Perceptions, Biases, and Inequality*

Dyotona Dasgupta† Anuradha Saha‡

December 26, 2020

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. Indian Statistical Institute, 7 SJS Sansanwal Marg, New Delhi, India 110016. 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

"*Successes build a robust belief in one's personal efficacy. Failures undermine it.*"
Bandura (1994)

"*A positive experience makes us happy, but it also renders similar experiences less exciting. A negative experience makes us unhappy, but it also helps us appreciate subsequent experiences that are less bad.*" Tversky and Griffin (2004)

It is well documented in the social sciences that socio-economic background plays an important role in shaping beliefs about self-efficacy and that can have long-run effects on inequality and poverty.[1] Filippin and Paccagnella (2012) find that children who lack self-confidence under accumulate human capital. They analyze how inherited beliefs about self-efficacy could contribute to growing inequality.[2] Our paper provides a mechanism in which self-efficacy beliefs augment behavioral inequality. We analyze the dynamics of human capital distribution and income inequality of economies with agents who have biased perceptions about their future.

In a dynamic setting, we build a novel framework where a decision maker's assessment of returns to investment is biased. She anchors the evaluation of the returns to her own experiences, and to the success or failure of people who she perceives similar to herself. Success makes an individual overconfident, in terms of the probability of future success, but also makes her underestimate the rewards from such a success (capturing the contrast effect introduced by Tversky and Griffin (2004)) 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. This generates two interesting results: behavioral anomalies may influence low-income households to invest more than high-income households, and it may even lower long-run income inequality.

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 aspiration (captured through 'milestone' utility in Ray (2006) 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 framework captures the empirical evidence that successful people display *overconfidence* (see Malmendier and Tate (2005) for example) while unsuccessful or poor

[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). We use this literature as a basis to claim that individuals identify with similar groups. In an inter-disciplinary work, Chowdry et al. (2011) document that differences in the attitudes of youngsters (around the age of 16) and their parents play a key role in explaining the rich-poor gap in secondary education attainment. Within the field of economics, Ross (2019) shows that parental education and economic activity influence their and their children's beliefs about future job prospects.

[2] We discuss this paper in more detail later.

persons display *under confidence* (see Bernard et al. (2011) for example). Why should we care about such biases? One distinct example where such biases matter is in human resource management within firms. Employees within a firm often come from different backgrounds and firms conduct diversity training to facilitate positive intergroup interaction. Lindsey et al. (2017) find that the efficacy of such training programs is conditional on the initial empathy levels within employees. Our paper highlights similar inter-linkages at a macroeconomic level. We characterize the effect of biases on inequality or investment rates, conditional on socio-economic conditions.

To this end, we construct an overlapping generations model where adults differ in terms of their education – educated or uneducated, and jobs – skilled or unskilled. Based on their education and jobs, adults can be classified into three groups – uneducated-unskilled, educated-unskilled, and educated-skilled. These groups are endogenously formed and can change every period depending on their education and luck. 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.[3] Each adult decides whether to invest in her child's education or not based on perceived expected return.

We say that the agents are biased when they form a belief about the probability of success based on their education and job.[4] We assume that such a decision maker gives greater weight to the experiences of the individuals who are similar to her and discounts the experiences of those individuals who she sees as different. Following Tversky (2004), we assume similarity with another individual is perceived higher when they share a higher number of common attributes. We capture this by a "degree of association" – the lower the degree of association, the more biased is the perceived probability of success. Due to higher dissimilarity with the skilled workers, an uneducated-unskilled worker is more under confident than an educated-unskilled worker about her child getting a skilled job. Conversely, an educated-skilled worker underestimates the probability of her educated child becoming an unskilled worker. All agents are non-Bayesian and there is no learning or convergence of beliefs. However, *given their beliefs*, each parent correctly calculates the equilibrium mass of skilled workers and their income. Unsuccessful people overestimate the reward in case of success and successful people do the converse.

The central objective of our paper is to compare the dynamics and steady-state properties of an economy with and without behavioral anomalies. At any initial income and degree of association, we calculate the probabilities of investment of each type of worker, such that no one has any incentive to deviate unilaterally (and their decision is consistent with their beliefs). We find that the weight a parent places on the utility from her child's perceived expected income relative to the utility from their own consumption plays a key role in the characterization. We call this the *child affinity parameter* (or intergenerational altruism as in Ghatak (2015)). The parameter is non-negative,[5] time-independent, and common for

[3]To capture the effect of behavioral anomaly solely, we abstract from any bequest motive.

[4]Given there is no inherent difference among the children, the education-job identity-based belief captures the bias in this model.

[5]As noted in Boca et al. (2014) children may be valued more or less than parents' own consumption,

all parents in an economy. In our characterization, we find that economic outcomes have distinct properties as per four child affinity ranges – low, moderate, high, and huge.

In absence of any behavioral anomalies, in the benchmark case, if the unskilled workers (with low incomes) invest in their children's education with a positive probability, then the skilled workers invest with certainty. This is because of the concavity of utility function and that without any behavioral anomaly, the expected benefit from investment is the same for all parents. Due to this, at any given degree of child affinity and initial skilled income, the equilibrium is unique. We also find, in the benchmark case, there is a poverty trap only when the degree of child affinity is low. Here, we define the poverty trap as a situation where there exists a positive mass of families that never become rich.

Behavioral anomalies significantly affect the economy. First, we consider when only uneducated workers are imprisoned in a behavioral trap – they do not believe that an educated child from their group would ever be able to get a skilled job. This implies they never invest in education. Educated parents take this into account while making their investment decision. They can rationally foresee that with fewer educated (and hence skilled) workers in the next period, the skilled income will be higher.[6] This incentivizes educated adults to invest with (weakly) higher probabilities than in the benchmark case. For low child affinity, the steady-state economic outcomes are the same as in the benchmark case. For moderate and high child affinity, due to non-investment of uneducated-unskilled workers, behavioral trap gives rise to multiple steady states. These multiple steady states can be ranked based on inequality. The steady-state inequality is at least as high as that in the benchmark case. In the behavioral trap case, there is no intergenerational mobility for the uneducated – once a child does not get an education, her family never gets the opportunity to earn a higher income.

Finally, we analyze the case where educated workers are also biased. They underestimate the probability of intergenerational mobility – educated-unskilled workers are underconfident about their children's odds of getting a skilled job, while the skilled workers are overconfident. Uneducated-unskilled workers, like before, are in a behavioral trap, and thus, do not invest. Perceived expected benefit from an educational investment for each type of parent is now different and can not be unambiguously ranked. Due to this, the skilled workers, despite their higher incomes, no longer invest with certainty whenever educated-unskilled workers invest with a positive probability. When the degree of behavioral bias is low, i.e. the degree of association is high, and child affinity is not huge, the dynamics of investment and skill distributions in the economy are similar to that of the case with only behavioral trap or even the benchmark case. However, when the degree of association is low or child affinity is huge, behavioral bias produces interesting effects – the educated-unskilled parents may invest with a higher probability than the skilled parents. Moreover, there could be multiple equilibria when the behavioral bias is high and child affinity is moderate. In a dynamic framework, behavioral biases could cause aggregate fluctuations in investment and income, or lead to behavior-driven business cycles. Also, in comparison to the benchmark case, the presence of behavioral bias can cause both over and under investment in children's

correspondingly the child affinity parameter may be a fraction or greater than unity. (See Browning et al. (2014) pp. 106-120 for further discussion).

[6] This is because we assume production function is strictly concave, hence the income of skilled workers is inversely related to their mass.

education by educated parents. Interestingly, when the degree of child affinity is moderate, behavioral biases could *reduce the steady-state income inequality*. However, in this framework, the opportunity to earn a higher income is limited to the educated families – once a family becomes not educated, it never gets the opportunity to earn higher income as a skilled worker.

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).[7] 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 of the difficulty of a task (getting a skilled job) is assumed to be group-identity specific. Thus biases are generated through socio-economic background and address both positive, and negative self-images. This conforms with empirical evidence that social background, gender, wealth, etc. can influence an individual's self-confidence.[8] Filippin and Paccagnella (2012) develop a theoretical model to look at how self-confidence affects inequality. Unlike us, they consider a lack of self-confidence as an information problem while we look at it as a behavioral problem. In their model, agents do not know their true types and form beliefs about their (unknown) ability from their parents. Given these beliefs, they choose tasks, which vary in their difficulty levels, to learn their true ability. The Bayesian learning mechanism is based on observing success or failure in the tasks, given that the probability of success depends on the true level of ability as well as on the difficulty of the task. They find that "wrong" initial beliefs can have long-lasting effects on earnings, even if agents eventually learn their true ability. In our model, people have self-deception; they believe that the probability of change in economic status is low; and there is no learning. This absence of learning is prominently shown through experiments in Tversky and Kaheneman (2004). Tversky and Kaheneman (2004) find that one of the *heuristics* employed by the people in assessing probabilities and in predicting values is *anchoring* to some initial value and *adjustments* from that initial value are typically insufficient. Whether external constraints (like lack of information) or internal constraints are more likely to engender prejudiced decision-making is an open question.

Of late there has been a growing interest in investigating how psychological and behavioral constraints may lead to inequality and poverty trap.[9] Our paper contributes to this literature. The distinctive feature of our paper is that we do not assume any *intrinsic difference* between a rich and a poor person – all agents are *ex ante* homogeneous – it is rather their *circumstances* that trap them in poverty. This aligns with Mani et al. (2013) who find that the same individual behaves differently due to poverty (see Balboni et al. (2020) for a more general discussion).

[7]See Benoît and Dubra (2011), and Bénabou and Tirole (2016) for an overview.

[8]Deshpande and Newman (2007) find that graduating students from reserved (backward) categories have significantly lower occupational *expectations* than their non-reservation counterparts. Filippin and Paccagnella (2012) find that socio-economic backgrounds of students of Bocconi University play an important role in shaping their wage expectations. Barber and Odean (2001) find men are overconfident and invest more.

[9]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).

This paper is also related to the literature on aspiration. Ray (2006) develops a model of socially dependent *reference point* based aspirations. Bogliacino and Ortoleva (2013) assume that the aspiration threshold for everyone at any date is the average income at that date. In a very general set-up, Genicot and Ray (2017) model socially dependent aspirations as endogenous thresholds that differ across incomes. Mookherjee et al. (2010) develop location-based aspirations. Like us, Dalton et al. (2016) use internal constraints to explain poverty trap. Aspiration and efforts are jointly determined, but an individual takes the former as given and that is the main source of the poverty trap. All of these papers model aspiration as 'milestone utility' or 'reference point', while we do not assume any such *non-standard* utility function. Further, we assume the sources of identity to be education and income, instead of a non-economic variable such as location (as in Mookherjee et al. (2010)). The relative strength of how an individual identifies, via economic classes or location, is an empirical question, which is outside the scope of this paper.

The plan of the paper is as follows. The next section lays out the general framework of the economy. Then, 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 discusses the implications 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).

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}).[10] 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 $m_{ut} = 1$.

[10]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)*[11] *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.[12] 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

[11] 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 $A = 1$.

[12] 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

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

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.[13]

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.*

[13] 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[14] 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 7.1.

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* $1 < \underline{b}_s(\delta)$.

[14] 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.*

$$\langle \lambda_{ut}, \lambda_{st} \rangle = \begin{cases} \langle 0, 1 \rangle & \text{if } m_{st} \geq \bar{b}_s(\delta) \\ \langle 0, (0,1) \rangle & \text{if } m_{st} \in (\underline{b}_s(\delta), \bar{b}_s(\delta)) \\ \langle 0, 0 \rangle & \text{if } m_{st} \in [1, \underline{b}_s(\delta)]. \end{cases}$$

$$\langle \lambda_{ut}, \lambda_{st} \rangle = \begin{cases} \langle (0,1), 1 \rangle & \text{if } m_{st} > \underline{b}_u(\delta) \\ \langle 0, 1 \rangle & \text{if } m_{st} \in [\bar{b}_s(\delta), \underline{b}_u(\delta)] \\ \langle 0, (0,1) \rangle & \text{if } m_{st} \in [1, \bar{b}_s(\delta)). \end{cases}$$

$$\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 1, 1 \rangle \quad \forall\, m_{st} \geq 1.$$

0 $\underline{\delta}$ $\bar{\delta}$ δ

Low Moderate High

Figure 1: Characterization of the Equilibria in the Benchmark Case

We prove this in Appendix 7.2 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 7.3.

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 $\eta < \theta \in [0, 1)$.[15] Thus, the degree of association of an educated-unskilled[16] 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

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

[16]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}[17] 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 7.4.
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 $\delta\in(0,\bar{\delta})$, $\bar{b}_u(\delta)=\infty$.*

[17]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.*

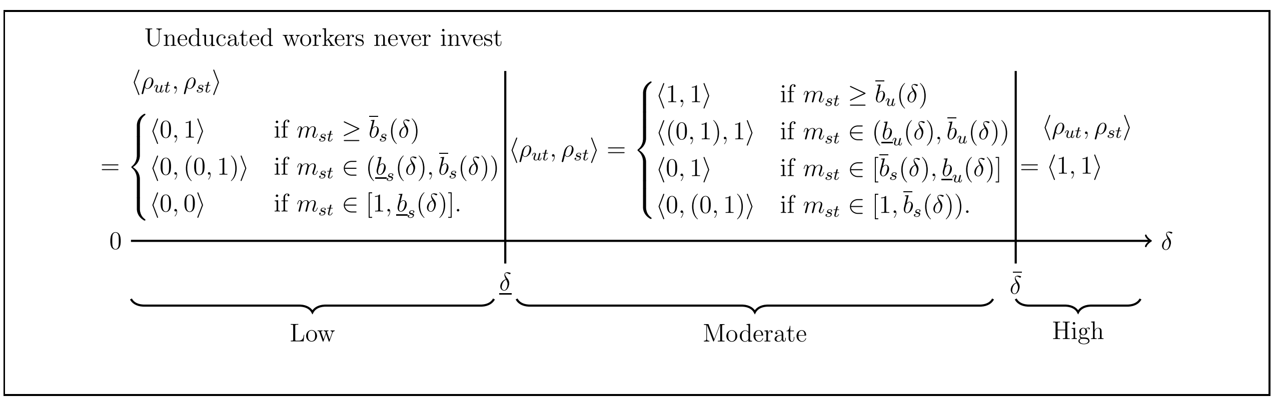


Figure 2: Characterization of the Equilibria with a Behavioral Trap

We prove this in Appendix 7.5 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 7.6.

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 $\theta(1-\beta)$[18] 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.[19] 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 $j \in \{u, s\}$,[20] invests with probability γ_{jt}. Then an educated-unskilled worker conjectures that the mass of skilled workers and their income would be

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

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

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

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

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

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

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

[18]Recall, $1-\beta$ is the probability with which an educated individual becomes an unskilled worker.
[19]$\theta\beta < \beta \leq 1-\theta(1-\beta)$.
[20]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

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

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

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

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 7.7.
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 $\underline{\mathbf{\Gamma}}_{\mathbf{s}}$ *for any m_{st}, γ_{st} is bounded below by $\underline{\gamma}_s(m_{st})$,*

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

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

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

The formal expressions are given in Appendix 7.8. 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. $\bar{\gamma}_s(m_{st})$.*

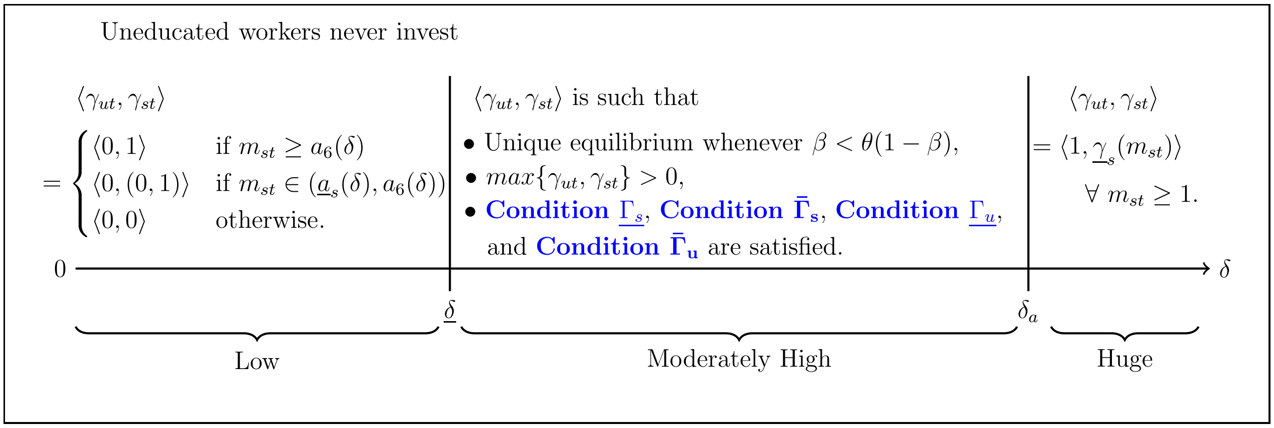
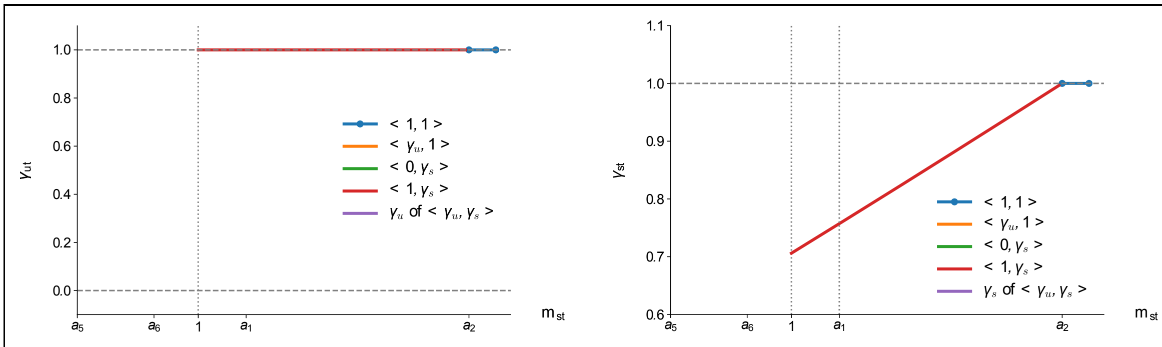


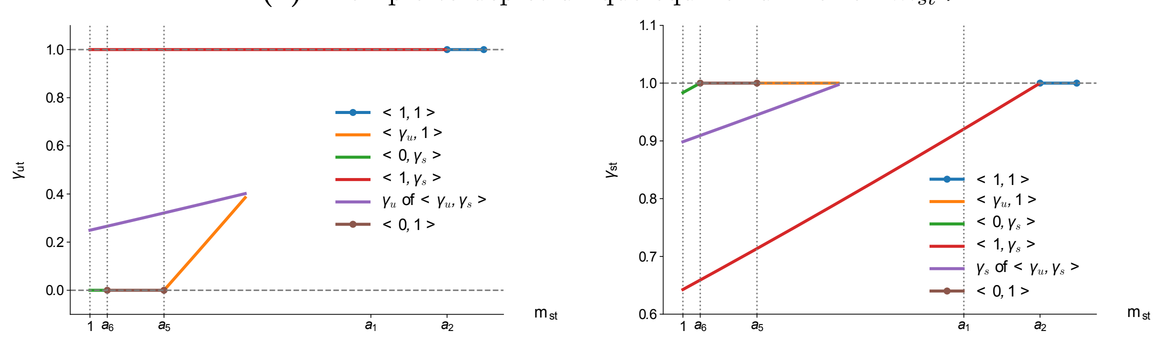
Figure 3: Characterization of the Equilibria with Behavioral Trap and Behavioral Bias

We prove this in Appendix 7.9 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.[21] 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

[21] 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 7.10.
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*:[22] 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.

[22]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 7.11.

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 7.12.
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)**.** $0 < \underline{\delta} < \bar{\bar{\delta}} < \delta_a$.

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 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 and they discount the experiences of those who they perceive as dissimilar. A behavioral trap would cause a poverty trap. When the society cares for children's future well-being, or child affinity is not low, the economy will have multiple steady states and the steady-state inequality is at least as high as in the benchmark. Behavioral biases may cause multiple equilibria. In a dynamic framework, it could result in fluctuations in investments and be a source of behavioral cycles. For economies with sufficiently high child affinity, the educated-unskilled parents may invest with a higher probability than the skilled parents. In some sense, biases beget parental aspirations. Under confident parents may overestimate the income of the skilled workers so much so that they find it incentive compatible to overinvest in their children's education. Behavioral biases may increase or decrease steady-state income inequality.

Behavioral trap limits intergenerational mobility. In such settings, our paper advocates for conditional cash transfer, free education, training. In reality, beliefs may not be as fatalistic but biased. In that case, a big push to incentivize poor people to invest in human capital would reduce poverty. There is ample evidence of the psychological benefits of poverty reduction. This paper brings this aspect into the realm of economic theory.

7 Appendix

7.1 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).

7.2 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. □

7.3 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. □

7.4 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

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

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

   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. □

7.6 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.). □

7.7 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).

7.8 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

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

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

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

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 $m_{st} \in [\underline{a}_s(\delta), a_6(\delta))$. □

7.9 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 7.1. *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 7.1. *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 7.1, 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 $\beta < \theta(1-\beta)$.

$\square$

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 7.1.

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. $\square$

7.10 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. □

7.11 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 $\lambda_{st} = \rho_{st}$.

7.12 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*. 4

Bandura, A. (1994). Self-efficacy. *Encyclopedia of Human Behavior*, 4:71–81. 1

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. 1

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. 4

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

Bénabou, R. and Tirole, J. (2016). Mindful Economics: The Production, Consumption, and Value of Beliefs. *Journal of Economic Perspectives*, 30(3):141–64. 4

Benoît, J.-P. and Dubra, J. (2011). Apparent Overconfidence. *Econometrica*, 79(5):1591–1625. 4

Bernard, T., Dercon, S., and Taffesse, A. S. (2011). Beyond Fatalism – An Empirical Exploration of Self-efficacy and Aspirations Failure in Ethiopia. 2

Bernheim, B., Ray, D., and Yeltekin, S. (2015). Poverty and Self-Control. *Econometrica*, 83(5):1877–1911. 1

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

Bogliacino, F. and Ortoleva, P. (2013). The Behavior of Others as a Reference Point. *Columbia Business School Research Paper*, (13-55). 5

Browning, M., Chiappori, P.-A., and Weiss, Y. (2014). *Economics of the Family*. Cambridge University Press. 3

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

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. 5

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. 4

Filippin, A. and Paccagnella, M. (2012). Family Background, Self-Confidence and Economic Outcomes. *Economics of Education Review*, 31(5):824 – 834. 1, 4

Genicot, G. and Ray, D. (2017). Aspirations and Inequality. *Econometrica*, 85(2):489–519. 5

Ghatak, M. (2015). Theories of Poverty traps and Anti-Poverty Policies. *The World Bank Economic Review*, 29(Suppl_1):S77–S105. 2, 4

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

Lindsey, A., King, E., Membere, A., and Cheung, H. K. (2017). Two Types of Diversity Training that Really Work. *Harvard Business Review*, 7:28. 2

Malmendier, U. and Tate, G. (2005). CEO Overconfidence and Corporate Investment. *The Journal of Finance*, 60(6):2661–2700. 1

Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, 341(6149):976–980. 4

Mookherjee, D., Ray, D., and Napel, S. (2010). Aspirations, Segregation, and Occupational Choice. *Journal of the European Economic Association*, 8(1):139–168. 5

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. 1, 5

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

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. 2, 11

Tversky, A. and Griffin, D. (2004). Endowment and Contrast in Judgements of Well-Being. In Shafir, E., editor, *Preference, Belief, and Similarity: Selected Writings by Amos Tversky*, chapter 38, pages 917–936. The MIT Press. 1

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. 4

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