Perceptions, Biases, and Inequality*

Dyotona Dasgupta† Anuradha Saha‡

October 2, 2020

Abstract

In an overlapping generations model, this paper looks at the effects of perceived self-efficacy beliefs, drawn from job and education networks, on human capital investments and skill distributions. Behavioral anomalies are such that adults underestimate the probability of intergenerational mobility. Uneducated-unskilled are imprisoned in a *behavioral trap* – they do not believe that an educated child from their community would get a skilled work, so they never invest. Educated parents suffer from a *behavioral bias* – skilled ones overestimate the chances of their educated children getting a skilled job while unskilled ones underestimate. Behavioral trap almost always causes poverty trap and also gives rise to multiple steady states, which can be ranked in terms of inequality. Depending on the degree of behavioral bias of the educated parents, steady state inequality could be lower or higher than that when they have correct beliefs. Behavioral bias may lead to multiple equilibrium paths to the steady state and even induce (poorer) educated-unskilled adults to invest with higher probability than (richer) skilled persons.

Keywords: Behavioral Welfare Economics, Human Capital Investment, Behavioral Bias, Poverty Trap, Behavioral Trap
JEL Codes: I3, D9, 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, Ratul Lahkar, Ajit Mishra, Dilip Mookherjee, Prabal Roy Chowdhury, 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.

†Centre for Development Economics, Delhi School of Economics. E-mail: dasgupta.dyotona@gmail.com

‡Ashoka University. 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)

This paper analyzes human capital investment and aggregate income distribution based on *perceived self-efficacy*. Perceived self-efficacy is an individual's belief about her own ability which is formed by her life-experiences and by the success or failure of people who she perceives similar to herself. The greater is the assumed kinship, the more persuasive is their influence on her perceived self-efficacy (Bandura, 1994). Tversky and Kaheneman (2004) document one of the *heuristics* employed by the people in assessing probabilities and in predicting values is *anchoring* on some initial value and *adjustments* from that initial value are typically insufficient. Indeed, empirical evidence shows that successful people display *overconfidence* (see Malmendier and Tate (2005) for example) while unsuccessful or poor persons display *under confidence* (see Bernard et al. (2011) for example). In a laboratory study and in a quasi-experimental set-up, Mani et al. (2013) find that poverty impedes cognitive function. The question we ask in this paper is what are the macroeconomic implications of such heterogeneous behavioral constraints.

In reality, an individual's choice problem is limited by both external and internal constraints. There is a large literature addressing the physical constraints such as credit market imperfection, non-convex technology (Banerjee and Newman (1993), Galor and Zeira (1993), Mookherjee and Ray (2002)) to explain persistent inequality and poverty trap. Of late there has been growing interest in investigating how psychological and behavioral constraints may lead to poverty trap (Dalton et al. (2016), Genicot and Ray (2017), more in Section 2). This paper contributes to the latter approach. Two distinctive features of our paper are as follows. First, we do not assume any *intrinsic difference* between a rich and a poor person, it is rather their *circumstances* that trap them in poverty. This conforms with the finding of 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). Second, while overconfidence and underconfidence are ubiquitous in reality, the literature mostly focuses on beliefs that reinforce a positive self-image.[1] In our model, biases are generated through socio-economic background and thus addresses both positive, and negative self-image.

In particular, we assume that perceptions are anchored on the decision maker's own experiences. She 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 more number of common attributes. In such a setting, unsuccessful people underestimate the probability of success of an investment as they relate more with unsuccessful persons. The opposite is true for successful people. Further, unsuccessful people overestimate the

[1]See Benoît and Dubra (2011), and Bénabou and Tirole (2016) for an overview. More in Section 2.

reward in case of success and successful people do the converse. These two opposing forces determine the investment decision of any individual. We analyze the equilibrium investment decisions of various individuals and their impacts on skill distribution and inequality in the economy.

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. They derive utility from their own consumption and the perceived expected income of their children.[2] 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. The exogenous probability of an educated child getting a skilled job is the same for all children, implying no intrinsic difference among them. Based on their education and jobs, adults can be classified into three groups – uneducated-unskilled, educated-unskilled, educated-skilled. These groups are endogenously formed and can change every period depending on their education and a random factor. Each adult decides whether to invest in her child's education or not.

We say that the agents are biased when they form a belief about the probability of success based on their education and job.[3] Adults from dissimilar groups identify less with each other. We capture this by a "degree of association" – lower the degree of association, more biased is the perceived probability of success. An unskilled worker underestimates the probability with which a child from her group, upon getting education, becomes a skilled worker. Due to higher dissimilarity, an uneducated-unskilled worker is more under confident than an educated-unskilled worker. Conversely, an educated-skilled worker under estimates 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.

The central objective of our paper is to compare the dynamics and the steady state properties of an economy with and without behavioral anomalies. At any initial skilled 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* (intergenerational altruism as in Ghatak (2015)). The parameter is non-negative,[4] time-independent, and common for all parents in an economy. In our characterization, we find that economic outcomes have distinct properties as per three child affinity ranges – low, moderate, and high.

In absence of any behavioral anomalies, in the benchmark case, if unskilled workers (with low incomes) invest in their children's education with a positive probability, then skilled workers invest with certainty. This is because of the concavity of utility function and that

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

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

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

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. Quite obviously, given any initial skilled income, the probability of investment by any parent weakly increases with child affinity. We also find, in the benchmark case, there is a poverty trap only when the degree of child affinity is low. Here, we define poverty trap as a situation where there exists a positive mass of families that never becomes rich. When child affinity is not low, at the steady state, the probability with which an adult from any family works as a skilled worker is a proper fraction. Thus, in the steady state any child becomes a skilled worker with a positive probability (and an unskilled worker with the complementary probability, keeping the mass of skilled workers constant). In this sense, there is intergenerational mobility at the steady state. The steady state probability of investment in a child's education and hence, the mass of skilled worker weakly decreases with decrease in the degree of child affinity.

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. The educated parents take this into account while making their investment decision. They are able to rationally foresee that with fewer educated (and hence skilled) adults in the next period, the future skilled incomes will be higher.[5] This incentivizes today's educated adults to invest with (weakly) higher probabilities than in the benchmark case. We find that the economic outcomes are same as in the benchmark case, for the low child affinity case. For moderate and high child affinity, due to non-investment of uneducated-unskilled workers, behavioral trap gives rise to multiple steady states. The multiple steady states can be ranked on the basis of 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. We also find once a child does not get education, her family never gets the opportunity to earn the higher income whereas in the benchmark case, that opportunity is available for all the families.

Finally, we analyze the case where educated workers are also biased and underestimate the probability of intergenerational mobility – educated-unskilled workers are under-confident about their children's odds of getting a skilled job, while the skilled workers are over-confident. Uneducated-unskilled workers, like before, are in a behavioral trap, and thus, do not invest. Perceived expected benefit from educational investment of each type of parent is now different. Due to this, the skilled workers, in spite of their high 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, the dynamics of investment and skill distributions in the economy are similar to that of the case with only behavioral trap. However, when the degree of association is not high, behavioral bias produces interesting effects – the educated-unskilled parents may invest with a higher probability than the skilled parents, and there could be multiple equilibria. In a dynamic framework, this could cause aggregate fluctuations in investment and income, or lead to behavior-driven business cycles. Also, in comparison to the benchmark case, the presence

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

of behavioral bias can cause both over and under investment in children's education by educated parents. Interestingly, when the degree of child affinity is moderate, over investment by educated parents could *reduce the steady state income inequality* compared to the case with only behavioral trap. However, even then the opportunity to earn 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. As in the previous case, the existence of behavioral trap leads to multiple steady states in economies with moderate or high child affinity. There is almost always a poverty trap in an economy.

In summary, this paper studies the long run implications of people forming beliefs and aspirations, based on *perceived* efficacy. Their perceptions are anchored on their own experiences. This framework captures that for different parameter values and initial conditions, individuals may maintain status-quo, be defeatist or even ambitious in their investment decisions.

The plan of the paper is as follows. Section 2 discusses relation to existing literature. Section 3 sets up the general framework of the model. Section 4 studies the benchmark case where there is no behavioral anomaly. Section 5 addresses two types of behavioral anomaly – Section 5.1 analyzes the case where only uneducated workers are under behavioral trap whereas Section 5.2 addresses the case where all types of workers are biased. Section 6 compares these cases and studies the welfare implications of behavioral anomaly. The final section concludes with some policy implications and future research. Main proofs are collected in respective Appendices within Section 8. The rest of the proofs are available in a Supplementary Appendix (available online).

2 Related Literature

This paper is related to various strands of literature. First, it brings the role of self-confidence in inter-generational investments. Traditionally, it was thought that "the most robust finding in the psychology of judgment is that people are overconfident" (De Bondt and Thaler, 1995). Overconfidence has been studied in the housing market (Case and Shiller, 2003), among CEOs (Malmendier and Tate, 2005), in financial investment (Biais et al., 2005), among others. It has been argued that self-confidence enhances motivation (Bénabou and Tirole, 2002), improves self-control (Bénabou and Tirole, 2004). Of late, it is found that overconfidence is not as a robust characteristic as it was thought to be (Clark and Friesen, 2009). Moore (2007) finds that people are under (over) confident when the task in consideration is difficult (easy). In our paper, the *perception* about the difficulty of a task (getting a skilled job) is assumed to be group-identity specific. There is empirical evidence that social background, gender, wealth etc. can influence individual's self-confidence. Deshpande and Newman (2007) find that graduating students from reserved (backwards) category have significantly lower occupational *expectations* than their non-reservation counterparts. Barber and Odean (2001) find men are overconfident and invest more.

This work also contributes to the emerging literature on behavioral poverty trap.[6] Present bias among poor people dampens savings which could generate poverty trap (Banerjee and

[6] 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).

Mullainathan (2010), Bernheim et al. (2015)). There is another thread of literature on how aspiration, or failure to achieve, could give rise to poverty trap. Appadurai (2004), Ray (2006) develop 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 which differ across incomes. Mookherjee et al. (2010) develop location based aspirations. The main difference with our paper is that we assume the sources of identity to be education and income, instead of a non-economic variable such as location. In our paper, economic status shapes beliefs about self-efficacy, which determine whether parents aim to equip their children to get skilled jobs or not. In psychology, there is a huge literature on self-efficacy, anchoring, and social identification (Gramzow et al. (2001), Cervone and Peake (1986), Van Veelen et al. (2016) to name a few). We use this literature as a basis to claim that individuals indentify with similar (income) groups. The relative strength of how an individual identifies, via economic classes or via location, is an empirical question, which is outside the scope of this paper. Like us, Dalton et al. (2016) considers internal constraints to explain poverty trap. They also model aspiration as reference points. Aspiration and efforts are jointly determined, but an individual takes former as given and this behavioral bias is the main source of poverty trap. In our paper, the source of internal constraint is the experience based perception about self-efficacy. There have been works on applications of a model with aspirations on political economy questions. In a citizen candidate model, Besley (2017) explores political economy of inequality where a child's aspiration can deliberately be affected by her parents. Ortoleva and Snowberg (2015) study the role of overconfidence in political behavior. Just like ours, in their paper, the citizens passively learn through their experiences, underestimate how correlated these experiences are and believe that their information is more precise than it actually is. However, the focus of their paper is to understand the process of evolution of (polarized) ideology whereas our aim is to provide an alternative explanation of income inequality and poverty trap.

3 Model

3.1 The Firms

We consider a single good economy comprised of a continuum of individuals of size 1. The good can be produced using only one input – labor. Labor is of two types – skilled (L_{st}) and unskilled (L_{ut}).[7] Let the production function of the skilled sector be 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 representative firms of the skilled and unskilled sectors, π_{st} and π_{ut}, are given by:

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

[7]In all notations, subscripts s and u designate skilled and unskilled workers, and subscript t denotes time.

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 the 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 $m_{ut} = 1$.

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

All claims with (S) notation are proved in the Supplementary Appendix. This one is shown in section SA.1.[8]

3.2 The Households

In a discrete time framework, consider an overlapping generations model with no population growth. An individual lives for two periods: first as a child and later as an adult. 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.[9] 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 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 expected income of her child respectively. Observe, the utility function is strictly increasing and strictly concave.

The investment decision on child's education is taken based on the *perceived* expected income of an educated child. It depends on the probability of her becoming a skilled worker upon getting 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. A parent's belief depends on her own experiences.

There is no inherent difference in the probability of getting a skilled job across educated children of different parent types. These odds are the same for all educated children, i.e., this probability is independent of her parent's type (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*

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

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

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^{ij}_{t+1} 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 worker would be L^{ij}_{st+1} and their income ω^{ij}_{st+1}. Thus, the *perceived* expected income of her educated child would be $E\omega^{ij}_{t+1} = p^{ij}_{t+1}\omega^{ij}_{st+1} + (1-p^{ij}_{t+1})\omega^{ij}_{ut+1} = p^{ij}_{t+1}\omega^{ij}_{st+1} + 1 - p^{ij}_{t+1}$. The expected income of an uneducated child is $E\omega^{ij}_{t+1} = 1$.

At any period t, a parent compares *perceived* expected utility from investing in her child's education with that from not investing and invests only when the former is (weakly) higher

$$
\begin{aligned}
& U^{ij}_t(\text{from investing in child's education}) \geq U^{ij}_t(\text{from not investing in child's education}) \\
\Rightarrow \quad & \frac{(m_{it}-\bar{s})^\sigma}{\sigma} + \delta\frac{[p^{ij}_{t+1}\omega^{ij}_{t+1} + (1-p^{ij}_{t+1})]^\sigma}{\sigma} \geq \frac{m^\sigma_{it}}{\sigma} + \frac{\delta}{\sigma} \\
\Rightarrow \quad & \delta\left[\frac{[p^{ij}_{t+1}\omega^{ij}_{st+1} + (1-p^{ij}_{t+1})]^\sigma}{\sigma} - \frac{1}{\sigma}\right] \geq \frac{1}{\sigma} - \frac{(1-\bar{s})^\sigma}{\sigma}. \qquad (2)
\end{aligned}
$$

The left hand side of the above inequality is the *perceived* expected net benefit from investing in child's education whereas the right hand side is the utility cost of making that investment.[10] 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.

4 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 worker 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

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

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

[10]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.

An equilibrium is denoted by $\langle \lambda_{ut}, \lambda_{st} \rangle$ which satisfies the features described in Section 3.2. 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 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 probability of investment of both types of workers (weakly) increase with increase in income of a skilled worker at that period.*

Intuitively, Part 1. is an immediate implication of our assumption of concave utility function. It implies the utility cost of investment decreases with income of a worker. The benefit from investment is the same for all the parents. Hence, whenever an unskilled worker invests, a skilled worker with higher income (see Observation 1) invests with certainty.

For Part 2., we refer to equation (3). When income of skilled workers increase (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. While the reason for the former is the concavity of utility function, the intuition behind the latter is a bit subtle. 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 are 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.

To see this positive relationship between the income of skilled workers of two consecutive periods, keeping investment decisions the same. Consider, two economies with two different income of skilled workers m^1_{st} and m^2_{st} with $m^1_{st} > m^2_{st}$ where investment decisions are the same, i.e. $\lambda^1_{st} = \lambda^2_{st} = \lambda_{st}$ and $\lambda^1_{ut} = \lambda^2_{ut} = \lambda_{ut}$. Then, we argue that m^1_{st+1} is greater than m^2_{st+1}. For that, first observe, income of the skilled workers is inversely related to the mass of skilled workers. So, the mass of skilled worker in Economy 1 is lower. Second, Part 1. of this observation tells us that $\lambda_{st} \geq \lambda_{ut}$, in fact, $\lambda_{st} = 1$, whenever $\lambda_{ut} > 0$. So, in Economy 2 higher mass of workers invest with higher probability (λ_{st} rather than λ_{ut}). Therefore, at $t + 1$, the mass of skilled workers would be lower in Economy 1 which implies m^1_{st+1} would be higher than m^2_{st+1}. Hence, the positive relationship.

The implications of the lemma are as follows. Part (a) implies that there can be five equilibria $\langle \lambda_{ut}, \lambda_{st} \rangle$ – three pure strategies and two mixed strategies. The pure strategy equilibria are: (i) both skilled and unskilled workers invest with certainty $\langle 1, 1 \rangle$, (ii) unskilled workers do not invest while skilled workers invest with certainty $\langle 0, 1 \rangle$, (iii) no worker invests $\langle 0, 0 \rangle$. The mixed strategy equilibria are (iv) unskilled workers invest with positive probability and skilled workers invest with certainty $\langle 0 < \lambda_{ut} < 1, 1 \rangle$, (v) unskilled workers do not invest while skilled workers invest with positive probability $\langle 0, 0 < \lambda_{st} < 1 \rangle$. Part (b) implies that the income of a skilled worker at any period is the state variable of that period.

The degree of child affinity[11] plays an important role in parent's investment decision. Next, we define three thresholds of child affinity which will be useful in further analyses.

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

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* $\underline{\delta} \equiv (1-\bar{s})^{\sigma}-1$*, or '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* $m_{st} \leq \underline{b}_u(\delta)$*.*

We provide the formal expressions of these thresholds of the state variable in Appendix A.1. Observe from Lemma 1 and equation (3), at $\underline{b}_u(\delta)$, the unique equilibrium is $\langle 0, 1 \rangle$. Similarly, the unique equilibrium at $\bar{b}_s(\delta)$ is $\langle 0, 1 \rangle$. And finally, at $\underline{b}_s(\delta)$, the unique equilibrium is $\langle 0, 0 \rangle$. 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) high, then* $\underline{b}_s(\delta) < \bar{b}_s(\delta) < \underline{b}_u(\delta) < 1$*, (ii) moderate, then* $\underline{b}_s(\delta) \leq 1 < \bar{b}_s(\delta) < \underline{b}_u(\delta)$*, and (iii) low, then* $\underline{b}_u(\delta) = \infty$ *and* $1 < \underline{b}_s(\delta) < \bar{b}_s(\delta) < \underline{b}_u(\delta)$*.*

Intuitively, as the parents become more child loving, the benefit from investment increases. At the thresholds, the parents are indifferent, to make them that, thus, the thresholds have to adjust accordingly. As observed earlier, the benefit from investment increases with the state variable. And, with increase in the state variable, the utility cost of investment decreases for the skilled workers and remains constant for the unskilled workers. Hence, the thresholds of state variable must decrease with increase in the degree of child affinity, such that the parents remain indifferent. The ranking of the thresholds, directly follows from Lemma 1. At $\underline{b}_u(\delta)$, unskilled workers are indifferent between investing and not investing. So from Part (a) of that lemma, we know that the skilled workers must invest with certainty at that threshold. The skilled workers invest with certainty as long as m_{st} is no less than $\bar{b}_s(\delta)$. Thus, from Part (b) of that lemma, $\underline{b}_u(\delta)$ is greater than $\bar{b}_s(\delta)$. As m_{st} falls below that threshold, again, from Part (b) of that lemma, we know that skilled workers no longer invest with certainty. They invest with a positive probability as long as m_{st} is no less than $\underline{b}_s(\delta)$. Hence, $\underline{b}_s(\delta)$ is lower than $\bar{b}_s(\delta)$.

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

Proposition 1. Characterization of the Equilibria

1. *Suppose child affinity is high. At the unique equilibrium all parents invest with certainty.*
2. *Suppose the degree of child affinity is moderate. The unique equilibrium is such that*
   a. $m_{st} > \underline{b}_u(\delta)$*: unskilled invest with a prob. s.t.* (3) *binds and skilled invest with prob. 1,*
   b. $m_{st} \in [\bar{b}_s(\delta), \underline{b}_u(\delta)]$*: unskilled workers do not invest and skilled invest with certainty,*
   c. $m_{st} < \bar{b}_s(\delta)$*: unskilled do not invest and skilled invest with a prob. s.t.* (3) *binds.*
3. *If and only if the degree of child affinity is low, there is a unique equilibrium such that*
   a. $m_{st} \geq \bar{b}_s(\delta)$*: unskilled workers do not invest and skilled workers invest with certainty,*
   b. $m_{st} \in (\underline{b}_s(\delta), \bar{b}_s(\delta))$*: unskilled do not invest and skilled invest with a prob. s.t.* (3) *binds,*
   c. $m_{st} < \underline{b}_s(\delta)$*: no worker invests.*

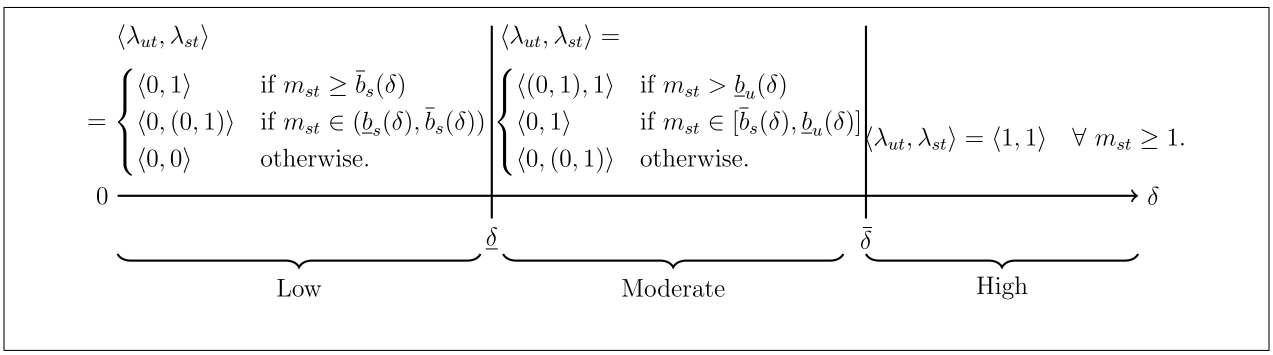

Figure 1: Characterization of the Equilibria in the Benchmark Case

We prove this in Appendix A.2 and depict the equilibria in Figure 1.

The intuition behind this proposition is, now, immediate. When the degree of child affinity is high, the parents care for their children so much that they invest at all relevant 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 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 degree of 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 the adult working as a skilled worker. Alternatively, there is no poverty trap if at any period, the probability with which a family becomes rich is positive.

Proposition 2. Dynamics and 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 worker 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 invest with a positive probability 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 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 converges to zero, correspondingly their income converges to infinity.*

   b. *no higher than $\underline{b}_s(\delta)$, then the economy immediately approaches a 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 A.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, i.e. the adult works as a skilled worker with probability β. 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 with the decrease in the degree of child affinity.

When child affinity is moderate, the unskilled workers no longer invest with certainty. Remember, the skilled incomes 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 is so lucrative that the unskilled workers invest with a positive probability. In this economy, at $\underline{b}_u(\delta)$ or higher incomes all workers have the incentives to invest in their children's education. 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 positive probability. 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 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 decrease in child affinity is quite obvious.

Given the state variable, higher child affinity implies greater probability of investment which lowers future skilled incomes. Hence, within the range of moderate child affinity, the steady state inequality is decreasing in the child affinity parameter.

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.

5 Behavioral Anomaly

Here, 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. Similarity between groups increases with addition of common features (following Tversky (2004) pp. 10-11). An individual feels less connected with more dissimilar groups. We capture this through *"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 $\eta < \theta \in [0, 1]$.[12] Thus, the degree of association of an educated-unskilled[13] worker with an educated-skilled worker is θ and that of an unskilled worker with a skilled worker is θ as education is necessary to become a skilled worker.

While forming the beliefs about the probability of her educated child becoming a skilled 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 degree of association. Recall, the 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.

5.1 Bias 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. The immediate implication of η being zero is

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

The educated parents take this into account and invest accordingly. Let the probability with which an educated-unskilled worker invests be ρ_{ut}[14] and that for a skilled worker be ρ_{st}.

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

[13] 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.

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

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

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

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 3.2. Like the benchmark case,

Observation 4. *Consider 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 behavioral trap, there does not exist any degree of child affinity where all parents invest. We define the following new threshold of state variable, the income of a skilled worker.

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 period t, 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 B.1. The thresholds stated in Definition 2 continue to be relevant here. At these thresholds, even in the benchmark case, unskilled workers do not invest – they invest only when the state variable is higher than the largest of the thresholds, $\underline{b}_u(\delta)$. So the effect of behavioral trap, via non-investment by uneducated workers, does not affect these thresholds. The following observation documents some features of the new threshold:

Observation 5 (S)**.** *(i) $\bar{b}_u(\delta)$ is decreasing in δ. (ii) $\bar{b}_u(\bar{\delta})=A\beta^{-(1-\phi)}$, when the degree of child affinity is moderate $\bar{b}_u(\delta)=\beta^{-(1-\phi)}\underline{b}_u(\delta)$ and when child affinity is low, then $\bar{b}_u(\delta)=\infty$.*

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. *The uneducated workers never invest.*
2. *If child affinity is high, at the unique equilibrium all educated workers invest with certainty.*
3. *Suppose the degree of child affinity is moderate. The unique equilibrium is such that*
   a. *$m_{st}\geq\bar{b}_u(\delta)$: all educated workers invest with probability 1,*
   b. *$m_{st}\in(\underline{b}_u(\delta),\bar{b}_u(\delta))$: edu-unskilled invest with prob. s.t. (4) binds, skilled with prob. 1*
   c. *$m_{st}\in[\bar{b}_s(\delta),\underline{b}_u(\delta)]$: edu-unskilled do not invest, and skilled invest with prob. 1,*
   d. *$m_{st}<\bar{b}_s(\delta)$: edu-unskilled do not invest and the skilled invest with a prob. s.t. (4) binds.*
4. *Suppose child affinity is low, unique equilibrium is such that*
   a. *$m_{st}\geq\bar{b}_s(\delta)$: edu-unskilled do not invest and skilled workers invest with certainty,*

b. $m_{st} \in (\underline{b}_s(\delta), \bar{b}_s(\delta))$*: edu-unskilled do not invest, skilled invest with a prob. s.t.* (4) *binds,*

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

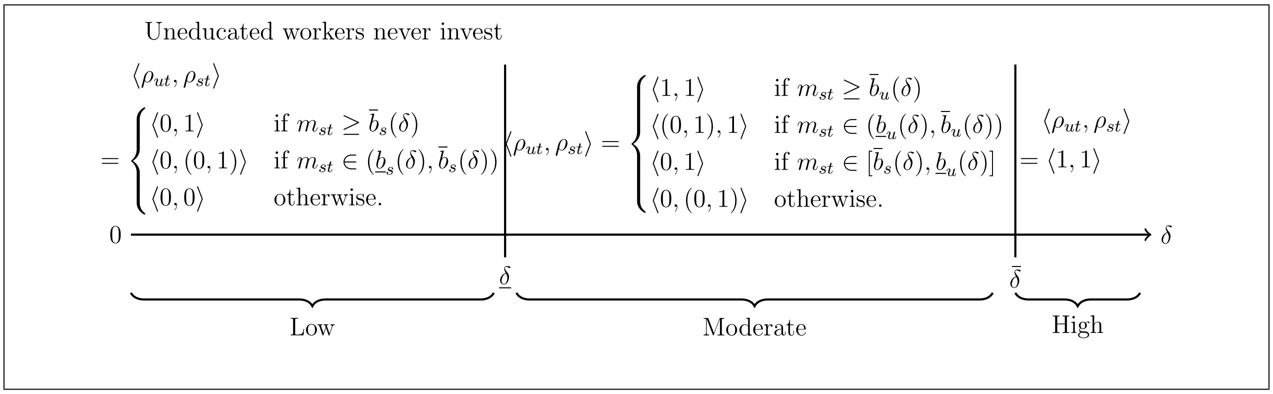

Figure 2: Characterization of the Equilibria with a Behavioral Trap

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

Observe, when child affinity is moderate, educated-unskilled workers invest with weakly higher probability than they would have in the benchmark case – they invest with strictly higher probability when they invest with a positive probability in the benchmark, and mass of uneducated workers is positive. We discuss the intuition now. Consider first, $m_{st} \in (\underline{b}_u(\delta), \bar{b}_u(\delta))$, then $\lambda_{ut}, \gamma_{ut} \in (0,1)$. From Equations (3), (4), Observation 2 and 5 we have

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

Thus, when the mass of uneducated workers is positive, $N_{et} > 0$ and they do not invest due to the behavioral trap, their non-investment is compensated by over investment of educated-unskilled workers, so that the benefits from investment remain the same. Now, it is clear from the definition that at $\bar{b}_u(\delta)$ we have $\lambda_u^*(1 - N_e) = (1 - \lambda_u^*)(1-\beta)N_e$ where λ_u^* is the steady state probability with which unskilled workers invest and N_e is the mass of educated workers corresponding to the skilled income level $\bar{b}_u(\delta)$. This implies at each δ, there is a unique value of N_{et} corresponding to $\bar{b}_u(\delta)$. If N_{et} is higher than N_e then educated-unskilled workers would not invest with probability 1, and if lower, then the benefit from investment of an unskilled worker would be strictly higher than her cost. Yet, due to the behavioral trap, uneducated workers would not invest. This gives rise to multiple steady states. The steady states can be ranked according to inequality – the difference between the income of a skilled worker and that of an unskilled worker. When child affinity is high, the steady state where $m_s^* = \bar{b}_u(\bar{\delta}) = A\beta^{-(1-\phi)}$, is called the 'least unequal steady state'. When child affinity is moderate, the steady state where $m_s^* = \bar{b}_u(\delta)$ is termed the 'least unequal steady state'.

Proposition 4. Dynamics and 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 certainty. 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 certainty. The steady state income of a skilled worker is $m_s^* = m_{st}$.

c. moderate and $m_{st} < \bar{b}_u(\delta)$, at least one type of workers do not invest with certainty. 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 converges to zero. The income of a skilled worker converges to infinity.

e. low and $m_{st} \leq \underline{b}_s(\delta)$, the unique steady state is immediately reached where no one invests.

3. Steady States Properties: 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 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 B.3.
Interestingly, with a behavioral trap, as uneducated workers never invest, it is not possible to have mixed strategies being played in any steady state. As that would decrease the mass skilled workers in the next period. For a similar reason, with a behavioral trap, we have multiple steady states – even if there are a few educated families in an economy which makes the income of skilled workers large, due to their behavioral imprisonment, uneducated workers do not invest and the economy stays there forever.
Next, we discuss the case where the educated parents are also biased.

5.2 Bias via Education & Job: Behavioral Trap & Behavioral Bias

All parents, here, are biased. An educated-unskilled worker believes with probability $\theta\beta$ an educated child from her group becomes a skilled worker. A skilled worker believes that with probability $\theta(1-\beta)$ 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 over confident.[15] 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

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

t, a worker of type j, where $j \in \{u, s\}$,[16] invests with probability γ_{jt}. Then an educated-unskilled worker conjectures that the mass of skilled workers and their income would be

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

recall, N_{et} is the mass of educated workers, $(1-\beta)N_{et}$ is mass of educated-unskilled workers 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

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

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

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

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

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

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) estimates the mass of future skilled workers and hence, over (under) estimates their income. At any period t, an educated-unskilled worker invests with probability γ_{ut} if and only if

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

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

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

The utility cost of investment is lower for a skilled worker. But, the *conjectured* net benefits from investment cannot be ranked, so Part (a) 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 $\gamma_{st} > 0$.*

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

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 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 thresholds, the mentioned $\langle \gamma_{ut}, \gamma_{st} \rangle$ may not be an equilibrium may or or there may exist many other equilibria at that threshold.

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

- *suppose $\langle \gamma_{ut}, \gamma_{st} \rangle$ is an equilibrium. $\underline{\mathbf{a}}_{\mathbf{s}}(\delta)$ 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 $\mathbf{a_6}(\delta)$ 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 $\mathbf{a_4}(\delta)$ 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 $\mathbf{a_2}(\delta)$ is the minimum value of the state variable at which skilled workers invest with certainty,*
- *suppose the skilled workers invest with probability 1, then $\mathbf{a_5}(\delta)$ is the maximum value of the state variable at which educated-unskilled workers do not invest,*
- *suppose the skilled workers do not invest, then $\mathbf{a_3}(\delta)$ 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* $\mathbf{a_1}(\delta)$ *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 C.1.
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*) $\underline{a}_s(\delta) < a_6(\delta) < a_2(\delta)$, (*iii*) $\underline{a}_s(\delta) < a_4(\delta) < a_2(\delta)$, *and* (*iv*) $a_4(\delta) > a_6(\delta)$ *if and only if* $\theta(1-\beta) > \beta$.
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*) $\underline{b}_u(\delta) < a_5(\delta)$.

Next, we discuss the intuition of this lemma. But before that a word about the parametric condition $\theta(1-\beta) > \beta$. This and the converse of it are important in optimal decisions of both types of workers, and hence in the characterization of the equilibria. N_{et} denotes the mass of educated workers, and due to the presence of behavioral trap, it also represents the mass of potential parents who may invest. Among them $(1-\beta)N_{et}$ is the mass of educated-unskilled workers and βN_{et} is that of the skilled workers. Now, $\theta(1-\beta) > \beta$ implies, first, the degree of association is higher than the proportion of skilled and unskilled workers among the educated individuals. Second, when $\gamma_{ut} = \gamma_{st}$

$$\underbrace{\theta\beta(1-\beta)N_{et}}_{\text{educated-unskilled workers' conjectured mass of skilled workers in the next period coming from their group}} > \underbrace{\beta^2 N_{et}}_{\text{educated-unskilled workers' conjectured mass of skilled workers coming from the group of skilled workers}} \quad \text{and} \quad \underbrace{\beta[1-\theta(1-\beta)]N_{et}}_{\text{skilled workers' conjectured mass of skilled workers in the next period coming from their group}} < \underbrace{\beta(1-\beta)N_{et}}_{\text{skilled workers' conjectured mass of skilled workers coming from the group of educated-unskilled workers}}$$

So, if both types of workers invest with the same probability, then workers of both types conjecture that the contribution of the educated-unskilled workers to the mass of skilled workers at the next period is higher than that contribution of the skilled workers.

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 investment of any type of worker is decreasing in the conjectured mass of skilled workers and is increasing with 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 primitives of the definitions. Higher is that mass higher is the threshold. For example, at the primitives of $\underline{a}_s(\delta)$ the conjectured mass of skilled worker is zero whereas that at $a_6(\delta)$ is positive. Hence, $a_6(\delta)$ must be higher than $\underline{a}_s(\delta)$ – the skilled workers are willing to invest at the primitive of $\underline{a}_s(\delta)$ when the state variable is low as the benefit is infinity whereas the state variable must be higher to make them invest at the primitive of $a_6(\delta)$. The other rankings depicted in Part 2. (ii), (iii) and those in Part 3. (iii) follow from similar reasoning. The ranking between $a_4(\delta)$ and $a_6(\delta)$ follows from the discussion on $\theta(1-\beta) > \beta$ and this logic. The same goes for the ranking between $a_3(\delta)$ and $a_5(\delta)$.

Second, observe the state variable, by assumptions, cannot be less than 1. So, if we find that any threshold of the state variable is less than 1, then when the primitive of the definition is satisfied, 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, irrespective of the value of state variable, the educated-unskilled workers optimally invest with certainty. Finally, the intuition behind the ranking between $a_4(\delta)$ and $a_1(\delta)$ follows directly from Lemma 3.

Now, 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* $\langle \gamma_{ut}, \gamma_{st} \rangle$

Condition $\underline{\boldsymbol{\Gamma}}_{\mathbf{s}}$. *If* $m_{st} \geq a_4(\delta)$, *then* γ_{st} *is bounded below by* $\underline{\gamma}_s(m_{st})$.

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

Condition $\underline{\boldsymbol{\Gamma}}_{\mathbf{u}}$. *If* $m_{st} \geq a_5(\delta)$, *then* γ_{ut} *is bounded below by* $\underline{\gamma}_u(m_{st})$.

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

The formal expressions are given in Appendix C.2. There we also show that $\underline{\gamma}_s(m_{st})$ 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

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

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

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

A word about how we get these boundary conditions. Let us consider **Condition $\underline{\boldsymbol{\Gamma}}_{\mathbf{s}}$.** and **Condition $\bar{\boldsymbol{\Gamma}}_{\mathbf{s}}$.**. To get the lower bound $\underline{\gamma}_s(m_{st})$, we use the fact that the probability of

investment of educated-unskilled workers can at most be 1, and for the upper bound $\bar{\gamma}_s(m_{st})$, the fact that it is at least zero. With these facts, Observation 7, and the definitions of the thresholds of the state variable addressing the optimum decision of the skilled workers we get these lower and upper bounds. For example, from the definition of $a_4(\delta)$ and Observation 7, we know that even when uneducated-unskilled workers invest with probability 1, at any $m_{st} > a_4(\delta)$, the skilled workers optimally invest with a positive probability. And from equation (6), observe that probability is unique. We call this unique probability $\underline{\gamma}_s(m_{st})$. Clearly, at $a_4(\delta)$, $\underline{\gamma}_s(m_{st})$ is zero. At any $m_{st} > a_4(\delta)$, the skilled workers invest with no less than $\underline{\gamma}_s(m_{st})$. Similar intuition gives us **Condition $\underline{\boldsymbol{\Gamma}}_{\mathbf{u}}$.**, and **Condition $\bar{\boldsymbol{\Gamma}}_{\mathbf{u}}$.**.
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 $\langle \gamma_{ut}, \gamma_{st} \rangle$: edu-unskilled workers invest with prob. 1, skilled invest with $\underline{\gamma}_s(m_{st})$ as in* **Condition $\underline{\boldsymbol{\Gamma}}_{\mathbf{s}}$.**.

3. *Suppose if 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. Further, if $a_1(\delta) \leq a_2(\delta)$ then $\gamma_{ut} = \underline{\gamma}_u(m_{st})$ and $\gamma_{st} = \underline{\gamma}_s(m_{st})$ $\forall$ $m_{st} \geq \min\{a_1(\delta), a_2(\delta)\}$. If $a_1(\delta) > a_2(\delta)$ then $\gamma_{ut} = \underline{\gamma}_u(m_{st})$ and $\gamma_{st} = \underline{\gamma}_s(m_{st})$ $\forall$ $m_{st} \geq a_5(\delta)$, and $\gamma_{ut} = 0$ and $\gamma_{st} = \underline{\gamma}_s(m_{st})$ $\forall$ $m_{st} \in [a_2(\delta), a_5(\delta))$, where $\underline{\gamma}_s(m_{st})$ and $\underline{\gamma}_u(m_{st})$ as in* **Condition $\underline{\boldsymbol{\Gamma}}_{\mathbf{s}}$.** *and* **Condition $\underline{\boldsymbol{\Gamma}}_{\mathbf{u}}$.**.

   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 positive probability and* **Condition $\underline{\boldsymbol{\Gamma}}_{\mathbf{s}}$.**, **Condition $\bar{\boldsymbol{\Gamma}}_{\mathbf{s}}$.**, **Condition $\underline{\boldsymbol{\Gamma}}_{\mathbf{u}}$.**, *and* **Condition $\bar{\boldsymbol{\Gamma}}_{\mathbf{u}}$.** *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 unique equilibrium $\langle \gamma_{ut}, \gamma_{st} \rangle$ is such that*

   a. *$m_{st} < a_6(\delta)$: edu-unskilled workers do not invest and skilled invest with prob. $\bar{\gamma}_s(m_{st})$,*

   b. *$m_{st} \geq a_6(\delta)$: edu-unskilled workers do not invest and skilled invest with certainty.*

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

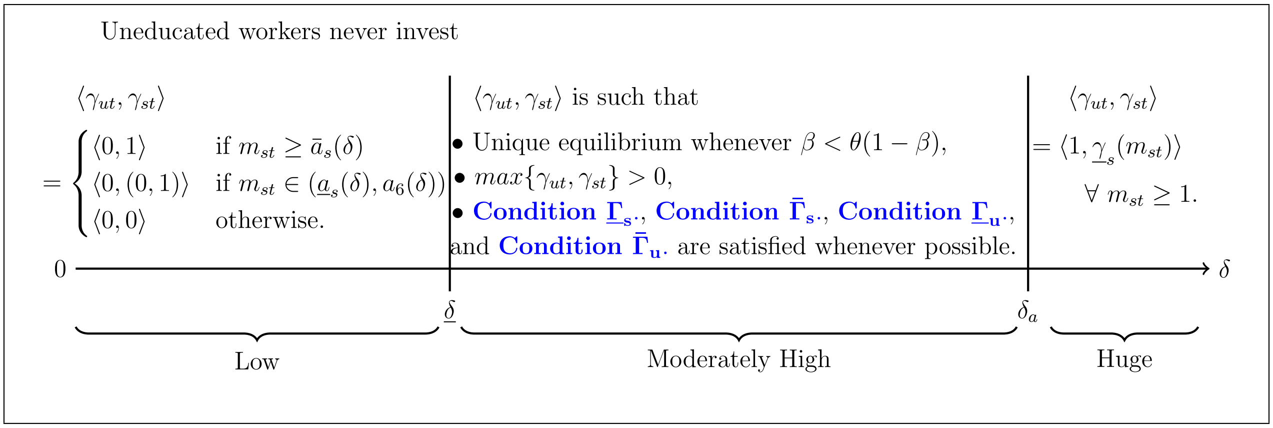

Figure 3: Characterization of the Equilibria with Behavioral Trap and Behavioral Bias

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.[17] Figure 4 shows unique equilibrium at any m_{st} when $\theta = 0.4$ and $\beta = 0.7$, and Figure 5 provides an example of multiple equilibria for the same values of θ and β. In this example, the difference in the two plots stem from differences in the values of ϕ and δ.

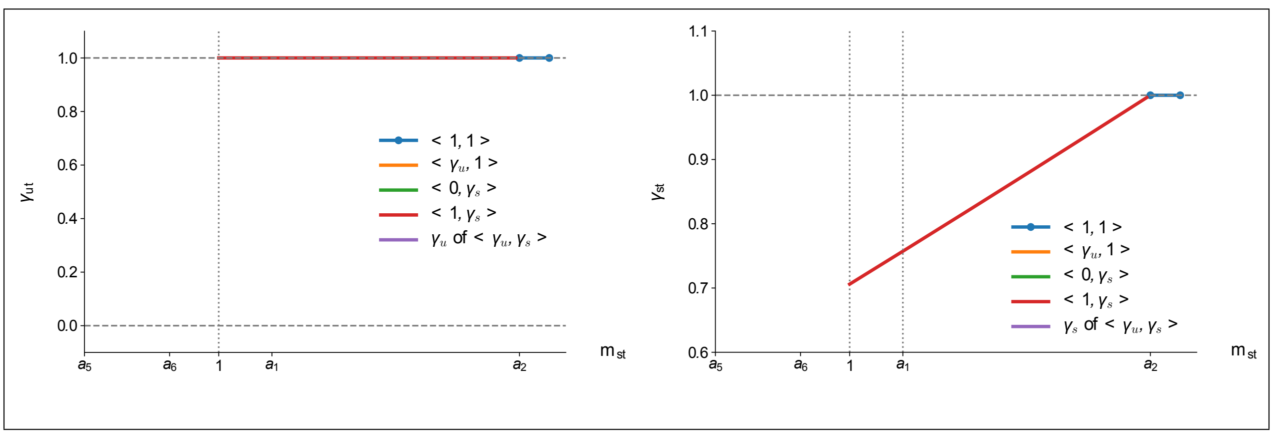

Figure 4: Example to depict unique equilibrium for all $m_{st} > 1$ and $\theta(1-\beta) < \beta$.

We, now, analyze the dynamics and steady states. Here also, as the uneducated workers 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 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 certainty. Steady state income of a skilled worker is initial income $m_s \equiv m_s^*$,*

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

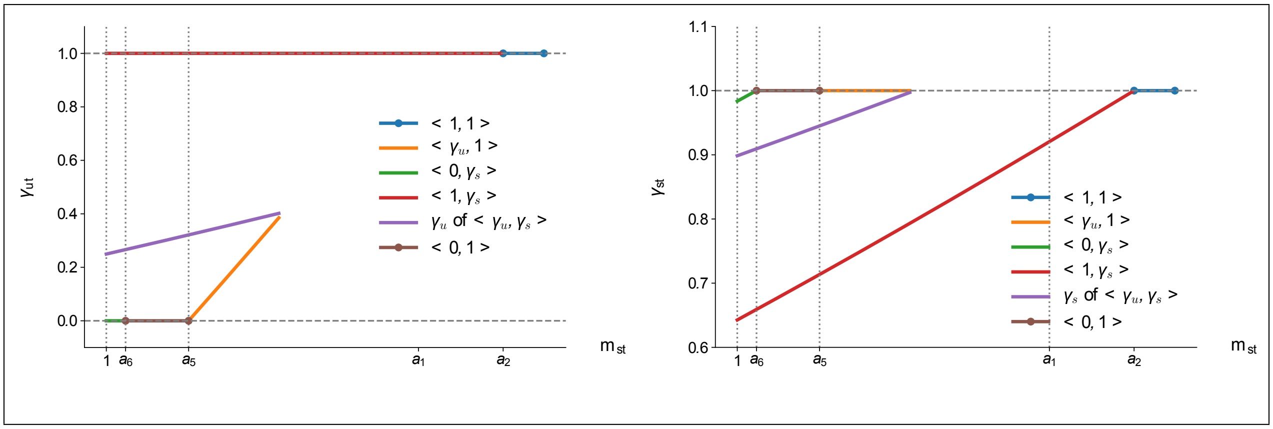

Figure 5: Example to depict multiple equilibria for some $m_{st} > 1$ and $\theta(1-\beta) < \beta$.

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. Their 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 converges to zero, and correspondingly the income of a skilled worker converges 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 decrease in child affinity,*

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

We prove in Appendix C.4.
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 5.1.
Next, we discuss the welfare implications of the behavioral anomalies studied so far.

6 Comparison: Implications of Behavioral Trap & Behavioral Bias

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

(i) *Distortions due to over and under investment*:[18] In the benchmark case, at any equilibrium when the unskilled workers invest with positive probability, the skilled workers invest with probability one. This feature holds for the economy with behavioral trap but, when both behavioral trap and behavioral bias are present, at any equilibrium, the skilled workers may under invest due to the *over* investment of educated-unskilled

[18]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.

workers. We explore this *crowding out* of investment due to *over* investment. Also, a parent *regrets* for over or under investing *ex post*, that is if she had believed the true probability of getting a skilled job she would have not under or over invested respectively.

(ii) *Poverty Trap and Inequality at the steady states*: We discuss the existence of the poverty trap and compare the mass of families under poverty trap whenever possible. We compare the inequality at the 'least unequal steady states' with that at the unique steady at the benchmark case. Even when the inequalities are equal, we can rank in terms of *opportunities* per se.

6.1 Implications of Behavioral Trap Only

We begin with the comparison between the benchmark case and 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.

Observation 8. *The distortions in the 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. *with behavioral trap educated-unskilled workers invest with 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 in the benchmark case and in the case with behavioral trap.*

We prove this in Appendix D.1.
When child affinity is moderate and the mass of skilled workers is sufficiently low, the expected future skilled incomes is sufficiently high for educated-unskilled parents to invest in their children's education. Otherwise, there are too many skilled workers in the economy, such that no unskilled workers invest. Hence, the existence of a behavioral trap does not affect the equilibrium strategies. When the degree of child affinity is high, all educated workers invest with certainty, while when child affinity is low no unskilled worker invests. The skilled workers invest with the same probability in the benchmark and the behavioral trap cases.

(ii) *Poverty Trap and Inequality at the steady states*: When the degree of child affinity is not low, there is no poverty trap in the steady state of the benchmark case, whereas with the behavioral trap, whenever an economy starts with a positive mass of uneducated workers, there is a poverty trap in the steady state. When the degree of child affinity is low, in the steady state all adults are uneducated and work as unskilled workers. Here, due both low incomes and extreme behavioral anomaly, the uneducated persons are in a poverty trap. We also find,

Observation 9. *When the degree of child affinity is not low, the inequality at the steady state is (weakly) higher when there is a behavioral trap.*

We prove this in Appendix D.2.

When all workers in the economy are educated and child affinity is high, all these workers invest in their children's education and hence the steady state inequality in the benchmark case and that at the 'least unequal steady state' in the behavioral trap case is equal. But for moderate child affinity, when all workers in the economy are educated, for sufficiently high state variable ($m_{st} > \bar{b}_u(\delta)$) all these workers invest in their children's education. Here again, the steady state inequality in the benchmark case and that at the 'least unequal steady state' in the behavioral trap case is equal. The steady state income of a skilled worker is $A\beta^{-(1-\phi)}$ or $\beta^{-(1-\phi)}\underline{b}_u(\delta)$ depending on the child affinity is high or moderate respectively. When there are some uneducated workers, with fewer skilled workers, their income is higher, and hence, the inequality is higher.

In the benchmark case, every family has a positive probability of investing in their children's education. Hence, every family can have a descendant who is a skilled worker. However, when there is a behavioral trap, only the educated families have this opportunity. Observe, this is not due to any external constraint, but only internal constraints.

6.2 Implications of Behavioral Trap & Behavioral Bias

Let us now consider the economy where educated workers are also biased. We compare the findings of Section 5.2 with those of Section 4 as well as Section 5.1. First, we cumulate the thresholds of child affinity defined in Section 4 and Section 5.2.

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

(i) *Distortions:*

For huge child affinity, i.e. $\delta \geq \delta_a$, all parents invest with certainty in the benchmark case, while all *educated* parents invest with certainty in the behavioral trap economy. The addition of behavioral bias, makes skilled workers over confident and lowers their incentives to invest with certainty. Now, skilled workers invest with certainty only for skilled incomes greater than $a_2(\delta)$. Thus, behavioral bias produces one distortion – for skilled income less that $a_2(\delta)$ skilled workers under invest. There is no over investment or crowding out – as all workers who believe the true probability of becoming a skilled worker β would invest with certainty.

When the degree of child affinity is high, i.e. $\delta \in [\bar{\delta}, \delta_a)$, investments in the benchmark and the behavioral trap are the same as the in the case with huge child affinity. With the inclusion of behavioral bias in educated workers, now all educated workers invest with certainty only for sufficiently high state variable, i.e. greater than $\max\{a_1, a_2\}$. Now, both educated-unskilled and skilled workers can under invest at lower state variable. Here also there is no over investment or crowding out because of the reason discussed in the previous case.

When the degree of child affinity is moderate, presence of behavioral bias may influence the educated parents to over or under invest. In general, educated-unskilled workers under invest (or over) invest at any m_{st} where γ_{ut} is respectively lower (or higher) than ρ_{ut}. Similarly, skilled workers under (or over) invest as γ_{st} is lower (or higher) than ρ_{st}. *Ex post*, they regret for this under (or over) investment. There is crowding out of investment whenever $\gamma_{ut} > \rho_{ut}$ and $\gamma_{st} < \rho_{st}$. Computationally we can find cases where crowding out occurs as well as cases where it does not occur when the economy has moderate child affinity.

When the degree of child affinity is low $\delta < \underline{\delta}$, unskilled workers never invest in any of the three models. When both behavioral trap and behavioral trap are present, the skilled workers over invest or under invest depending on $a_6(\delta)$ is lower or greater than $\bar{b}_s(\delta)$ respectively. Precisely, skilled workers over invest at $m_{st} \in [a_6(\delta), \bar{b}_s(\delta))$ and under invests at $m_{st} \in [\bar{b}_s(\delta), a_6(\delta))$. There is no crowding out as no unskilled workers invest.

(ii) *Poverty Trap and Inequality at the steady state:* When the degree of child affinity is high and the economy has only educated workers, then there is no poverty trap in the benchmark case or the behavioral trap case. However, when both behavioral trap and behavioral bias are present in this case, there is a poverty trap for the parametric condition, $\max\{a_1(\delta), a_2(\delta)\} > A\beta^{-(1-\phi)}$. Thus, inclusion of behavioral bias with behavioral trap may also beget a poverty trap. This is an especially powerful result. Due to behavioral bias, educated adults will invest with certainty in their children's education only for sufficiently high skilled incomes. In the interim, however, when these adults invest with less than unit probability, some children become uneducated and (due to behavioral trap) remain uneducated and unskilled forever. Thus, in this scenario behavioral bias gives rise to a poverty trap.

We know that in general the existence of behavioral trap creates a poverty trap as well as multiple steady states. In addition, the steady state income inequality is weakly higher in the behavioral trap. Only when the degree of child affinity is moderate, the behavioral bias may *decrease* this steady state income inequality. This also implies that in the presence of both behavioral trap and behavioral bias, the mass of families under poverty trap could be higher, equal or lower compared to the case when only behavioral trap exists, depending on $\beta(\underline{b}_u(\delta)/A)^{-(1-\phi)} \gtreqless \max\{a_1(\delta), a_2(\delta)\}$.

Due to behavioral bias of the educated workers, there are both over and under investments, so we find the following

Observation 11. *In the presence of both behavioral bias and behavioral trap:*

1. *When the degree of child affinity is high $\delta \geq \bar{\delta}$, the steady state inequality in the benchmark case is equal to that at the 'least unequal steady state' when both behavioral trap and behavioral bias are present only if the economy starts with all educated workers and $\max\{a_1(\delta), a_2(\delta)\} \leq A\beta^{-(1-\phi)}$. Otherwise, compared to the benchmark and the behavioral trap only cases, the inequality is higher due to behavioral bias and behavioral trap.*

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

economy with only behavioral bias may increase, decrease or have no effect on the least unequal steady state.

3. *When the degree of child affinity is low, there is no inequality at the steady states in both the cases.*

Only when the educated-unskilled workers over invest due to their bias, the behavioral bias may lower inequality in the least unequal steady state. Though educated-unskilled regret *ex post*, but in this scenario they have a greater opportunity to become rich at the steady state.

7 Conclusion

Homo Sapiens, unlike *Homo Economicus* does get affected by experiences of her own or that of individuals she perceives as similar. We provide a behavioral explanation of inequality where individuals form belief about efficacy of their children based on their own experiences and they discount the experiences of those who they perceive as dissimilar. 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 causes multiple equilibria. In a dynamic framework, it could results in fluctuations in investments and be a source of behavioral cycles. For economies with sufficiently high child affinity, educated unskilled parents may invest with a higher probability than skilled parents. In some sense, biases beget parental aspirations. Under confident parents may over estimate the skilled incomes so much so that they find it incentive compatible to over invest 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 would also incentivize poor people to invest in human capital which would reduce poverty. There is ample evidence of psychological benefits of poverty reduction. This paper brings this aspect in the realm of economic theory.

8 Appendices

A Appendix for Benchmark Case

A.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

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

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

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

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

$$\underline{b}_u(\delta) : \quad \delta \left[\frac{[\beta^\phi \underline{b}_u + 1 - \beta]^\sigma}{\sigma} - \frac{1}{\sigma} \right] = \frac{1}{\sigma} - \frac{(1 - \bar{s})^\sigma}{\sigma}. \tag{A.3}$$

A.2 Proof of Proposition 1

1. 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 $\forall\ \delta \geq \bar{\delta}$.

2. Now consider the case $\delta \in [\underline{\delta}, \bar{\delta})$. 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.

   2.a., 2.b., and 2.c. follow directly from the definitions of $\underline{b}_u(\delta)$, $\bar{b}_s(\delta)$ and from the Observation 2 that $\underline{b}_s(\delta) \leq 1$. Note, it follows from definition of mixed strategy that (3) must bind for $\lambda_{ut} \in (0, 1)$. Similarly, when $\lambda_{st} \in (0, 1)$, (3) must bind.

3. Now consider the case $\delta < \underline{\delta}$. Again it can be seen from (3) that at $\langle \lambda_{ut}, 1 \rangle$ where $\lambda_{ut} > 0$, the benefit from investment of an unskilled worker is strictly lower than her cost of investment. So, she has an incentive to deviate. Hence, there does not exist any equilibrium where $\lambda_{ut} > 0$.

   3.a., 3.b. and 3.c. follow from the definitions of $\bar{b}_s(\delta)$ and $\underline{b}_s(\delta)$. Following the aforementioned argument, it can be shown that when $\lambda_{st} \in (0,1)$ $\lambda_{st} \in (0,1)$, (3) must bind. $\square$

A.3 Proof of Proposition 2

1.a. In this case, from Proposition 1 (subpoint 1.), 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 $A(L_s^*)^{-(1-\phi)} = A\beta^{-(1-\phi)}$.

Now we argue at the steady state, the probability with which an adult works as a skilled worker is β:

$$\begin{aligned}\beta \cdot [\lambda_s^* \cdot \ & \text{the probability that her parent was a skilled worker} \\ & + \lambda_u^* \cdot \ \text{the probability that her parent was an unskilled worker }] = \beta\end{aligned}$$

since $\lambda_s^* = \lambda_u^* = 1$.

1.b. Observe $\langle 0, 1 \rangle$ or $\langle (0,1), 1 \rangle$ cannot be the equilibrium strategy at any steady state because in those cases, the mass of skilled workers decreases over time. Also observe from Proposition 1, these equilibrium strategies exist only when m_{st} is lower than $\underline{b}_u(\delta)$ or the mass of skilled workers is higher than $\beta \left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}}$.

In other words, 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 with a positive probability and hence, the mass of skilled workers decreases over time and their income increases over time and reaches $\underline{b}_u(\delta)$.

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

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

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

Observe, $\underline{b}_u(\delta)$ is time independent, hence L_{st+1} is time independent. So, if an economy is such that $\lambda_{ut} = \hat{\lambda}_{ut}$ and $\lambda_{st} = 1$, then the economy is at a steady state at $t + 1$. 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 $\lambda_u^* \equiv \dfrac{\left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}} - L_s^*}{1 - L_s^*}$.

To check that indeed this is a steady state, we need $\lambda_u^* \in (0,1)$. Now, $\lambda_u^* < 1$ because the degree of child affinity is not high, i.e. $\delta < \bar{\delta}$. And, $\lambda_u^* > 0$ because $m_s^* \equiv \beta^{-(1-\phi)}\underline{b}_u(\delta) > \underline{b}_u(\delta)$.

At the steady state, the probability with which an adult works as a skilled worker is

$$\begin{aligned} &\beta \cdot [\lambda_s^* \cdot \text{ the probability that her parent was a skilled worker} \\ &\quad + \lambda_u^* \cdot \text{ the probability that her parent was an unskilled worker }] \\ &< 1. \end{aligned}$$

where the 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 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 the degree of 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 steady state inequality increases with decrease in δ.

2.a. From Part 3. of Proposition 1., 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 converges to infinity.

2.b. From Part 3. of Proposition 1., 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. □

B Appendix for Bias via Education: Behavioral Trap

B.1 Formal Statement of Definition 3

Given the definition of $\bar{b}_u(\delta)$ and Lemma 1 1., we know at $\bar{b}_u(\delta)$, an educated-unskilled worker is just indifferent between investing and not investing, when all educated workers invest with certainty. Thus, the mass of educated worker at $t+1$ would remain N_{et}. So, $m_{st+1} = m_{st}$ and

$$\bar{b}_u(\delta): \quad \delta\left[\frac{[\beta\bar{b}_u(\delta) + 1 - \beta]^\sigma}{\sigma} - \frac{1}{\sigma}\right] = \frac{1}{\sigma} - \frac{(1-\bar{s})^\sigma}{\sigma}. \tag{A.4}$$

B.2 Proof of Proposition 3

1. See Observation 3.

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

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

Now, the benefit, i.e. the L.H.S. increases with decrease in N_{et}, whose maximum value N_{et} is 1. Similarly, the L.H.S. increases with increase in δ. So, to prove the claim, it sufficient to show that L.H.S. is no less than R.H.S. at $N_{et} = 1$ and $\delta = \bar{\delta}$:

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

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

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

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

B.3 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 probability less than 1. Hence, in the next period, there will be a positive mass of uneducated workers. Following the previous statement, we see that there will be poverty trap in the economy.
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 exists a poverty trap.

2.a. From Proposition 3., we have that 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 $\bar{b}_u(\bar{\delta})$. If the mass of uneducated workers is positive, then that would be strictly higher than $\bar{b}_u(\bar{\delta})$. All the educated workers invest forever from the beginning. Hence, the result.

2.b. From Proposition 3., when the degree of child affinity is moderate and $m_{st} \geq \bar{b}_u(\delta)$, all educated workers invest. 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., we have that when $m_{st} < \bar{b}_u(\delta)$, educated-unskilled workers invest with probability less than 1. Hence, the mass of educated workers, and hence the mass of skilled workers decrease over time. Since, the income of a skilled worker is inversely related to the mass of skilled workers, this implies the income of a skilled worker increases over time. This happens till at some time until $m_{st} \geq \bar{b}_u(\delta)$. Then, we are in the region described in Part 2.b., and hence reach a steady state, $m_s^* \geq \bar{b}_u(\delta)$.

2.d. For low child affinity, the case is same as the benchmark model. The proof is similar to that of Proposition 2 (subpoint 2).

2.e. For low child affinity, the case is same as the benchmark model. The proof is same as that for Proposition 2 (subpoint 2).

3.a. Multiplicity of steady state follows from the discussion of the dynamics in Proposition 4. The steady state inequality, which is the difference between the income of a skilled worker and an unskilled worker, increases with increase in the income of a skilled worker (as the unskilled worker's income is unity).

For high child affinity, the least unequal steady state occurs at $m_s^* = 1$. For moderate child affinity, the least unequal steady state is $\bar{b}_u(\delta) > 1$, which as we have noted in Observation 5, is decreasing in δ. Hence the claim.

3.b. For low child affinity, the unskilled workers do not invest in the benchmark as well as in the behavioral trap case. Hence, this proof is very similar to the proof of Proposition 2 (subpoint 2.), so we skip the details.

□

C Appendix for Bias via Education and Job Network

C.1 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

$$\underline{a}_s(\delta): \quad \underline{a}_s^\sigma - (\underline{a}_s - \bar{s})^\sigma + \delta = 0. \tag{A.5}$$

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

$$a_6(\delta): \quad \frac{\delta}{\sigma}\left[\left[[1-\theta(1-\beta)]^\phi a_6 + \theta(1-\beta)\right]^\sigma - 1\right] = \frac{a_6^\sigma - (a_6 - \bar{s})^\sigma}{\sigma}. \tag{A.6}$$

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

$$a_4(\delta): \quad \frac{\delta}{\sigma}\left[\left[[1-\theta(1-\beta)](1-\beta)^{-(1-\phi)} a_4 + \theta(1-\beta)\right]^\sigma - 1\right] = \frac{a_4^\sigma - (a_4 - \bar{s})^\sigma}{\sigma}. \tag{A.7}$$

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

$$a_2(\delta): \quad \frac{\delta}{\sigma}\Big[\big[[1-\theta(1-\beta)][1+(1-\theta)(1-\beta)]^{-(1-\phi)}a_2+\theta(1-\beta)\big]^{\sigma}-1\Big] = \frac{a_2^{\sigma}-(a_2-\bar{s})^{\sigma}}{\sigma}. \tag{A.8}$$

- At $a_5(\delta)$ an educated-unskilled worker is just indifferent between investing and not investing, when all skilled workers are investing and no other educated-unskilled worker is investing. So, from (5)

$$a_5(\delta): \frac{\delta}{\sigma}\left[\left[\theta\beta^{\phi}a_5+1-\theta\beta\right]^{\sigma}-1\right] = \frac{1-(1-\bar{s})^{\sigma}}{\sigma}. \tag{A.9}$$

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

$$a_3(\delta): \quad \frac{\delta}{\sigma}\left[\left[\theta\beta[\theta(1-\beta)]^{-(1-\phi)}a_3+1-\theta\beta\right]^{\sigma}-1\right] = \frac{1-(1-\bar{s})^{\sigma}}{\sigma} = 0. \tag{A.10}$$

- At $a_1(\delta)$ an educated-unskilled worker is just indifferent between investing and not investing, when all other educated workers are investing with certainty. So, from (6)

$$a_1(\delta): \quad \frac{\delta}{\sigma}\Big[\left[\theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}a_1+1-\theta\beta\right]^{\sigma}-1\Big] = \frac{1-(1-\bar{s})^{\sigma}}{\sigma}. \tag{A.11}$$

C.2 Derivation of Boundary Conditions for Equilibrium Probabilities

1. <u>**Condition $\underline{\Gamma}_s$:**</u> We will show $\underline{\gamma}_s(m_{st})$ is such that

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

   Suppose, $m_{st} \geq a_4(\delta)$. From the definitions of $a_4(\delta)$ and $a_2(\delta)$, it is clear from the expression of Condition $\underline{\Gamma}_s$, stated in the main paper, that at $m_{st} = a_4(\delta)$, $\underline{\gamma}_s(m_{st}) = 0$, and at $m_{st} = a_2(\delta)$, $\underline{\gamma}_s(m_{st}) = 1$.

   Now we prove several claims sequentially.

   To show: For $m_{st} \in (a_4(\delta), a_2(\delta))$. Here $0 < \underline{\gamma}_s(m_{st}) < 1$. Also, at any equilibrium $\langle \gamma_{ut}, \gamma_{st} \rangle$, we find $\gamma_{st}(m_{st}) > \underline{\gamma}_s(m_{st})$.

Suppose we have $0 = \underline{\gamma}_s(m_{st})$. Then, we find that it violates

$$
\begin{aligned}
&\frac{a_4(\delta)^\sigma}{\sigma} - \frac{(a_4(\delta)-\bar{s})^\sigma}{\sigma} \\
>&\delta\left[\frac{\Big[\big[1-\theta(1-\beta)\big]\cdot\big[(1-\beta)+[1-\theta(1-\beta)]\underline{\gamma}_s\big]^{-(1-\phi)}m_{st}+\theta(1-\beta)\Big]^\sigma}{\sigma} - \frac{1}{\sigma}\right] \quad \text{from def. } a_4(\delta) \\
\geq&\frac{m_{st}^\sigma}{\sigma} - \frac{(m_{st}-\bar{s})^\sigma}{\sigma}
\end{aligned}
$$

But if $m_{st} > a_4(\delta)$ implies $\dfrac{m_{st}^\sigma}{\sigma} - \dfrac{(m_{st}-\bar{s})^\sigma}{\sigma} > \dfrac{a_4(\delta)^\sigma}{\sigma} - \dfrac{(a_4(\delta)-\bar{s})^\sigma}{\sigma}$, hence a contradiction.
Similarly, we can show $\underline{\gamma}_s(m_{st}) < 1 = \underline{\gamma}_s(a_2(\delta))$.
In this income range, consider any equilibrium $\langle\gamma_{ut}, \gamma_{st}\rangle$. If $\gamma_{ut} = 1$ then we have shown that $\gamma_{st}(m_{st}) = \bar{\gamma}_s(m_{st})$. If $\gamma_{ut} < 1$ and $\gamma_{st}(m_{st}) \leq \bar{\gamma}_s(m_{st})$, then from (6) we find, for the skilled worker utility benefit from investment is strictly higher than the utility cost. Hence, γ_{st} must be one, which violates that $\bar{\gamma}_s(m_{st})$ is a strict fraction for $m_{st} \in (a_4(\delta), a_2(\delta))$. Thus, we get that for $m_{st} \in (a_4(\delta), a_2(\delta))$, the equilibrium $\langle\gamma_{ut}, \gamma_{st}\rangle$ will be such that $\gamma_{ut} \in [0, 1]$ and $\gamma_{st}(m_{st}) \leq \bar{\gamma}_s(m_{st})$.

To show: For $m_{st} \in [a_2(\delta), \infty)$, at any equilibrium $\langle\gamma_{ut}, \gamma_{st}\rangle$, we find $\gamma_{st}(m_{st}) = 1$.
At $m_{st} = a_2(\delta)$, if $\gamma_{ut} = 1$ then $\gamma_{st} = 1$. Using the same argument as in the previous proof, we find that at $m_{st} = a_2(\delta)$, $\gamma_{st} = 1\ \forall\ \gamma_{ut} \leq 1$. Also, for any $m_{st} > a_2(\delta)$, the benefit of investment is higher than the cost for the skilled workers. Hence, for this range of the state variable: $\gamma_{st} = 1\ \forall\ \gamma_{ut} \leq 1$.

To show: For $m_{st} = a_4(\delta)$, at any equilibrium $\langle\gamma_{ut}, \gamma_{st}\rangle$, we find $\gamma_{st}(m_{st}) = 0$.
Using same logic as in the previous two cases.

To show: 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

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

But it is not possible as

$$\frac{m_{st}^2}{\sigma} - \frac{(m_{st}^2 - \bar{s})^\sigma}{\sigma} < \frac{m_{st}^1}{\sigma} - \frac{(m_{st}^1 - \bar{s})^\sigma}{\sigma}.$$

Therefore, when $m_{st} \in (a_4(\delta), a_2(\delta))$ we have $\underline{\gamma}_s(m_{st})$ is increasing in m_{st}.

2. **Condition $\bar{\boldsymbol{\Gamma}}_{\mathbf{s}}$:** Based on the expression of Condition $\bar{\Gamma}_s$, stated in the main paper, we prove that

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

From the definitions of $\underline{a}_s(\delta)$ and $a_6(\delta)$, now it is clear that at $m_{st} = \underline{a}_s(\delta)$ we get $\bar{\gamma}_s(m_{st}) = 0$, and at $m_{st} = a_6(\delta)$, the probability of investment by skilled worker is $\bar{\gamma}_s(m_{st}) = 1$.

To show: For $m_{st} \in (\underline{a}_s(\delta), a_6(\delta))$. Here $0 < \bar{\gamma}_s(m_{st}) < 1$. Also, at any equilibrium $\langle \gamma_{ut}, \gamma_{st} \rangle$, we find $\gamma_{st}(m_{st}) < \bar{\gamma}_s(m_{st})$.

Suppose we have $0 = \bar{\gamma}_s(m_{st})$. Then we find that this violates the claim that the probability of investment must be zero at $\underline{a}_s(\delta)$. Similarly if $\bar{\gamma}_s(m_{st})$ were one, it would violate the property that probability of investment is one at $a_6(\delta)$.

Also, the proof of $\bar{\gamma}_s(m_{st})$ being increasing in m_{st} for $m_{st} \in (\underline{a}_s, a_6)$ is similar to the one stated in the Condition $\underline{\Gamma}_s$.

For $m_{st} \in [1, a_6(\delta))$, suppose an equilibrium $\langle \gamma_{ut}, \gamma_{st} \rangle$ exists. We can apply the same reasoning as in the previous case to show that $\gamma_{st}(m_{st})$ is bounded above by $\bar{\gamma}_s(m_{st})$ and $\gamma_{ut} \in [0, 1]$.

3. Can be proved following the argument used in the proof of Condition $\underline{\Gamma}_s$.

4. Can be proved following the argument used in the proof of Condition $\bar{\Gamma}_s$. $\square$

C.3 Characterization of Equilibria

We introduce an observation and a lemma which we use to prove Proposition 5 in Section C.3.1.

Observation C.1. *Suppose there are two equilibria $\langle \gamma_{ut}, \gamma_{st} \rangle$ and $\langle \tilde{\gamma}_{ut}, \tilde{\gamma}_{st} \rangle$ at any $m_{st} \geq 1$. 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). $\square$

Lemma C.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 $\gamma_{st} > \gamma_{ut}$, then at all $\tilde{m}_{st} \in (m_{st}, min\{a_1(\delta), a_2(\delta)\})$ $\tilde{\gamma}_{st} > \tilde{\gamma}_{ut}$ 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 $\langle \gamma_{ut}, \gamma_{st} \rangle$ $\langle \tilde{\gamma}_{ut}, \tilde{\gamma}_{st} \rangle$ 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
   $$\theta(1-\beta)\gamma_{ut} + \beta\gamma_{st} = \theta(1-\beta)\tilde{\gamma}_{ut} + \beta\tilde{\gamma}_{st} \tag{A.12}$$
   And, from the decision of skilled workers, given by (6), we must have
   $$(1-\beta)\gamma_{ut} + [1-\theta(1-\beta)]\gamma_{st} = (1-\beta)\tilde{\gamma}_{ut} + [1-\theta(1-\beta)]\tilde{\gamma}_{st} \tag{A.13}$$
   As $\theta(1-\beta) \neq \beta$, from (A.12) and (A.13), we have $\gamma_{ut} = \tilde{\gamma}_{ut}$ and $\gamma_{st} = \tilde{\gamma}_{st}$ – a contradiction.

2. We find necessary condition for the existence of multiple equilibria at any $m_{st} \in [1, \min\{a_1(\delta), a_2(\delta)\})$. Given Definition of $a_1(\delta)$, $a_2(\delta)$, Lemma 3, and part (i) of this lemma, 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. Using (5) and (6), we get

$$\begin{aligned}
& \theta(1-\beta)\tilde{\gamma}_{ut} + \beta\tilde{\gamma}_{st} \geq \theta(1-\beta)\gamma_{ut} + \beta\gamma_{st} \\
\Rightarrow \quad & \beta[\tilde{\gamma}_{st} - \gamma_{st}] \geq \theta(1-\beta)[\gamma_{ut} - \tilde{\gamma}_{ut}]
\end{aligned} \tag{A.14}$$

$$\begin{aligned}
& (1-\beta)\gamma_{ut} + [1-\theta(1-\beta)]\gamma_{st} \geq (1-\beta)\tilde{\gamma}_{ut} + [1-\theta(1-\beta)]\tilde{\gamma}_{st} \\
\Rightarrow \quad & (1-\beta)[\gamma_{ut} - \tilde{\gamma}_{ut}] \geq [1-\theta(1-\beta)][\tilde{\gamma}_{st} - \gamma_{st}].
\end{aligned} \tag{A.15}$$

Both conditions (A.14) and (A.15) 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:

$$\beta \geq \theta[1-\theta(1-\beta)] \qquad \Rightarrow \quad \beta \geq \theta(1-\beta).$$

3. Suppose not. $\gamma_{st} > \gamma_{ut}$ 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, at m_{st} and $\tilde{m}_{st}$ less than $\min\{a_1(\delta), a_2(\delta)\}$ implies $\gamma_{ut} < 1$ and $\tilde{\gamma}_{st} < 1$.

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

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

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

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

   From these two conditions we get

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

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

□

C.3.1 Proof of Proposition 5

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

2. $\delta > \delta_a$, so from Lemma 4. iii. (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 no less than $\underline{\gamma}_s(m_{st})$ $\forall$ $m_{st} \geq 1$.

The equilibrium is unique as at any $m_{st} \geq 1$, $\gamma_{ut} = 1$. This implies equilibrium γ_{st} would be equal to $\underline{\gamma}_s(m_{st})$, which is unique.

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

Now, by Lemma 4. (subpoint 4.), we have $a_4(\delta) \leq a_1(\delta)$. Hence, due to Condition $\underline{\Gamma}_s$, it follows from the previous argument that the γ_{st} must be no less than $\underline{\gamma}_s(m_{st})$ and the equilibrium is unique as at any $m_{st} \geq a_1(\delta)$.

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

By the definition of $a_5(\delta)$, eq. (5) and Condition $\underline{\Gamma}_u$, we have that

$$\gamma_{ut} = 0 \quad \forall\ m_{st} \in [a_2(\delta), a_5(\delta)) \qquad \gamma_{ut} = \underline{\gamma}_u(m_{st}) \quad \forall\ m_{st} \geq a_5(\delta).$$

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. For $\beta < \theta(1-\beta)$, the uniqueness of equilibrium follows from Part 3. of Lemma C.1. For $\beta > \theta(1-\beta)$ there is multiple equilibria. There is at most on equilibrium at a m_{st}, such that both educated parents randomize. This follows from Part 1. and 2. of Lemma C.1.

4. From Lemma 4, we have $a_1(\delta), a_3(\delta), a_5(\delta)$ tend to infinity for $\delta \leq \bar{\delta}$. Hence, the educated-unskilled workers do not invest for any finite state variable: $\gamma_{ut} = 0$ $\forall$ $\gamma_{st} \in [0, 1]$ and $m_{st} > 1$.

   From Lemma 2. (subpoint 3.) and Lemma 4 (subpoints 2. and 5.), we have

$$1 < \underline{b}_s(\delta) = \underline{a}_s(\delta) < a_6(\delta).$$

   So, due to Condition $\underline{\Gamma}_s$, for any $m_{st} \in [1, a_6(\delta)]$, $\gamma_{st} \leq \bar{\gamma}_s(m_{st})$.

   The equilibrium is unique as $\gamma_{ut} = 0$ for any finite m_{st}. Hence, by Condition $\underline{\Gamma}_s$, we must have $\gamma_{st} = \bar{\gamma}_s(m_{st})$ $\forall$ $m_{st} \in [1, a_6(\delta)]$ and by definition of $a_6(\delta)$, $\gamma_s = 1$ $\forall$ $m_{st} > a_6(\delta)$. □

C.4 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.

2.d. We 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. Now, the inequality at the 'least unequal steady state' remains constant for all δ, which are no less than the particular degree of child affinity ($\hat{\delta}$) at which $\max\{a_1(\delta), a_2(\delta)\} = A\beta^{-(1-\phi)}$.

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. □

D Appendix for Comparison

D.1 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}, $\rho_{ut} > 0$, $\rho_{st} = \lambda_{st} = 1$.

1.b. From Proposition 3, we have that when $\delta \geq \bar{\delta}$ then $\rho_{ut} = 1$. Now, we argue when $\delta \in (\underline{\delta}, \bar{\delta})$, the mass of uneducated workers is positive and $m_{st} > \underline{b}_u(\delta)$ then $\rho_{ut} > \lambda_{ut} > 0$.

Lemma 2 implies $\underline{b}_u(\delta) > \bar{b}_s(\delta)$. So, for $m_{st} > \underline{b}_u(\delta)$, we get $\lambda_{st} = \rho_{st} = 1$ and $\lambda_{ut} \in (0, 1)$. Using these probabilities of investment, from (3) and (4) we can derive

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

where we have used the fact that the mass of uneducated workers is positive, i.e. $1-N_{et} > 0$ and $\lambda_{ut} > 0$.

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

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

D.2 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 inequality is higher. Hence, the result.

References

Appadurai, A. (2004). The Capacity to Aspire: Culture and the Terms of Recognition in Vijayendra Rao and Michael Walton (eds), Culture and Public Action. 5

Balboni, C., Bandiera, O., Burgess, R., Ghatak, M., and Heil, A. (2020). Why do People Stay Poor? 1, 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. 4

Banerjee, A. V. and Newman, A. F. (1993). Occupational Choice and the Process of Development. *Journal of Political Economy*, 101(2):274–298. 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. 1

Benoît, J.-P. and Dubra, J. (2011). Apparent Overconfidence. *Econometrica*, 79(5):1591–1625. 1

Bernard, T., Dercon, S., and Taffesse, A. S. (2011). Beyond Fatalism – An Empirical Exploration of Self-efficacy and Aspirations Failure in Ethiopia. 1

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

Besley, T. (2017). Aspirations and the Political Economy of Inequality. *Oxford Economic Papers*, 69(1):1–35. 5

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

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

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. 2

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

Cervone, D. and Peake, P. K. (1986). Anchoring, Efficacy, and Action: The Influence of Judgmental Heuristics on Self-efficacy Judgments and Behavior. *Journal of Personality and Social Psychology*, 50(3):492. 5

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. 1, 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

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

Genicot, G. and Ray, D. (2017). Aspirations and Inequality. *Econometrica*, 85(2):489–519. 1, 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. 5

Malmendier, U. and Tate, G. (2005). CEO Overconfidence and Corporate Investment. *The Journal of Finance*, 60(6):2661–2700. 1, 4

Mani, A., Mullainathan, S., Shafir, E., and Zhao, J. (2013). Poverty Impedes Cognitive Function. *Science*, 341(6149):976–980. 1

Mookherjee, D. and Ray, D. (2002). Contractual Structure and Wealth Accumulation. *American Economic Review*, 92(4):818–849. 1

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

Ortoleva, P. and Snowberg, E. (2015). Overconfidence in Political Behavior. *American Economic Review*, 105(2):504–35. 5

Ray, D. (2006). Aspirations, Poverty, and Economic Change. In *Understanding Poverty*. Oxford University Press. 5

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. 1, 12

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. 1

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. 5

Supplementary Appendix

SA Appendix for Benchmark Case

SA.1 Proof of Observation 1.

In an economy with both skilled and unskilled workers, $0 < L_{st} < 1$. Since $0 < \phi < 1$, $L_{st}^{-(1-\phi)} > 1$. Thus, $m_{st} = AL_{st}^{-(1-\phi)} \geq 1 = m_{ut}$. $\square$

SA.2 Proof of Lemma 1.

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

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

   Since $\sigma < 0$ it implies $y'(x) < 0$ and $y''(x) > 0$. Since $m_{st} > m_{ut}$, it implies $y(m_{st}) < y(m_{ut})$. Thus if

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

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

   Hence, if an unskilled worker invests in her child's education with positive probability, then a skilled worker invest in her child's education with certainty. $\square$

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

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

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

   Now, as m_{st} increases, the utility cost of investment, i.e., the RHS of the above inequality decreases for skilled workers and remains the same for unskilled workers. Increase in m_{st} implies decrease in L_{st}. It can be easily derived that given $\langle \lambda_{ut}, \lambda_{st} \rangle$, the L.H.S. of the above inequality (weakly) increases with decrease in L_{st} as $\lambda_{st} \geq \lambda_{ut}$.

   Thus, with increase in the income of the skilled worker at period t, the benefit of investment (weakly) increases and the utility cost of investment (weakly) decreases. Hence, the probability of investment must (weakly) increase. $\square$

SA.3 Proof of Observation 2.

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

SA.4 Proof of Lemma 2.

1. This is immediate from comparing definitions, (A.1), (A.2), and (A.3). Also from differentiating these equations with respect to δ, we get that these income-cutoffs are decreasing in δ.

2. From (A.3) we get that $b_u(\delta) < 1$ if and only if

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

   We now show that $\bar{b}_s(\delta) > 1$ if and only if $\delta < \bar{\delta}$.

   $\bar{b}_s(\delta) > 1$ if and only if the expected net benefit (and utility cost) from the investment of a skilled parent is higher (and lower respectively) at $\bar{b}_s(\delta)$ than at 1, when all other skilled workers are investing and no unskilled worker is investing. That is, $\bar{b}_s(\delta) > 1$ if and only if

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

   We show if and only if $\delta \in [0, \underline{\delta})$, then $\underline{b}_s(\delta) > 1$. It follows directly from (A.1).

   That $\underline{b}_u(\delta) = \infty$ when $\delta \in [0, \underline{\delta})$, also follows directly from (A.3). Hence, proved. □

SA.5 Equilibria Strategies

Given Lemma 1 1., there can be five equilibria:

χ1. Both unskilled and skilled workers invest with certainty: $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 1, 1 \rangle$.

χ2. Unskilled workers invest with positive probability and skilled workers invest with certainty: $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle (0,1), 1 \rangle$.

χ3. Unskilled workers do not invest and skilled workers invest with certainty: $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 0, 1 \rangle$.

χ4. Unskilled workers do not invest and skilled workers invest with positive probability: $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 0, (0,1) \rangle$.

χ5. Both unskilled and skilled workers do not invest: $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 0, 0 \rangle$.

We consider all the strategies below. Recall the notation: L_{ut} : Mass of unskilled workers at period t, L_{st} : Mass of skilled workers at period t, N_{et} : Mass of educated adult at period t.

χ1. Both unskilled and skilled workers invest with certainty: $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 1, 1 \rangle$

All skilled and unskilled parents invest in their children's education. So,

$$N_{et+1} = 1$$
$$L_{st+1} = \beta(L_{st} + L_{ut}) = \beta, \qquad m_{st+1} = AL_{st+1}^{-(1-\phi)} = A\beta^{-(1-\phi)}$$
$$L_{ut+1} = 1 - L_{st+1} = 1 - \beta, \qquad m_{ut+1} = 1.$$

As observed in Lemma 1.1., whenever an unskilled worker has an incentive to invest, a skilled worker also has an inventive to invest with certainty. So, this case $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 1, 1 \rangle$ arises when the parametric condition is such that an unskilled worker has an incentive to invest with certainty, when all other workers invest with certainty. We, now, characterize such parametric condition.

At any t, both unskilled and skilled workers invest with certainty is an equilibrium if and only if

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

So, when the parents are very child loving, all workers invest with certainty.

Observe, this condition is time independent, hence if this condition satisfies at certain t, then it holds at all other t. This implies that the economy enters in to a steady state at $t = 2$, we summarize this in the following observation.

Observation SA.1. *If and only if the degree of child affinity is high , i.e. $\delta \geq \bar{\delta}$*

(i) All parents always invest with certainty.

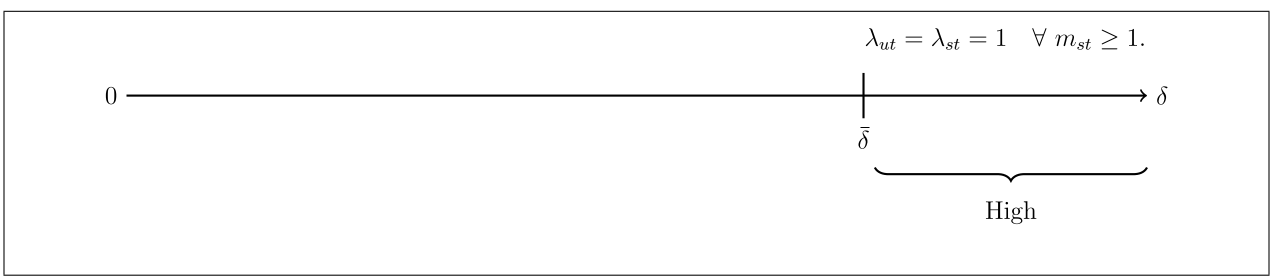

Figure 6: If and only if $\delta \geq \bar{\delta}$, $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 1, 1 \rangle$

(ii) At $t = 2$, the economy immediately enters into a steady state.

At the steady state, all types of parents invest with certainty. Hence, the number of skilled and unskilled workers and their respective incomes remain constant over time.

At steady state: the mass of educated individual $N_e^* = 1$,
the mass of skilled workers $L_s^* = \beta$, and wage of a skilled worker $m_s^* = A\beta^{-(1-\phi)}$,
the mass of unskilled workers $L_u^* = 1 - \beta$, and wage of an unskilled worker $m_s^* = 1$.

χ**2. Unskilled workers invest with positive probability:** $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle (0,1), 1 \rangle$

Unskilled workers invest with positive probability $\lambda_{ut} \in (0,1)$ and skilled workers invest with certainty. So,

$$\begin{aligned} N_{et+1} &= \lambda_{ut} L_{ut} + L_{st} \\ L_{st+1} &= \beta(\lambda_{ut} L_{ut} + L_{st}), \qquad m_{st+1} = A\left(\beta(\lambda_{ut} L_{ut} + L_{st})\right)^{-(1-\phi)} \\ L_{ut+1} &= (1-\beta\lambda_{ut})L_{ut} + (1-\beta)L_{st}, \qquad m_{ut+1} = 1. \end{aligned}$$

We are looking for parametric condition such that an unskilled worker is indifferent between investing and not investing when all other unskilled workers are investing with probability $\lambda_{ut} \in (0,1)$ and all skilled workers are investing with certainty. Whenever the degree of child affinity is low i.e. $\delta < \bar{\delta}$ unskilled parents invest with probability strictly less than 1. They invest with positive probability only when m_{st} is higher than $\underline{b}_u(\delta)$. From Lemma 2 we know $\underline{b}_u(\delta)$ is infinite when the degree of child affinity is low, i.e. $\delta < \underline{\delta}$. This gives us the part (a) of the following observation.

Now, coming to the dynamics, we show that λ_{ut} is such that the economy enters into a steady state at $t+1$. For that consider, again, the incentive constraint of an unskilled worker when when all other unskilled workers are investing with probability λ_{ut} and all skilled workers are investing with certainty

$$\frac{(m_{ut} - \bar{s})^{\sigma}}{\sigma} + \delta \frac{[\beta m_{st+1} + (1-\beta)m_{ut+1}]^{\sigma}}{\sigma} = \frac{m_{ut}^{\sigma}}{\sigma} + \delta \frac{m_{ut+1}^{\sigma}}{\sigma}$$

At the steady state, $L_{st+1} = L_{st}$ which implies

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

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

Observe, $\underline{b}_u(\delta)$ is time independent, hence L_{st+1} is time independent. So, if an economy is such that $\lambda_{ut} = \dfrac{\left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}} - L_{st}}{1 - L_{st}}$ and $\lambda_{st} = 1$, then the economy is at a steady state at $t+1$. 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 $\lambda_u^* \equiv \dfrac{\left(\underline{b}_u(\delta)/A\right)^{-\frac{1}{1-\phi}} - L_s^*}{1 - L_s^*}$.

To check that indeed this is a steady state, we need $\lambda_u^* \in (0,1)$. Now, $\lambda_u^* < 1$ because the degree of child affinity is not high, i.e. $\delta < \bar{\delta}$. And, $\lambda_u^* > 0$ because $m_s^* \equiv \beta^{-(1-\phi)}\underline{b}_u(\delta) > \underline{b}_u(\delta)$. Hence, the part (b) of the following observation.

Observation SA.2. *If and only if the degree of child affinity is moderate i.e.* $\delta \in [\underline{\delta}, \bar{\delta})$ *and* $m_{st} \in (\underline{b}_u(\delta), \infty)$*, then*

(a) unskilled workers invest with probability λ_{ut}*, where* $\lambda_{ut} \in (0,1)$*, and all skilled workers*

invest with certainty is the unique equilibrium.

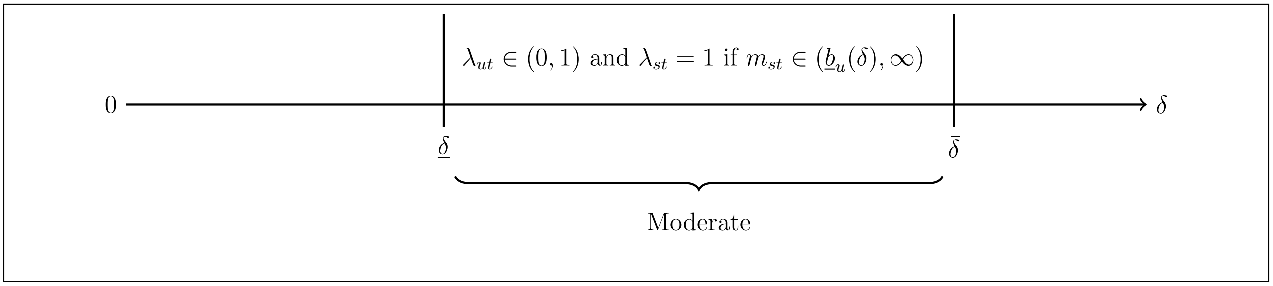

Figure 7: At period t, $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle (0,1), 1 \rangle$ is an equilibrium

(b) *The economy enters into a steady state at $t+1$. At steady state,*

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

χ3. Unskilled workers do not invest and skilled workers invest with certainty: $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 0, 1 \rangle$

All skilled parents invest with probability 1 and no unskilled parents invest. So,

$$
\begin{aligned}
N_{et+1} &= L_{st} \\
L_{st+1} &= \beta L_{st}, \qquad m_{st+1} = A L_{st+1}^{-(1-\phi)} = \beta^{-(1-\phi)} m_{st} \\
L_{ut+1} &= 1 - L_{st+1}, \qquad m_{ut+1} = 1.
\end{aligned}
$$

It follows from the definition of $\bar{b}_s$ that the income at which skilled workers invest with unit probability is $m_{st} > \bar{b}_s$. Similarly, income range for which unskilled parents do not invest with certainty is $m_{st} < \underline{b}_u$. Now, $\underline{b}_u(\delta)$ is infinite when $\delta < \underline{\delta}$. This gives us the part (a) of the observation. Now, at this parametric condition, all skilled workers invest with probability 1 and no unskilled workers invest, so in the next period, the mass of educated individual would be equal to the mass of skilled workers at this period. Among those educated individuals β part will become skilled workers. Hence, the part (b) of the following observation.

Observation SA.3. *(i) At any period t, all skilled workers invest with certainty and no unskilled workers invest if*

(a) *either the degree of child affinity is moderate, i.e.* $\delta \in [\underline{\delta}, \bar{\delta})$ *and* $m_{st} \in [\bar{b}_s, \underline{b}_u)$
(b) *or the degree of child affinity is low, i.e.* $\delta \in (0, \underline{\delta})$ *and* $m_{st} \in [\bar{b}_s, \infty)$.

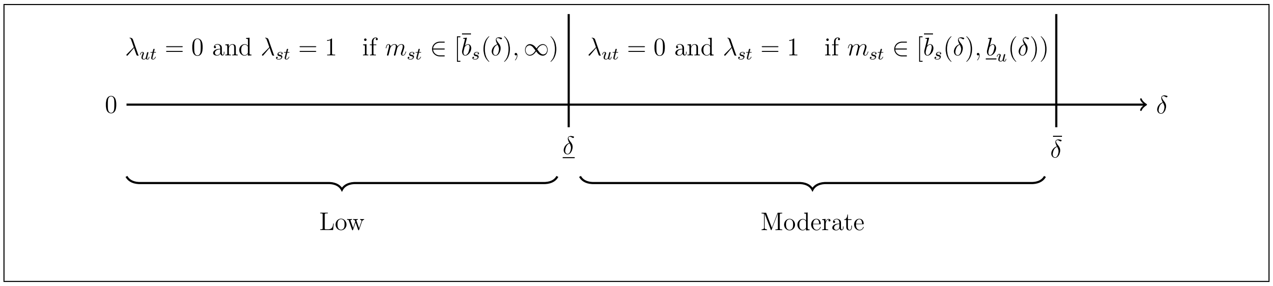

Figure 8: At period t, $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 0, 1 \rangle$ is an equilibrium

(ii) If either of the above two conditions is satisfied, then the number of skilled workers fall at the rate β while skilled income rises at the rate $\beta^{-(1-\phi)}$.

χ4. Unskilled workers do not invest and skilled invest with positive probability: $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 0, (0, 1) \rangle$

Unskilled workers do not invest and skilled workers invest with probability $\lambda_{st} \in (0, 1)$. So,

$$\begin{aligned} N_{et+1} &= \lambda_{st} L_{st} \\ L_{st+1} &= \beta \lambda_{st} L_{st}, \qquad m_{st+1} = A \left(\beta \lambda_{st} L_{st}\right)^{-(1-\phi)} = \left(\beta \lambda_{st}\right)^{-(1-\phi)} m_{st} \\ L_{ut+1} &= 1 - \beta \lambda_{st} L_{st}, \qquad m_{ut+1} = 1. \end{aligned}$$

We are looking for parametric condition such that a skilled worker is indifferent between investing and not investing when all other skilled workers are investing with probability λ_{st} and no unskilled workers are investing. Hence, m_{st} must be higher than $\underline{b}_s(\delta)$ but lower than $\bar{b}_s(\delta)$. Now, in Lemma 2, we have seen that $\underline{b}_s(\delta) > 1$ if and only if $\delta < \underline{\delta}$ and in Lemma 1, we have seen that $m_{st} \geq 1$. This gives us the part (a) of the observation. Now, at this parametric condition, skilled workers invest with probability λ_{st} and no unskilled workers invest, so in the next period, the mass of educated individual would be less than the mass of skilled workers at this period. Among those educated individuals β part will become skilled workers. Hence, the part (b) of the following observation.

Observation SA.4. *(i) At any period t, no unskilled workers invest and skilled workers invest with probability λ_{st} such that $\lambda_{st} \in (0, 1)$ and if*

(a) either the degree of child affinity is moderate, i.e. $\delta \in [\underline{\delta}, \bar{\delta})$ and $m_{st} \in [1, \bar{b}_s(\delta))$

(b) or the degree of child affinity is low, i.e. $\delta \in (0, \underline{\delta})$ and $m_{st} \in (\underline{b}_s(\delta), \bar{b}_s(\delta))$.

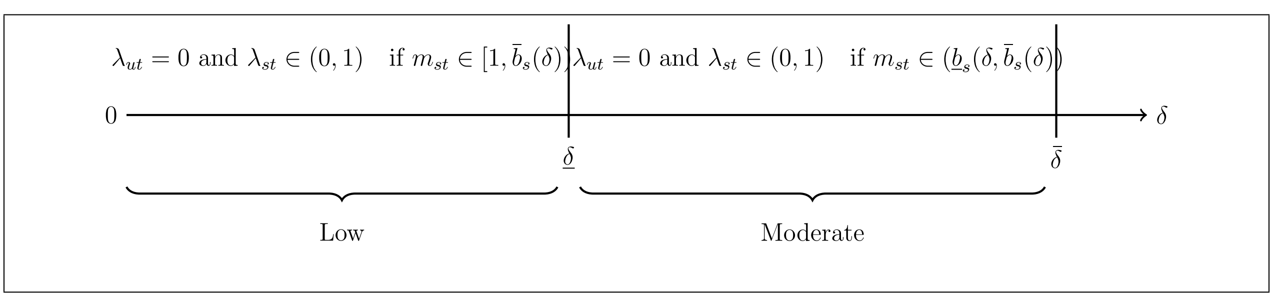

Figure 9: At period t, $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 0, (0, 1) \rangle$ is an equilibrium

(ii) If either of the above two conditions is satisfied, then the number of skilled workers fall and the income of skilled workers rises over time.

χ5. Both unskilled workers and skilled workers do not invest: $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 0, 0 \rangle$

No worker invests, so

$$\begin{aligned} N_{et+1} &= 0 \\ L_{st+1} &= 0, \qquad m_{st+1} \to \infty \\ L_{ut+1} &= 1, \qquad m_{ut+1} = 1. \end{aligned}$$

It follows from the definition of $\underline{b}_s(\delta)$ that m_{st} must be lower than $\underline{b}_s(\delta)$. Now, in Lemma 2, we have seen that $\underline{b}_s(\delta) > 1$ if and only if $\delta < \underline{\delta}$. Together with Lemma 1, we find that neither skilled nor unskilled parents invest in their children education for $\delta < \underline{\delta}$ and $1 \leq m_{st} < \underline{b}_s$ – explains part (a) of the following observation. Now, we have observed that a skilled worker does not have any incentive to invest when no other worker is investing. In the next period, thus, all workers would be unskilled and income of an unskilled worker is lower than that of a skilled worker, so from the next period onwards also, no parent would ever invest in her child's education. Hence, the part (b) of the following observation.

Observation SA.5. *If and only if degree of child affinity is low $\delta < \underline{\delta}$ and $m_{st} \in [1, \underline{b}_s(\delta))$*

(i) No workers invest.

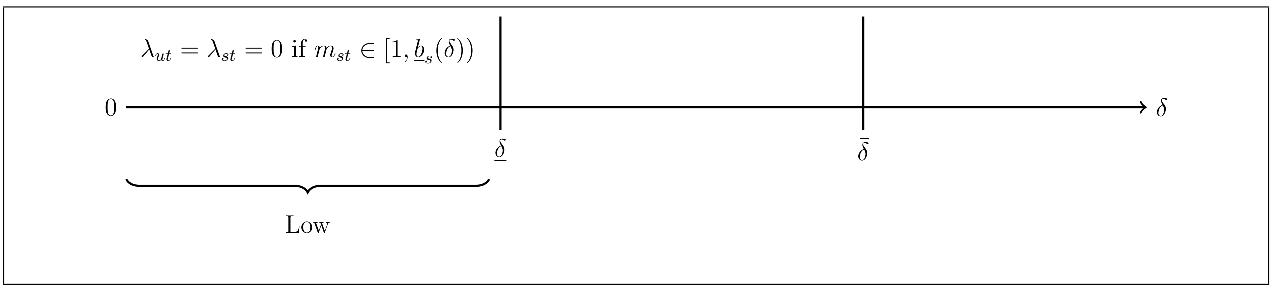

Figure 10: At period t, $\langle \lambda_{ut}, \lambda_{st} \rangle = \langle 0, 0 \rangle$ is an equilibrium

(ii) The economy immediately jumps to steady state with no skilled or educated worker. Further as $m_{st+1} \to \infty$.

SB Appendix for Bias via Education: Behavioral Trap

SB.1 Proof of Observation 5

By differentiating equation (A.4) with respect to δ, we get that the income-cutoff is decreasing in δ.
Now about the value of this cut-off for different child affinity ranges. Substituting $\bar{\delta}$ in (A.4), we have $\bar{b}_u(\delta) = A\beta^{-(1-\phi)}$. Suppose $\delta \in (\underline{\delta}, \bar{\delta})$. From the definitions of $\underline{b}_u(\delta)$ and $\bar{b}_u(\delta)$ we get $\bar{b}_u(\delta) = \beta^{-(1-\phi)}\underline{b}_u(\delta)$, which obviously is greater than $\underline{b}_u(\delta)$. That $\bar{b}_u(\delta) = \infty$ when the degree of child affinity is low follows directly from (A.4). $\square$

SC Appendix for Bias via Education and Job Network

SC.1 Proof of Lemma 3

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

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

$$\begin{aligned}
\delta\Big[1 - \big[\theta\beta[\theta\gamma_{ut}(1-\beta)]^{-(1-\phi)}m_{st} + 1 - \theta\beta\big]^{\sigma}\Big] &\geq (1-\bar{s})^{\sigma} - 1 > (m_{st}-\bar{s})^{\sigma} - m_{st}^{\sigma} \quad \text{(since } m_{st} > 1)\\
&> \delta\Big[1 - \big[[1-\theta(1-\beta)][\gamma_{ut}(1-\beta)]^{-(1-\phi)}m_{st} + \theta(1-\beta)\big]^{\sigma}\Big] \quad \text{(since } \gamma_{st} = 0)\\
\Rightarrow \quad (1-\theta) &\geq [\gamma_{ut}(1-\beta)]^{-(1-\phi)}\big[[1-\theta(1-\beta)] - \theta^{\phi}\beta\big]m_{st}. \qquad \text{(S.2)}
\end{aligned}$$

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

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

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

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

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

Hence, $[\gamma_{ut}(1-\beta)]^{-(1-\phi)}\big[[1-\theta(1-\beta)-\theta^{\phi}\beta\big]m_{st} > (1-\beta)^{-(1-\phi)}\big[[1-\theta(1-\beta)-\theta^{\phi}\beta\big] \geq 1-\theta$ which contradicts (S.2). □

SC.2 Proof of Observation 6

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

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

SC.3 Proof of Lemma 4

1. We want to show that all income cut-offs are decreasing in δ. Consider the income-cutoff $a_6(\delta)$, which is determined by (A.6). The L.H.S. of (A.6) is increasing in a_6 and δ and the R.H.S is decreasing in a_6. Differentiating the equation gives $da_6/d\delta < 0$. We get the

same result for the other income cut-offs $\underline{a}_s$, a_1, a_2, a_3, a_4 and a_5 by differentiating their respective equations (A.5), (A.7)–(A.11).

2. We show $a_2(\delta) > 1$. Suppose not. Since $a_2(\delta) \leq 1$, $(a_2(\delta)-\bar{s})^\sigma - a_2(\delta)^\sigma \geq (1-\bar{s})^\sigma - 1 > 0$. We get the utility cost of investment is positive. Since,

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

which implies that the benefit from investment is positive and strictly greater than the cost of investment, which violates the property that at $a_2(\delta)$ both should be equal.

We first show $a_2(\delta) > a_4(\delta)$. Suppose not and $a_2(\delta) \leq a_4(\delta)$. Hence, from Lemma 1 it follows that the utility cost of investment at $a_2(\delta)$ is no less than that at $a_4(\delta)$: $(a_4(\delta) - \bar{s})^\sigma - a_4(\delta)^\sigma \leq (a_2(\delta) - \bar{s})^\sigma - a_2(\delta)^\sigma$. Since, $1 + (1-\theta)(1-\beta) > 1-\beta$, we get

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

which contradicts the claim that utility cost of investment at $a_2(\delta)$ is weakly higher than that at $a_4(\delta)$.

Now, we show $a_2(\delta) > a_6(\delta)$. We follow similar steps as in the previous case and prove this by contradiction.

Finally, we show $\underline{a}_s(\delta) < a_4(\delta)$ and $\underline{a}_s(\delta) < a_6(\delta)$.

Comparing the cutoff levels, we find that the benefit of investment at $\underline{a}_s(\delta)$ is infinity whereas that at the primitive of $a_4(\delta)$ or of $a_6(\delta)$ is finite. Hence, to make a skilled worker indifferent at the respective income cut-offs, we must have $\underline{a}_s(\delta)$ lower than $a_4(\delta)$ as well as $a_6(\delta)$.

$a_4(\delta) > a_6(\delta)$ if and only if $\theta(1-\beta) > \beta$.

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

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

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

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

3. We show $a_1(\delta)$, $a_3(\delta)$ and $a_5(\delta)$ are finite if and only if $\delta > \underline{\delta} \equiv (1-\bar{s})^{\sigma} - 1$.

   Let us write the income cut-off $a_1(\delta)$ expression (A.11) as:

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

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

   We follow analogous reasoning and use equations (A.9) and (A.10) to show the same for the income-cutoffs $a_5(\delta)$ and $a_3(\delta)$.

   $1 < a_1(\delta)$ if and only if $\delta < \delta_a$.

   Suppose $a_1(\delta) \gtreqless 1$. Using this in (A.11) we get

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

   This proves the claim.

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

   First we show $a_5(\delta) < a_1(\delta)$ Comparing (A.11) and (A.9) we get,

$$\begin{aligned}
&\delta\Big[1 - \big[\theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}a_1(\delta) + 1 - \theta\beta\big]^{\sigma}\Big] = \delta\Big[1 - \big[\theta\beta^{\phi}a_5(\delta) + 1 - \theta\beta\big]^{\sigma}\Big] \\
&\Rightarrow \quad \theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}a_1(\delta) = \theta\beta^{\phi}a_5(\delta) \\
&\Rightarrow \quad [\theta(1-\beta)+\beta]^{-(1-\phi)}a_1(\delta) = \beta^{-(1-\phi)}a_5(\delta) \\
&\Rightarrow \quad a_1(\delta) > a_5(\delta) \quad \text{as } \theta(1-\beta)+\beta > \beta
\end{aligned}$$

Now, we show $a_3(\delta) < a_1(\delta)$ Comparing (A.11) and (A.10), we get

$$\delta\Big[1 - \big[\theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}a_1(\delta) + 1 - \theta\beta\big]^{\sigma}\Big] = \delta\Big[1 - \big[\theta\beta[\theta(1-\beta)]^{-(1-\phi)}a_3(\delta) + 1 - \theta\beta\big]^{\sigma}\Big]$$
$$\Rightarrow \quad \theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}a_1(\delta) = \theta\beta[\theta(1-\beta)]^{-(1-\phi)}a_3(\delta)$$
$$\Rightarrow \quad a_1(\delta) > a_3(\delta) \quad \text{as } \theta(1-\beta)+\beta > \theta(1-\beta)$$

$a_3(\delta) > a_5(\delta)$ if and only if $\theta(1-\beta) > \beta$ Comparing (A.10) and (A.9) we get,

$$\delta\Big[1 - \big[\theta\beta[\theta(1-\beta)]^{-(1-\phi)}a_3(\delta) + 1 - \theta\beta\big]^{\sigma}\Big] = \delta\Big[1 - \big[\theta\beta^{\phi}a_5(\delta) + 1 - \theta\beta\big]^{\sigma}\Big]$$
$$\Rightarrow \quad \theta\beta[\theta(1-\beta)]^{-(1-\phi)}a_3(\delta) = \theta\beta^{\phi}a_5(\delta)$$

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

4. First, we show $\forall\ \delta \leq \delta_a$, $a_4(\delta) \leq a_1(\delta)$.

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

   Now, we show $\forall\ \delta > \delta_a$, $a_4(\delta) < 1$.

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

4. Follows from comparing definitions (A.1) and (A.5) we get $\underline{b}_s(\delta) = \underline{a}_s(\delta)$.

   $\underline{b}_u(\delta) < a_5(\delta)$.

   We show that the benefit from investment with behavioral anomaly at $\underline{b}_u(\delta)$ when $\theta < 1$ is strictly lower than that when $\theta = 1$, i.e.

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

   so, $a_5(\delta) \geq \underline{b}_u(\delta)\ \ \forall\ \theta \in [0,1]$, and strictly higher $\ \forall\ \theta < 1$.
   Let $x = \theta\beta^{\phi}\underline{b}_u(\delta) + 1 - \theta\beta$. Then,

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

Observe at $\underline{b}_u = \beta^{1-\phi}$, the benefit from investment when $\theta = 1$ is zero and cost is positive, but $\underline{b}_u(\delta)$ should be such that the benefit is equal to cost. The benefit is increasing in m_{st}, hence, $\underline{b}_u(\delta)$ must be greater than $\beta^{1-\phi}$. Hence, we retrace the steps to find $\underline{b}_u < a_5$. $\square$

SC.4 Figures

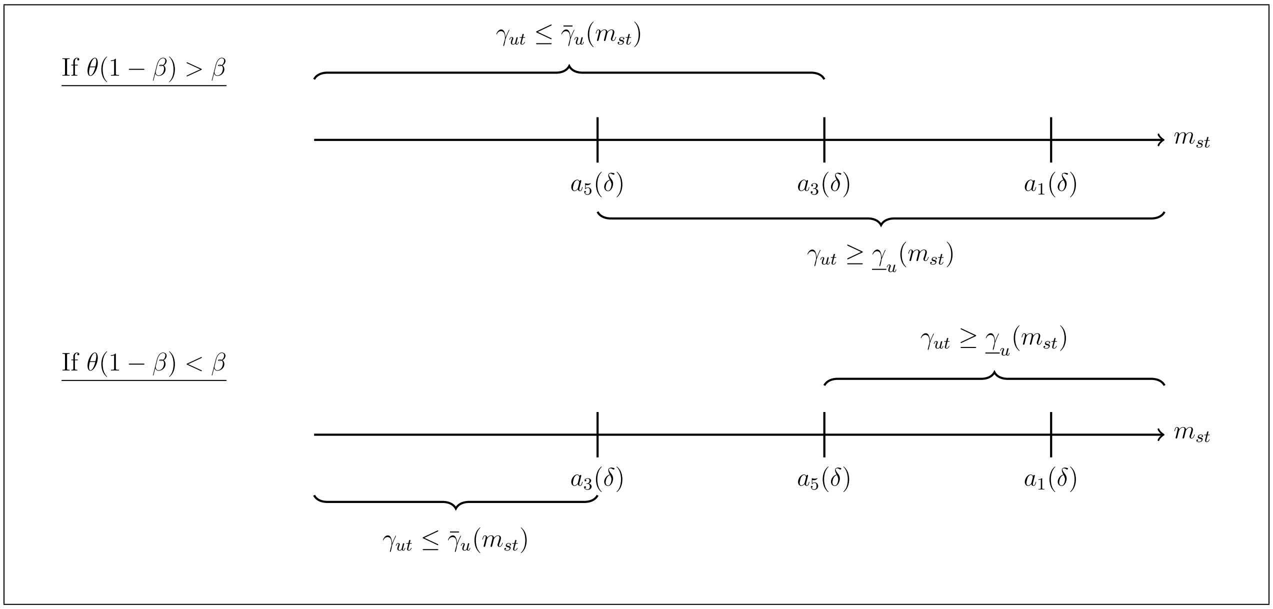

Figure 11: Condition $\underline{\Gamma}_u$ and $\bar{\Gamma}_u$

Figure 12 depicts an example of equilibrium trajectory for $\theta(1-\beta) > \beta$. Interestingly, here, the probability of investment by skilled workers weakly increases over time but not for the educated-unskilled worker.

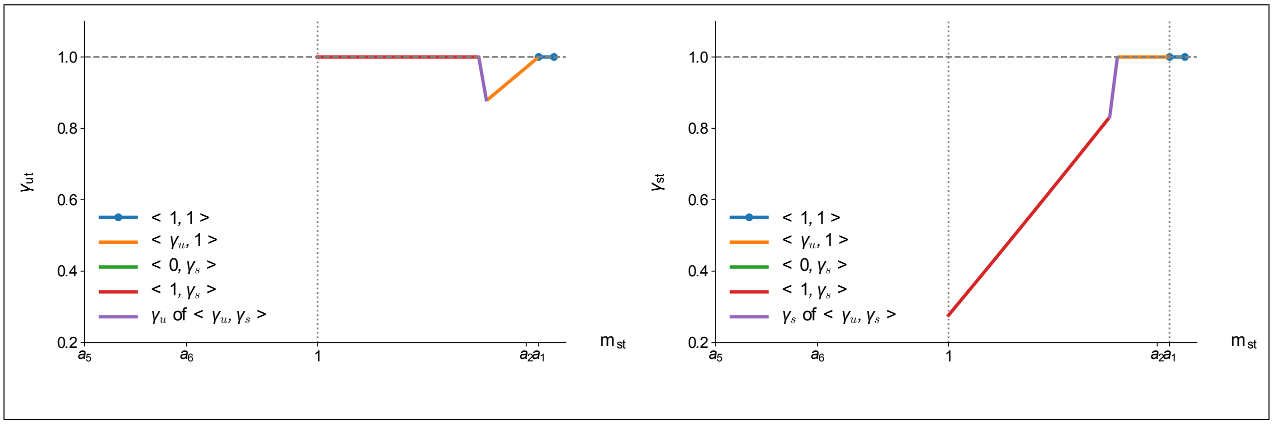

Figure 12: Example to depict equilibrium for $\theta(1-\beta) > \beta$.

SD Appendix for Comparison

SD.1 Proof of Observation 10

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

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

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

$$\theta\beta(\theta(1-\beta)+\beta)^{-(1-\phi)}+1-\theta\beta < A\beta^{\phi}+1-\beta \tag{S.3}$$

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

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

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

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

Thus, equation (S.3) always holds and hence $\bar{\delta} < \delta_a$. □