004). Features of Similarity. In Shafir, E., editor, *Preference, Belief, and Similarity: Selected Writings by Amos Tversky*, chapter 1, pages 7–45. The MIT Press. 10

Tversky, A. and Kaheneman, D. (2004). Judgement under Uncertainty: Heuristics and Biases. In Shafir, E., editor, *Preference, Belief, and Similarity: Selected Writings by Amos Tversky*, chapter 8, pages 203–220. The MIT Press. 2

Van Veelen, R., Otten, S., Cadinu, M., and Hansen, N. (2016). An Integrative Model of Social Identification: Self-stereotyping and Self-anchoring as Two Cognitive Pathways. *Personality and Social Psychology Review*, 20(1):3–26. 1

SB Supplementary Appendix

SB.1 Proof of Observation 1.

By assumption, the mass of adults is 1, each of them works either as a skilled worker or as an unskilled worker: $L_{ut} + L_{st} = 1$. Hence, $0 \leq L_{st} \leq 1$. Since $0 < \phi < 1$, $L_{st}^{-(1-\phi)} \geq 1$. Thus, $m_{st} = AL_{st}^{-(1-\phi)} \geq 1 = m_{ut}$.
The inequality binds only when $L_{st} = 1$ and $A = 1$. □

SB.2 Proof of Lemma 1.

1. Let us define $y(x)$ such that

$$y(x) = \frac{x^\sigma}{\sigma} - \frac{(x-\bar{s})^\sigma}{\sigma}, \quad x > 1 > \bar{s}.$$

$\sigma < 0$ implies $y'(x) < 0$ and $y''(x) > 0$. Thus, Observation 1, i.e., $m_{st} \geq m_{ut}$ implies $y(m_{st}) \leq y(m_{ut})$:

$$\frac{m_{ut}^\sigma}{\sigma} - \frac{(m_{ut}-\bar{s})^\sigma}{\sigma} = \frac{1}{\sigma} - \frac{(1-\bar{s})^\sigma}{\sigma} \geq \frac{m_{st}^\sigma}{\sigma} - \frac{(m_{st}-\bar{s})^\sigma}{\sigma}$$

i.e. utility cost of investment for the unskilled workers is (weakly) higher.
It is strictly higher whenever the mass of unskilled workers is positive: $L_{ut} > 0$.

Now, consider an economy with both types of workers i.e. $L_{ut} > 0$ and $L_{st} > 0$. Let $\langle \lambda_{ut}, \lambda_{st} \rangle$ be any equilibrium, where $\lambda_{ut} > 0$, then we want to show $\lambda_{st} = 1$. We start with the case $\lambda_{ut} \in (0, 1)$. Then,

$$\delta \left[\frac{\left[\beta^\phi A[\lambda_{st}L_{st} + \lambda_{ut}L_{ut}]^{-(1-\phi)} + 1 - \beta\right]^\sigma}{\sigma} - \frac{1}{\sigma} \right] = \frac{1}{\sigma} - \frac{(1-\bar{s})^\sigma}{\sigma} > \frac{m_{st}^\sigma}{\sigma} - \frac{(m_{st}-\bar{s})^\sigma}{\sigma}$$

where the first equality is from the investment decision of the unskilled workers and the inequality is due to $m_{st} > 1$.

This implies λ_{st} must be 1, otherwise a skilled worker would have an incentive to deviate and invest with a higher probability. Now, it is evident that if $\lambda_{ut} = 1$, then $\lambda_{st} = 1$. □

2. Recall (3), a worker of type j, where $j \in \{u, s\}$, invests with probability λ_{jt} if

$$\delta \left[\frac{\left[\beta^\phi A[\lambda_{st}L_{st} + \lambda_{ut}(1-L_{st})]^{-(1-\phi)} + 1 - \beta\right]^\sigma}{\sigma} - \frac{1}{\sigma} \right] \geq \frac{m_{jt}^\sigma}{\sigma} - \frac{(m_{jt}-\bar{s})^\sigma}{\sigma}.$$

where the inequality binds for j^{th} type when $\lambda_{jt} \in (0, 1)$. From part (a), we also know $\lambda_{ut} > 0$ only when $\lambda_{st} = 1$.

Now, as m_{st} increases, the utility cost of investment, i.e., the R.H.S of the above inequality decreases for the skilled workers and remains the same for the unskilled workers. Next, observe, increase in m_{st} implies decrease in L_{st}. It can be shown that given $\langle \lambda_{ut}, \lambda_{st} \rangle$, the L.H.S. of the above inequality (weakly) increases with decrease in L_{st} as $\lambda_{st} \geq \lambda_{ut}$. Thus, with increase in the income of the skilled workers at period t, the benefit of investment (weakly) increases and the utility cost of investment (weakly) decreases. Hence, the probability of investment must (weakly) increase. $\square$

SB.3 Proof of Observation 2.

Since $\bar{s} \in (0,1)$, clearly $\underline{\delta} = (1-\bar{s})^\sigma - 1 > 0$. We know $A \geq 1$, $\beta \in (0,1)$ and that $(A\beta^\phi + 1 - \beta)$ is a weighted average of $A\beta^{-(1-\phi)}$ and 1. Hence, $A\beta^\phi + 1 - \beta \geq 1$, which together with $\sigma < 0$ yields $\underline{\delta} < \bar{\delta}$. $\square$

SB.4 Formal Expression for the Definition 2

- At $\bar{b}_s(\delta)$ a *skilled* worker is indifferent between investing and not investing, when all other skilled worker invest with probability 1 and no unskilled worker invests. Thus, $L_{st+1} = \beta L_{st}$, $m_{st+1} = AL_{st+1}^{-(1-\phi)} = \beta^{-(1-\phi)}\bar{b}_s(\delta)$ and it must be that

$$\bar{b}_s(\delta): \quad \delta\left[\frac{[\beta^\phi \bar{b}_s + 1 - \beta]^\sigma}{\sigma} - \frac{1}{\sigma}\right] = \frac{\bar{b}_s^\sigma}{\sigma} - \frac{(\bar{b}_s - \bar{s})^\sigma}{\sigma}. \tag{S.1}$$

- At $\underline{b}_u(\delta)$ an *unskilled* worker is indifferent between investing and not investing, when all the skilled workers invest with probability 1 and no other unskilled worker invests. Thus, $L_{st+1} = \beta L_{st}$, $m_{st+1} = AL_{st+1}^{-(1-\phi)} = \beta^{-(1-\phi)}\underline{b}_u(\delta)$ and it must be that

$$\underline{b}_u(\delta): \quad \delta\left[\frac{[\beta^\phi \underline{b}_u + 1 - \beta]^\sigma}{\sigma} - \frac{1}{\sigma}\right] = \frac{1}{\sigma} - \frac{(1 - \bar{s})^\sigma}{\sigma}. \tag{S.2}$$

SB.5 Proof of Lemma 2.

1. The ranking is immediate from comparing definitions, (8.3), (S.1), and (S.2). Differentiating these equations with respect to δ, we get that the thresholds are decreasing in δ.

2. (i) We, first, show that if the degree of child affinity is moderate then $\bar{b}_s(\delta) > 1$.

$$\delta < \bar{\delta} \equiv \frac{(1-\bar{s})^\sigma - 1}{1 - (A\beta^\phi + 1 - \beta)^\sigma} \leq \frac{(1-\bar{s})^\sigma - 1}{1 - (\beta^\phi + 1 - \beta)^\sigma} \quad \text{as} \quad A \geq 1 \text{ and } \sigma < 0,$$

$$\Rightarrow \quad \delta\left[\frac{(\beta^\phi + 1 - \beta)^\sigma}{\sigma} - \frac{1}{\sigma}\right] < \frac{1}{\sigma} - \frac{(1-\bar{s})^\sigma}{\sigma}$$

   which yields that the expected net benefit from investment for a skilled parent is lower than the utility cost at $m_{st} = 1$. Since, at $\bar{b}_s(\delta)$ the expected net benefit for a skilled parent equals the utility cost of investment, it must be $\bar{b}_s(\delta) > 1$.

   Similarly, from (8.3) it can be shown that $\underline{b}_s(\delta) > 1$ if and only if $\delta < \underline{\delta}$.

(ii) When $\delta \in [0, \underline{\delta})$, that $\underline{b}_u(\delta) = \infty$, follows directly from (S.2). $1 < \underline{b}_s(\delta)$ has already been shown above. Hence, proved. $\square$

SB.6 Proof of Observation 5

Differentiating equation (A.1) with respect to δ, we find that $\bar{b}_u(\delta)$ is decreasing in δ. Now about the value of this threshold for different child affinity ranges. Substituting $\bar{\delta}$ in (A.1), we have $\bar{b}_u(\bar{\delta}) = A\beta^{-(1-\phi)}$. Suppose $\delta \in [\underline{\delta}, \bar{\delta})$. From the definitions of $\underline{b}_u(\delta)$ and $\bar{b}_u(\delta)$ we get $\bar{b}_u(\delta) = \beta^{-(1-\phi)}\underline{b}_u(\delta)$, which obviously is greater than $\underline{b}_u(\delta)$. That $\bar{b}_u(\delta) = \infty$ when the degree of child affinity is low follows directly from (A.1). $\square$

SB.7 Proof of Lemma 3

We prove by contradiction. Suppose not, there exists an equilibrium $\langle \gamma_{ut}, \gamma_{st} \rangle$ such that $\gamma_{ut} > 0$ and $\gamma_{st} = 0$.

We show that at such an equilibrium, an educated-unskilled worker would have a unilateral incentive to deviate. Formally, at equilibrium $\gamma_{ut} > 0$ and $\gamma_{st} = 0$ imply

$$
\begin{aligned}
\delta\Big[1 - \big[\theta\beta[\theta(1-\beta)\gamma_{ut}]^{-(1-\phi)}m_{st} + 1 - \theta\beta\big]^{\sigma}\Big] &\geq (1-\bar{s})^{\sigma} - 1 > (m_{st} - \bar{s})^{\sigma} - m_{st}^{\sigma} \quad (\text{since } m_{st} > 1) \\
&> \delta\Big[1 - \big[[1-\theta(1-\beta)][(1-\beta)\gamma_{ut}]^{-(1-\phi)}m_{st} + \theta(1-\beta)\big]^{\sigma}\Big] \quad (\text{since } \gamma_{st} = 0) \\
\Rightarrow \quad (1-\theta) &\geq [(1-\beta)\gamma_{ut}]^{-(1-\phi)}\big[[1-\theta(1-\beta)] - \theta^{\phi}\beta\big]m_{st}. \qquad \text{(S.3)}
\end{aligned}
$$

Now, define a function $L(\theta) = (1-\beta)^{-(1-\phi)}(1 - \theta(1-\beta) - \theta^{\phi}\beta) - 1 + \theta$.

Observe, $L(0) = (1-\beta)^{-(1-\phi)} - 1$ and $L(1) = 0$. Further,

$$
\begin{aligned}
L'(\theta) &= -(1-\beta)^{-(1-\phi)}(1 - \beta + \phi\theta^{-(1-\phi)}\beta) + 1 \\
L'(\theta) &= 0 \quad \text{at} \quad \theta = \left[\frac{(1-\beta)^{(1-\phi)} - (1-\beta)}{\phi\beta}\right]^{-\frac{1}{1-\phi}} > \left[\frac{1}{\phi}\right]^{-\frac{1}{1-\phi}} > 1 \\
L''(\theta) &= (1-\beta)^{-(1-\phi)}\phi(1-\phi)\beta\theta^{-(2-\phi)} > 0
\end{aligned}
$$

Since $L'(\theta) < 0$ for all $\theta \in [0, 1]$ and the boundary values of $L(\theta)$ at 0 and 1 are non-negative, $L(\theta) > 0$ for all $\theta \in [0, 1)$. Thus,

$$
(1-\beta)^{-(1-\phi)}\big[[1 - \theta(1-\beta) - \theta^{\phi}\beta\big] > 1 - \theta.
$$

Hence, $[(1-\beta)\gamma_{ut}]^{-(1-\phi)}\big[[1-\theta(1-\beta)-\theta^{\phi}\beta\big]m_{st} > (1-\beta)^{-(1-\phi)}\big[[1-\theta(1-\beta)-\theta^{\phi}\beta\big] > 1-\theta$ which contradicts (S.3). $\square$

SB.8 Proof of Observation 6

We have already shown in Observation 2 that $0 < \underline{\delta}$. Here we show, $\underline{\delta} < \delta_a$. The weighted average of $(\theta(1-\beta) + \beta)^{-(1-\phi)}$ and 1 will be greater than 1. It follows,

$$\underline{\delta} = (1-\bar{s})^{\sigma} - 1 < \frac{(1-\bar{s})^{\sigma} - 1}{1 - [\theta\beta(\theta(1-\beta)+\beta)^{-(1-\phi)} + 1 - \theta\beta]^{\sigma}} = \delta_a. \qquad \square$$

SB.9 Formal Expressions for Definition 5

- For a given degree of child affinity δ, at $a_6(\delta)$ a skilled worker is indifferent between investing and not investing, when all other skilled workers invest with certainty and no educated-unskilled worker invests. So, from (6)

$$a_6(\delta): \quad \frac{\delta}{\sigma}\left[\left[[1-\theta(1-\beta)]^{\phi} a_6 + \theta(1-\beta)\right]^{\sigma} - 1\right] = \frac{a_6^{\sigma} - (a_6 - \bar{s})^{\sigma}}{\sigma}. \tag{S.4}$$

- At $a_4(\delta)$ a skilled worker is indifferent between investing and not investing, when no other skilled worker invests and all educated-unskilled workers invest with certainty. From (6)

$$a_4(\delta): \quad \frac{\delta}{\sigma}\left[\left[[1-\theta(1-\beta)](1-\beta)^{-(1-\phi)} a_4 + \theta(1-\beta)\right]^{\sigma} - 1\right] = \frac{a_4^{\sigma} - (a_4 - \bar{s})^{\sigma}}{\sigma}. \tag{S.5}$$

- At $a_2(\delta)$ a skilled worker is indifferent between investing and not investing, when all other educated workers invest with certainty. So, from (6)

$$\begin{aligned} a_2(\delta): \quad & \frac{\delta}{\sigma}\left[\left[[1-\theta(1-\beta)][1+(1-\theta)(1-\beta)]^{-(1-\phi)} a_2 + \theta(1-\beta)\right]^{\sigma} - 1\right] \\ = \; & \frac{a_2^{\sigma} - (a_2 - \bar{s})^{\sigma}}{\sigma}. \end{aligned} \tag{S.6}$$

- At $a_5(\delta)$ an educated-unskilled worker is indifferent between investing and not investing, when all skilled workers invest and no other educated-unskilled worker invests. From (5)

$$a_5(\delta): \frac{\delta}{\sigma}\left[\left[\theta\beta^{\phi} a_5 + 1 - \theta\beta\right]^{\sigma} - 1\right] = \frac{1 - (1-\bar{s})^{\sigma}}{\sigma}. \tag{S.7}$$

- Here at $a_3(\delta)$ an educated-unskilled worker is indifferent between investing and not investing, when all other educated-unskilled workers invest with certainty and no skilled worker invests. So, from (5)

$$a_3(\delta): \quad \frac{\delta}{\sigma}\left[\left[\theta\beta[\theta(1-\beta)]^{-(1-\phi)} a_3 + 1 - \theta\beta\right]^{\sigma} - 1\right] = \frac{1 - (1-\bar{s})^{\sigma}}{\sigma} = 0. \tag{S.8}$$

- At $a_1(\delta)$ an educated-unskilled worker is indifferent between investing and not investing, when all other educated workers invest with certainty. So, from (5)

$$a_1(\delta): \quad \frac{\delta}{\sigma}\left[\left[\theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)} a_1 + 1 - \theta\beta\right]^{\sigma} - 1\right] = \frac{1 - (1-\bar{s})^{\sigma}}{\sigma}. \tag{S.9}$$

SB.10 Proof of Lemma 4

1. We want to show that all the thresholds are decreasing in δ. Consider the income-cutoff $a_6(\delta)$, which is determined by (S.4). The L.H.S. of (S.4) is increasing in a_6 and δ and the R.H.S is decreasing in a_6. Hence, the claim. Following similar argument, this negative relationship can be shown for the other thresholds as well.

2. (i) Suppose not and $a_2(\delta) \leq 1$. Since $y(x)$ is decreasing in x as in Appendix SB.2 we have

$$\frac{a_2(\delta)^\sigma - (a_2(\delta) - \bar{s})^\sigma}{\sigma} \geq \frac{1 - (1 - \bar{s})^\sigma}{\sigma} > 0$$

that is, the utility cost of investment is positive. So, it is enough to show that the benefit from investment at the premises of the definition is negative, i.e.

$$\frac{\delta}{\sigma}\Big[\big[[1-\theta(1-\beta)][1+(1-\theta)(1-\beta)]^{-(1-\phi)}a_2(\delta)+\theta(1-\beta)\big]^\sigma - 1\Big] < 0.$$

The following steps give us that

$$\begin{aligned}
& & [1+(1-\theta)(1-\beta)]^{-(1-\phi)} &< 1 \\
\Rightarrow \quad & [1-\theta(1-\beta)][1+(1-\theta)(1-\beta)]^{-(1-\phi)}a_2(\delta)+\theta(1-\beta) & &< [1-\theta(1-\beta)]a_2(\delta)+\theta(1-\beta) \\
& & &< \quad 1-\theta(1-\beta)+\theta(1-\beta) = 1.
\end{aligned}$$

(ii) The benefit of investment at $\underline{a}_s(\delta)$ is higher than that at the premise of $a_4(\delta)$. Hence, to make a skilled worker indifferent at the thresholds, $\underline{a}_s(\delta)$ must be lower than $a_4(\delta)$.

<u>We show $a_2(\delta) > a_4(\delta)$.</u> Suppose not and $a_2(\delta) \leq a_4(\delta)$. Hence, from Lemma 1 it follows that the utility cost of investment at $a_2(\delta)$ is no less than that at $a_4(\delta)$:

$$\frac{a_4(\delta)^\sigma - (a_4(\delta) - \bar{s})^\sigma}{\sigma} \leq \frac{a_2(\delta)^\sigma - (a_2(\delta) - \bar{s})^\sigma}{\sigma} \quad \Rightarrow (a_4(\delta)-\bar{s})^\sigma - a_4(\delta)^\sigma \leq (a_2(\delta)-\bar{s})^\sigma - a_2(\delta)^\sigma.$$

Since, $1+(1-\theta)(1-\beta) > 1-\beta$, we get

$$\begin{aligned}
\Rightarrow \quad & \delta\Big[1 - \big[[1-\theta(1-\beta)][1+(1-\theta)(1-\beta)]^{-(1-\phi)}a_2(\delta)+\theta(1-\beta)\big]^\sigma\Big] \\
< \quad & \delta\Big[1 - \big[[1-\theta(1-\beta)](1-\beta)^{-(1-\phi)}a_4(\delta)+\theta(1-\beta)\big]^\sigma\Big] \\
\Rightarrow \quad & (a_2(\delta)-\bar{s})^\sigma - a_2(\delta)^\sigma < (a_4(\delta)-\bar{s})^\sigma - a_4(\delta)^\sigma \quad \text{from definition of } a_2(\delta) \text{ and } a_4(\delta),
\end{aligned}$$

which contradicts the above.

(iii) Similar steps can be used to prove it.

(iv) We show by contradiction that when $\theta(1-\beta) > \beta$ then $a_4(\delta) > a_6(\delta)$. The converse can be shown similarly which we skip.

Suppose $\theta(1-\beta) > \beta$ and $a_4(\delta) \leq a_6(\delta)$ which implies utility cost of investment at $a_4(\delta)$ is no less than that at $a_6(\delta)$: $(a_6(\delta)-\bar{s})^\sigma - a_6(\delta)^\sigma \leq (a_4(\delta)-\bar{s})^\sigma - a_4(\delta)^\sigma$ Since $\theta(1-\beta) > \beta$, we get

$$\begin{aligned} \Rightarrow \quad & \delta\Big[1 - \big[[1-\theta(1-\beta)](1-\beta)^{-(1-\phi)}a_4(\delta) + \theta(1-\beta)\big]^\sigma\Big] \\ < \quad & \delta\Big[1 - \big[[1-\theta(1-\beta)]^\phi a_6(\delta) + \theta(1-\beta)\big]^\sigma\Big] \end{aligned}$$

or the benefit at the primitive of $a_4(\delta)$ is strictly lower than that of $a_6(\delta)$. Hence, the statements of both the definitions of $a_4(\delta)$ and $a_6(\delta)$ cannot simultaneously be true.

3. a. Let us write the threshold $a_1(\delta)$ expression (S.9) as:

$$\big[\theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}a_1 + 1 - \theta\beta\big]^\sigma = \frac{1-(1-\bar{s})^\sigma + \delta}{\delta}.$$

If $\delta \leq (1-\bar{s})^\sigma - 1$, then R.H.S is negative and hence there does not exist any finite $a_1(\delta)$ which satisfies the above equation. If $\delta > (1-\bar{s})^\sigma - 1$, then R.H.S is a positive fraction, L.H.S. is decreasing in $a_1(\delta)$, and for all positive values of $a_1(\delta)$, the L.H.S. is bounded in $[0, (1-\theta\beta)^\sigma)$, where $(1-\theta\beta)^\sigma > 1$. Thus, there exists a finite $a_1(\delta)$ at which L.H.S. equals R.H.S.

Similarly using equations (S.7) and (S.8) the same can be shown for $a_5(\delta)$ and $a_3(\delta)$.

b. Suppose $a_1(\delta) \gtreqless 1$. Using this in (S.9) we get

$$\delta \gtreqless \frac{(1-\bar{s})^\sigma - 1}{1 - [\theta\beta(\theta(1-\beta)+\beta)^{-(1-\phi)} + 1 - \theta\beta]^\sigma} \equiv \delta_a.$$

c. The cost of investment for the educated-unskilled worker is independent of the state variable, so we compare the benefits at $a_1(\delta)$, $a_3(\delta)$, $a_5(\delta)$. Since $a_1(\delta)$, $a_3(\delta)$ and $a_5(\delta)$ are finite if and only if $\delta > \underline{\delta}$, so the following ranking holds only for $\delta > \underline{\delta}$.

First we show $a_5(\delta) < a_1(\delta)$ Comparing (S.9) and (S.7) we get,

$$\frac{\delta}{\sigma}\Big[\big[\theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}a_1(\delta) + 1 - \theta\beta\big]^\sigma - 1\Big] = \frac{\delta}{\sigma}\big[\big[\theta\beta^\phi a_5(\delta) + 1 - \theta\beta\big]^\sigma - 1\big]$$

$$\Rightarrow \quad [\theta(1-\beta)+\beta]^{-(1-\phi)}a_1(\delta) = \beta^{-(1-\phi)}a_5(\delta) \qquad \Rightarrow \quad a_1(\delta) > a_5(\delta) \quad \text{as } \theta(1-\beta)+\beta > \beta$$

Now, we show $a_3(\delta) < a_1(\delta)$ Comparing (S.9) and (S.8), we get

$$\begin{aligned}
&\frac{\delta}{\sigma}\left[\left[\theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}a_1(\delta)+1-\theta\beta\right]^{\sigma}-1\right]=\frac{\delta}{\sigma}\left[\left[\theta\beta[\theta(1-\beta)]^{-(1-\phi)}a_3(\delta)+1-\theta\beta\right]^{\sigma}-1\right]\\
&\quad\Rightarrow\quad \theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}a_1(\delta)=\theta\beta[\theta(1-\beta)]^{-(1-\phi)}a_3(\delta)\\
&\quad\Rightarrow\quad a_1(\delta)>a_3(\delta)\quad \text{as } \theta(1-\beta)+\beta>\theta(1-\beta)
\end{aligned}$$

$a_3(\delta) > a_5(\delta)$ if and only if $\theta(1-\beta) > \beta$ Comparing (S.8) and (S.7) we get,

$$\begin{aligned}
&\frac{\delta}{\sigma}\left[\left[\theta\beta[\theta(1-\beta)]^{-(1-\phi)}a_3(\delta)+1-\theta\beta\right]^{\sigma}-1\right]=\frac{\delta}{\sigma}\left[\left[\theta\beta^{\phi}a_5(\delta)+1-\theta\beta\right]^{\sigma}-1\right]\\
&\Rightarrow\quad \theta\beta[\theta(1-\beta)]^{-(1-\phi)}a_3(\delta)=\theta\beta^{\phi}a_5(\delta)
\end{aligned}$$

Hence, $a_3(\delta) > a_5(\delta)$ if and only if $[\theta(1-\beta)]^{-(1-\phi)} < \beta^{-(1-\phi)} \Leftrightarrow \theta(1-\beta) > \beta$.

4. First, we show $\forall\ \delta \le \delta_a,\ a_4(\delta) \le a_1(\delta)$. Suppose not. $\exists\ \delta \le \delta_a$ such that $a_1(\delta) < a_4(\delta)$. Consider any $m_{st} \in [a_1(\delta), a_4(\delta)]$. Since, $m_{st} \ge a_1(\delta)$, $\gamma_{ut} = 1\ \forall\ \gamma_{st} = [0,1]$. From the definition of $a_4(\delta)$, γ_{st} must be equal to zero for $m_{st} < a_4(\delta)$, which contradicts Lemma 3.

   Now, we show $\forall\ \delta > \delta_a,\ a_4(\delta) < 1$. It can again be proved by contradiction following the aforementioned argument in the range $m_{st} \in [1, a_4(\delta)]$.

5. Comparing definitions (8.3) and (A.2) we get $\underline{b}_s(\delta) = \underline{a}_s(\delta)$.

   $\underline{b}_u(\delta) < a_5(\delta)$. First, observe that $a_5(\delta)$ coincides with $\underline{b}_u(\delta)$ when $\theta = 1$. So, to establish the claim it is sufficient to show that

$$\delta\left[\frac{[\theta\beta^{\phi}\underline{b}_u(\delta)+1-\theta\beta]^{\sigma}}{\sigma}-\frac{1}{\sigma}\right]<\delta\left[\frac{[\beta^{\phi}\underline{b}_u(\delta)+1-\beta]^{\sigma}}{\sigma}-\frac{1}{\sigma}\right]. \tag{S.10}$$

   As this implies

$$\begin{aligned}
&\underbrace{\delta\left[\frac{[\theta\beta^{\phi}\underline{b}_u(\delta)+1-\theta\beta]^{\sigma}}{\sigma}-\frac{1}{\sigma}\right]}_{\substack{\text{Perceived Benefit from investment at}\\ b_u(\delta)\text{ at the premise of } a_5(\delta)}}<\delta\left[\frac{[\beta^{\phi}\underline{b}_u(\delta)+1-\beta]^{\sigma}}{\sigma}-\frac{1}{\sigma}\right]=\frac{1-(1-\bar{s})^{\sigma}}{\sigma}\\
&\qquad=\frac{\delta}{\sigma}\left[\left[\theta\beta^{\phi}a_5+1-\theta\beta\right]^{\sigma}-1\right]
\end{aligned}$$

   that is the *perceived* benefit from investment is lower than the cost of investment at $\underline{b}_u(\delta)$. Thus, an educated-unskilled worker with behavioral anomaly would not invest at $\underline{b}_u(\delta)$, she would start investing at a higher state variable. Hence, $a_5(\delta)$ must be strictly higher than $\underline{b}_u(\delta)$ at all $\theta \in (0,1)$.

To prove the inequality, let $z = \theta\beta^{\phi}\underline{b}_u(\delta) + 1 - \theta\beta$. Then,

$$\frac{\partial z}{\partial \theta} = \beta^{\phi}\underline{b}_u(\delta) - \beta > 0 \qquad \text{if } \underline{b}_u > \beta^{1-\phi}.$$

Observe if $\underline{b}_u = \beta^{1-\phi}$, the benefit from investment is zero and cost is positive, but $\underline{b}_u(\delta)$ should be such that the benefit is equal to cost. The benefit is increasing in m_{st}, hence, $\underline{b}_u(\delta)$ must be greater than $\beta^{1-\phi}$. Hence, the inequality (S.10) is true $\forall\theta < 1$. □

SB.11 Figure from Numerical Example

Figure 5 depicts an example of equilibrium trajectory for $\theta(1-\beta) > \beta$. Interestingly, here, the probability of investment by skilled workers weakly increases over time but not for the educated-unskilled worker.

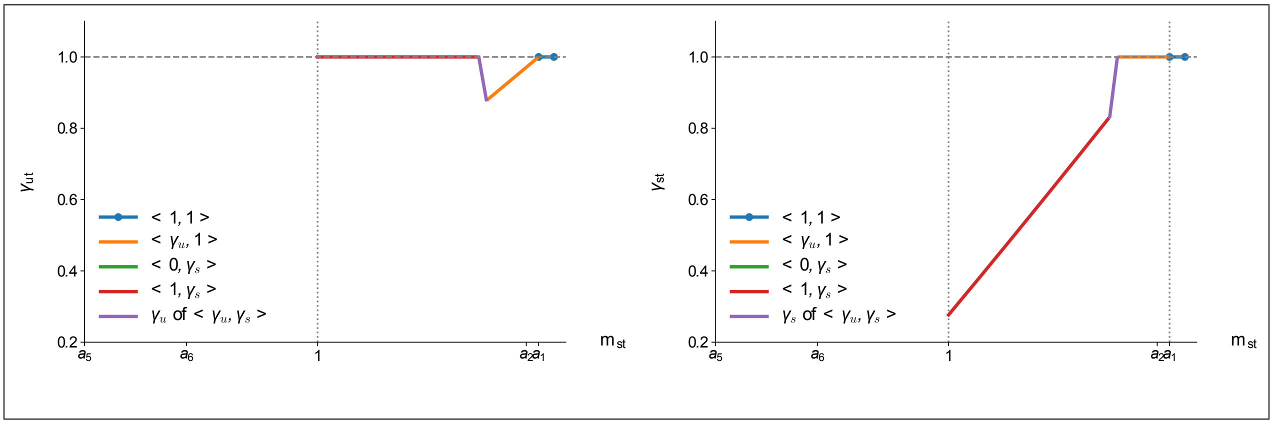


Figure 5: Example to depict equilibrium for $\theta(1-\beta) > \beta$. [Colored Graphs]

SB.12 Proof of Observation 10

We have already noted in Observation 2 that $\underline{\delta} < \bar{\delta}$. Now we show

$$\bar{\delta} \equiv \frac{(1-\bar{s})^{\sigma} - 1}{1 - [A\beta^{\phi} + 1 - \beta]^{\sigma}} < \frac{(1-\bar{s})^{\sigma} - 1}{1 - [\theta\beta(\theta(1-\beta)+\beta)^{-(1-\phi)} + 1 - \theta\beta]^{\sigma}} \equiv \delta_a.$$

Comparing these δ values we get that this statement is true if and only if

$$\theta\beta(\theta(1-\beta)+\beta)^{-(1-\phi)} + 1 - \theta\beta < A\beta^{\phi} + 1 - \beta \tag{S.11}$$

We define a function $L(\theta)$ and derive its properties:

$$
\begin{aligned}
L(\theta) &= \theta(\theta(1-\beta)+\beta)^{-(1-\phi)} - \theta + 1 - \beta^{-(1-\phi)} \quad \text{and } L(0) = L(1) = 1 - \beta^{-(1-\phi)} < 0 \\
L'(\theta) &= (\beta + \theta(1-\beta))^{-(2-\phi)}(\beta + \phi\theta(1-\beta)) - 1 \\
L'(0) &= \beta^{-(1-\phi)} - 1 > 0 \quad \text{and } L'(1) = -(1-\beta)(1-\phi) < 0 \\
L'(\bar{\theta}) &= 0 \quad \text{where} \quad \bar{\theta}\ : (\beta + \bar{\theta}(1-\beta))^{-(1-\phi)} = \frac{\beta + \bar{\theta}(1-\beta)}{\beta + \phi\bar{\theta}(1-\beta)} < \beta^{-(1-\phi)} \\
L(\bar{\theta}) &= \underbrace{\frac{\beta + \bar{\theta}(1-\beta)(\phi + \bar{\theta}(1-\phi))}{\beta + \bar{\theta}(1-\beta)}}_{fraction} \underbrace{(\beta + \bar{\theta}(1-\beta))^{-(1-\phi)}}_{<\beta^{-(1-\phi)}} - \beta^{-(1-\phi)} < 0 \\
L''(\theta) &= -(1-\phi)(1-\beta)(\beta + \theta(1-\beta))^{-(3-\phi)}(2\beta + \theta\phi(1-\beta)) < 0
\end{aligned}
$$

Thus L is (a) a strictly concave function, (b) has a maxima at $\bar{\theta}$, (c) the maxmimum value is negative. Thus, $L(\theta)$ is negative values for all $\theta \in [0, 1]$. Thus,

$$
\begin{aligned}
& \theta(\theta(1-\beta)+\beta)^{-(1-\phi)} - \theta + 1 - \beta^{-(1-\phi)} < 0 \\
\Rightarrow \quad & \theta\beta(\theta(1-\beta)+\beta)^{-(1-\phi)} + 1 - \theta\beta < \beta^{\phi} + 1 - \beta < A\beta^{\phi} + 1 - \beta
\end{aligned}
$$

Thus, equation (S.11) always holds and hence $\bar{\delta} < \delta_a$. $\square$