Perceptions, Biases, and Inequality*

Dyotona Dasgupta† Anuradha Saha‡

May 20, 2022

Abstract

This paper introduces *perceived self-efficacy beliefs* in an overlapping generations model and studies the effects of behavioral anomalies on human capital investments and aggregate inequality. These beliefs are not inherent but based on a person's socio-economic background. *Ex ante* children are homogeneous, but parents are biased about the returns to their children's education. Based on parental education and job status, parents may be over or under confident but otherwise, they are rational. We witness *spillover effects* that shift the incentives to educate children from the pessimistic parents to unbiased parents. Biases may influence an under-confident parent to invest less or, at times, even more. Same holds for the overconfident parent. Over investment by one parent type may even *crowd-out investments* of other parent types. We see interesting effects of biases on the aggregate economy when parental warm glow is not low. For moderate parental warm glow, behavioral biases may give rise to multiple equilibria as well as lower the steady-state inequality. For huge parental warm glow, poor adults may invest with a higher probability than the rich.

Keywords: Behavioral Inequality, Human Capital Investment, Behavioral Bias, Intergenerational Mobility

JEL Codes: J62, D91, E2

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

†Corresponding author. Jindal Global Business School, O.P. Jindal Global University, Sonepat, Haryana 131029. India. E-mail: dasgupta.dyotona@gmail.com

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

1 Introduction

Socio-economic background plays an important role in investment decisions – through access to opportunities as well as how they influence *beliefs* on the benefits from different opportunities. For example, a high-school educated parent who owns a small business may invest less in their child's schooling due to their limited financial capabilities or because they *do not think* education would significantly help their child's employment prospects. Chowdry et al. (2011) show that attitudes of youngsters and their parents play a key role in explaining the rich-poor gap in secondary education attainment. Beliefs and biases matter.

In this paper, we construct a theoretical model to capture the fact that socio-economic background affects beliefs about self-efficacy and we study their long-run effects.[1] In the model, we depict both over and under-confident parents. As we would expect, behavioral anomalies may lead to lower educational investment and worsen inequality in the economy. More interestingly, though, it does not always lead to negative outcomes – behavioral biases may reduce income inequality. Off the steady state, along the dynamic path, biases present other interesting outcomes. Own biases may result in both higher and lower investments. Biases of other types of parents may crowd-in or crowd-out investments of a parent. These effects may influence some less affluent parents to invest in their children's education with a higher probability relative to the richer parents.

In our model, agents experience internal constraints. This framework allows us to identify how behavioral anomalies interact with socio-economic attributes to create differential outcomes. For example, in an empirical paper Chetty et al. (2019), establish that black-white inter-generational income gap for children, at the same level of parental income, is substantial for men *but not for women.*[2] Even after controlling for family or children's characteristics, the authors are not able to explain why black-white inter-generational gap is so small for females. In an environment where biases against black men are more widely prevalent compared to those against black women (Opportunity Insights), incorporating these behavioral biases appears essential to explain the differential racial gaps across gender.

While research on the effects of behavioral anomalies is limited, it is increasingly becoming imperative to study behavioral inequality. We should do so for two reasons. First, in reality individuals are multi-faceted and they do not rank the same across different attributes. People could be fortuitous in one dimension but not in others, for example a rich uneducated carpenter versus a poor college teacher. There is a need to formulate a generalized theory which captures such multi-dimensional people and depicts how such individuals access and evaluate their choices. We build such a theory where it is fundamentally *beliefs* which influence how people evaluate different opportunities. By incorporating both under confident and overconfident groups, we are able to

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

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

highlight how such disparate biases affect the aggregate economy. Second, inequality stemming from external constraints limits the opportunities available to a person while behavioral inequality limits how individuals respond to opportunities. Our paper advocates for behavioral public policies that help change culture and which is usually amiss in conventional poverty alleviation programmes.

In an overlapping generations model, we build a theoretical framework where a decision maker's assessment of returns to investment maybe biased. Adults differ in terms of their education – educated or uneducated, and jobs – skilled or unskilled. There is a fixed cost of acquiring education. Education is necessary, though not sufficient, for getting a skilled job. There are no intrinsic differences among educated children, i.e. the exogenous probability of an educated child getting a skilled job is the same for all. Naturally, skilled workers earn more than the unskilled workers. Adults derive utility from their own consumption and the perceived expected income of their children, and thus, they decide whether to invest in their children's education.

A biased adult forms beliefs about the probability of their educated child getting a skilled job. These beliefs are based on the parent's education and job identity. When two adults share fewer common attributes, the "degree of association" is weaker and they believe that their own child is less likely to work in the dissimilar adult's occupation. Due to higher dissimilarity with the skilled workers, an uneducated-unskilled is extremely pessimistic about her child getting a skilled job, and an educated-unskilled worker is under confident, while an educated-skilled worker is overconfident.[3] All agents are non-Bayesian and there is no learning[4] or convergence of beliefs. However, *given their beliefs*, each parent correctly calculates the equilibrium mass of future skilled workers and their income. Less confident people overestimate the reward in case of success and over confident people do the converse. These two opposing forces – confidence and estimated rewards – determine the investment decision of any individual. While all agents have economic resources (though with varying utility costs) to invest, some adults do not invest as they perceive the returns not to be sufficiently high. This captures the effect of cognitive limitations on investment.

We find that the relative weight a parent places on the utility from her child's perceived expected income plays a key role in the characterization of dynamic and steady state properties. We call this the *parental warm glow* (or intergenerational altruism as in Ghatak (2015)). In our characterization, the economic outcomes have distinct properties as per three ranges of parental warm glow, namely, low, moderately high, and huge.

We highlight four of our main findings. First, behavioral biases cause poverty traps. When parents do not experience any behavioral anomalies and when warm glow is not low, there does not exist any poverty trap. However, the inclusion of extremely pessimistic adults in this range of warm glow, gives rise to poverty trap.[5] Second, contrary to common opinion, biases may lower

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

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

[5]In another paper Dasgupta and Saha (2022), we show a stronger result. When warm glow is not-low, even small

long-run income inequality. We find that for moderately high warm glow, educated parents may over invest in education, relative to the benchmark (where no one is biased). Here, since the mass of skilled workers is higher, the skilled income and income inequality is lower in the steady state. Third, biases not only affect oneself but also other biased and unbiased agents. For example, under confidence may influence educated-unskilled parents to invest more than the skilled parents because the perceived gains appear larger to them. Or, pessimism of uneducated workers may create spillover effects whereby educated workers tend to invest more in education. Further, over investment of one type of parents may crowd-out investment of other type of parents. Finally, there could be multiple equilibria, which stems from the fact that the perceived expected benefit from an educational investment for each type of parent is different and can not be unambiguously ranked. In a dynamic setting, this would cause aggregate fluctuations in investment and income, or lead to *behavior-driven business cycles*.

In terms of policy making, our paper provides a framework to understand systematic behavioral inequality, for example biases against women in STEM careers (Sterling et al. (2020)) and in terms of black-white investment in education in the United States (Klugman and Xu (2008)). These are not problems of resource constraints. The confidence gaps in female STEM or black students remains uncorrected because biases are difficult to change. Here, *behavioral policies* are required and we advocate for creating policies appropriate for local conditions. Our paper provides a theoretical foundation of behavioral policy-making, which has been advocated by Chetty et al. (2019), BRAC, corporate professionals (Forbes), among others.

The plan of the paper is as follows. Next, we discuss the related literature. Section 2 lays out the general framework of the economy. Section 3 studies the benchmark case where there is no behavioral anomaly. Section 4 addresses the case where all workers are biased. In section 5 we discuss the effects of the behavioral anomaly. Section 6 concludes with some policy implications. Main proofs are collected in an Appendix. We provide the formal expressions of various definitions, and conditions, lemmas, and observations along with their proofs in the Supplementary Appendix.

1.1 Related Literature

This paper is related to various strands of literature. There is a large literature addressing the physical constraints such as credit market imperfection, non-convex technology (Banerjee and Newman (1993), Galor and Zeira (1993), Mookherjee and Ray (2002)) to explain persistent inequality and poverty trap. Like Galor and Zeira (1993), we have a fixed cost of education but unlike them, in our model parents make educational investments and they always have sufficient income to do so. Without any biases, only at low parental warm glow there is a poverty trap. Otherwise, when parents care enough about their children's future, there isn't any poverty trap. The inclusion of behavioral anomalies, in our paper, brings a new source of poverty trap. We find

biases may cause a poverty trap.

that when parental warm glow is not-low there is a poverty trap. Hence, we contribute to the emerging literature on behavioral inequality.[6]

We depict behavioral biases differently from the existing literature. Behavioral theory on inequality and poverty, to the best of our knowledge, has relied on non-standard utility functions – time inconsistency (captured through quasi-hyperbolic discounting as in Bernheim et al. (2015)), or temptation (as in Banerjee and Mullainathan (2010)), or aspiration (captured through 'milestone' utility in Ray (2006), Dalton et al. (2016) and subsequent papers). In contrast, the agents, in our model, maximize a standard lifetime utility function. The anomaly comes from the fact that their judgement about future prospects is prejudiced by their own life experiences.

This paper brings the role of self-confidence in inter-generational investments. Traditionally, it was thought that "the most robust finding in the psychology of judgment is that people are overconfident" (De Bondt and Thaler, 1995).[7] 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.

Moreover, there is empirical evidence that social background, gender, wealth, among others can influence individual's self-confidence. Filippin and Paccagnella (2012) develop a theoretical model where small differences in self-confidence lead to gaps in human capital. Unlike us, they consider a lack of self-confidence as an information problem while we look at it as a behavioral problem. In our model, people engage in self-deception; they believe that the probability of change in economic status is low.

2 Model

2.1 The Firms

In a discrete-time framework, we consider a two-sector economy. The production function of the skilled sector is AL_{st}^{ϕ}, where L_{st} is the mass of skilled workers,[8] $0 < \phi < 1$, and $A \geq 1$ – the production function is strictly increasing and strictly concave. At any period t, a skilled worker earns wage and also the profit of the skilled sector which is divided among the skilled workers equally. Hence, the income of a skilled worker is $AL_{st}^{-(1-\phi)}$. The production function of the unskilled sector is L_{ut}. The unskilled sector does not make any profit, and the income of

[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).

[7] Overconfidence has been studied in housing market (Case and Shiller, 2003), among CEOs (Malmendier and Tate, 2005), in financial investment (Biais et al., 2005), among others. It has been argued that self-confidence enhances motivation (Bénabou and Tirole, 2002), improves self-control (Bénabou and Tirole, 2004).

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

an unskilled worker is just her wage, 1. The total mass of workers is normalized to 1. In the Supplementary Appendix, we show that a skilled worker earns (weakly) more than an unskilled worker.

2.2 The Households

We consider an overlapping generations model with no population growth. An individual lives for two periods: first as a child and later as a parent. The parent derives utility from her own consumption[9] and her child's *conjectured* income. The conjectured income depends on their beliefs about the probabilities with which an educated child would get a skilled job.

Education is necessary but not sufficient for becoming a skilled worker. A parent decides whether to educate their child or not. An educated individual gets a job in the skilled sector with probability β, whereas an uneducated person gets work in the unskilled sector with certainty. Required investment in education is fixed at $\bar{s}$, where $\bar{s} \in (0,1)$. The utility of a parent of type ij, where i denotes her education $i \in \{e,n\}$ and j denotes her skill $j \in \{s,u\}$, is

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

Here $\delta(>0)$ is the parental warm glow, c_t^{ij} and $E\omega_{t+1}^{ij}$ denote the parent's consumption and the conjectured income of her only child respectively. Observe, the utility function is strictly increasing and strictly concave.[10]

The investment decision on a child's education is made on the basis of the conjectured income of a child. There is no inherent difference in the probability of getting a skilled job across educated children of different parent types. 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 beliefs about the probability of a child from her community (education-job based community) becoming a skilled worker*, she accurately calculates the mass of skilled workers in the next period and makes the investment decision accordingly. An equilibrium has two features:

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

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

We start our analysis with unbiased parents. Section 4 analyses the case with biased parents.

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

[10]This is special case of a CRRA utility function, where $u(c) = (c^{1-\rho} - 1)/(1-\rho)$, and $\rho \geq 0, \rho \neq 1$. For sharpness of results, we have assumed that the parents are quite risk averse, $\rho > 1$.

3 Benchmark Case

In this section, parents are not biased – all types of parents believe that the probability of an educated child getting work in the skilled sector is β, which is the true probability. Thus, parental decisions differ only due to differences in their incomes.

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

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

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

$$\delta \left[\frac{\big[\beta \cdot m_{st+1} + (1-\beta) \cdot 1\big]^{\sigma}}{\sigma} - \frac{1}{\sigma} \right] \geq \frac{m_{jt}^{\sigma}}{\sigma} - \frac{(m_{jt} - \bar{s})^{\sigma}}{\sigma}. \tag{1}$$

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

An equilibrium is denoted by $\langle \lambda_{ut}, \lambda_{st} \rangle$ which satisfies the features described in Section 2.2. As the utility cost of investment for the skilled workers is strictly lower than the unskilled workers, at any equilibrium, if an unskilled worker invests with a positive probability, the skilled workers invest with certainty.

The warm glow parameter plays an important role in the parent's investment decision. We define three key thresholds of this parameter which will be useful in further analyses.

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

Observe, these thresholds are directly related to the fixed cost of education, $\bar{s}$.

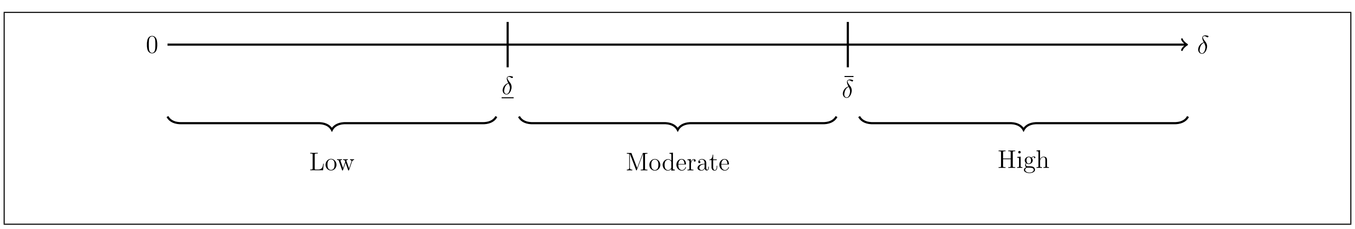


Figure 1: Ranges of Parental Warm Glow Parameter

The thresholds of warm glow parameter is defined in such a way that when it is high, investing with probability one is a strictly dominating strategy of the unskilled workers. When δ is low, the unskilled workers never invest. Further for any given parental warm glow, we find that the equilibrium probability of investment (weakly) decreases with the income of a skilled worker.

Accordingly, we define three thresholds of skilled income[11] in order to characterize the equilibria completely.

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

- $\underline{b}_s(\delta)$ *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* $m_{st} \leq \underline{b}_u(\delta)$.

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

Proposition 1. Characterization of the Equilibria *For any parameter value and at any state variable m_{st}, the equilibrium is unique:*

1. *For high warm glow at any m_{st}, all parents invest with probability 1.*
2. *When warm glow is moderate, at any m_{st}, the unskilled parents never invest with certainty while the skilled parents always invest with a positive probability. Specifics are in Figure 2.*
3. *When warm glow is low, at any m_{st}. Unskilled parents never invest while the probability of investment of skilled parents increases with m_{st}, as depicted in Figure 2.*

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

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

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

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

Low | Moderate | High

Figure 2: Characterization of the Equilibria in the Benchmark Case

We prove this in Appendix 7.1.

Uniqueness of equilibrium follows from the fact that the benefits for both types of workers are equal while the utility cost of investment for the skilled workers is strictly lower. Thus when the warm glow is high as the unskilled workers invest with probability one (from the definition of high warm glow), the skilled workers also invest with probability one. When the warm glow factor

[11] The formal expressions, the properties of these thresholds, and the proof of above mentioned equilibrium property, $\lambda_{st} = 1$ whenever $\lambda_{ut} > 0$, can be found in the Supplementary Appendix.

is high, the parents care for their children so much that they invest irrespective of their income. When the warm glow factor is moderate, the unskilled workers no longer invest with certainty and the probability of investment decreases with a decrease in the state variable. If the state variable falls below $\underline{b}_u(\delta)$, the unskilled workers do not invest at all. The investment decisions of the skilled worker can be explained along the similar lines. For low warm glow, using analogous reasoning we find that unskilled workers never invest and skilled workers invest only for sufficiently high skilled incomes.

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

Proposition 2. Dynamics and the Steady States

1. *When warm glow is not low, there is no poverty trap in the economy.*

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

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

2. *When warm glow is low there is a poverty trap. The economy converges to the unique steady state where no parents invest and all workers are unskilled.*

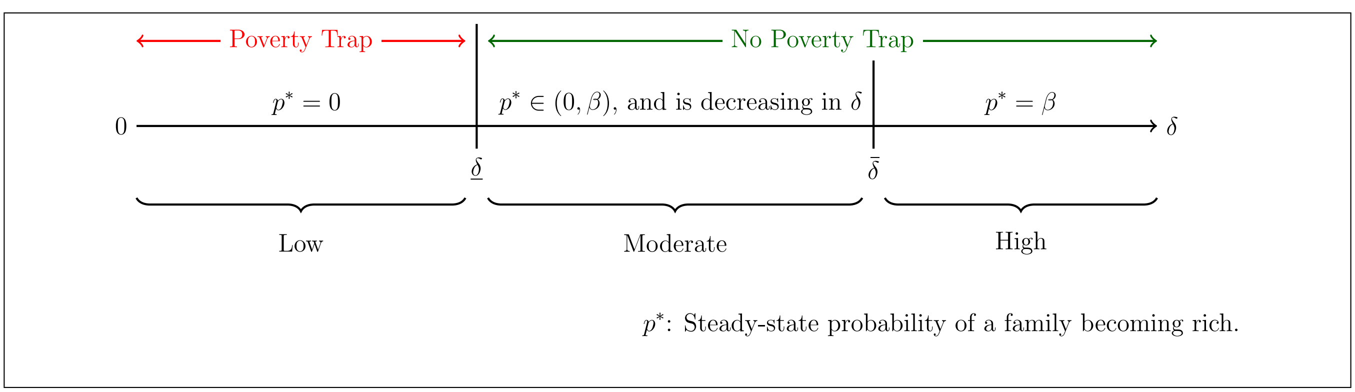


Figure 3: Steady State Properties without Behavioral Bias

We prove this in Appendix 7.2.

The intuition for the dynamics is easy to understand from the equilibrium strategies. For high warm glow, since all workers invest with certainty, all children are educated, mass of skilled workers is β and the income of skilled workers immediately converges to $A\beta^{-(1-\phi)}$. For moderate warm glow, over time the skilled income increases at a decreasing rate. When initial skilled income is less than $\bar{b}_s(\delta)$, as unskilled workers do not invest skilled, skilled income unequivocally increase with time. It continues to do so until it hits the unique steady state of $\beta^{-(1-\phi)}\underline{b}_u(\delta)$. Finally, when warm glow is low, as unskilled never invest, the mass of educated declines over time. In the long run, everyone is uneducated and unskilled.

In terms of steady-state inequality, economies with moderate warm glow are more unequal than economies with high warm glow. Since the steady-state mass of educated workers is increasing in warm glow parameter, skilled workers earn less in economies with higher warm glow. For low warm glow, everyone is unskilled in the steady state, and hence, there is zero income inequality.

The focus of this paper is to understand the implications of behavioral anomalies on investment decisions and their long-term impact on poverty trap and inequality. In the benchmark, i.e. in the absence of any behavioral anomalies, we find that there is a poverty trap only for economies with low parental warm glow. In the remaining sections, we focus on the case where warm glow is not low. This allows us to analyze whether behavioral anomalies give rise to poverty traps. In further analyses we investigate how behavioral anomalies produce differential outcomes.

4 Behavioral Anomaly

In this section, we model behavioral anomaly whereby parents underestimate the probability of their children's intergenerational mobility. This stems from the fact that parents identify themselves with some groups and they find it unlikely for their children to be part of other groups. A group is represented by a set of features or attributes, namely education and job. The similarity between groups increases with the addition of common features. An individual feels less connected with more dissimilar groups (following Tversky (2004) pp. 10-11). We capture this through a *"degree of association"*. The degree of association between two individuals belonging to the same group is normalized to 1. Let the degree of association between two individuals belonging to two groups which differ by one attribute be θ and that when they differ by two attributes be η, so $\eta < \theta \in [0, 1)$.[12] Thus, the degree of association of an educated-unskilled worker with an educated-skilled worker is θ and that of an uneducated-unskilled worker with an educated-skilled worker is η.[13]

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

[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.

worker of a different type than herself via the degree of association. Recall, the true probability with which an educated child becomes a skilled worker is independent of her parent's group identity. So, discounting the probability of intergenerational mobility captures bias in our model which has been documented in several empirical findings (Filippin and Paccagnella (2012), Goel and Deshpande (2020), for example).

We assume that $\eta = 0$, that is uneducated parents *extremely pessimist* – they believe that an educated child from their group would never get a skilled job. As parent invests only when that provides her (weakly) higher utility, the immediate implication of η being zero is that an uneducated worker never invests.

An educated-unskilled worker believes that an educated child from her group becomes a skilled worker with probability $\theta\beta$. Analogously, a skilled worker believes that with probability $\theta(1-\beta)$ an educated child from her group becomes an unskilled worker.[14] So, she believes the probability with which such a child becomes a skilled worker is $1-\theta(1-\beta)$. As $\theta < 1$, educated-unskilled workers are under confident and skilled workers are overconfident.[15] We assume that parents are unbiased in assessing the probability with which an educated child from a different group becomes a skilled worker. It is as if parents believe the statistical evidence for other groups but not their own.

Since a parent's belief about the probability of success – an educated child from her own group becomes a skilled worker – is type dependent, the 'conjectured' mass of skilled workers and their income would also be type dependent. Suppose, at period 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

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

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

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

Similarly, a skilled worker conjectures that the mass of skilled workers and their income:

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

[14] Recall, $1-\beta$ is the probability with which an educated individual becomes an unskilled worker.
[15] $\theta\beta < \beta \leq 1-\theta(1-\beta)$.
[16] Here also, we use subscript u for educated-unskilled workers and subscript s for the skilled workers.

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

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

At any period t, an educated-unskilled worker invests with probability γ_{ut} if and only if

$$\frac{(1-\bar{s})^{\sigma}}{\sigma}+\delta\frac{\Big[\theta\beta[\theta(1-\beta)\gamma_{ut}+\beta\gamma_{st}]^{-(1-\phi)}m_{st}+1-\theta\beta\Big]^{\sigma}}{\sigma}\geq\frac{1}{\sigma}+\frac{\delta}{\sigma}$$
$$\Rightarrow\quad \delta\left[\frac{\Big[\theta\beta[\theta(1-\beta)\gamma_{ut}+\beta\gamma_{st}]^{-(1-\phi)}m_{st}+1-\theta\beta\Big]^{\sigma}}{\sigma}-\frac{1}{\sigma}\right]\geq\frac{1}{\sigma}-\frac{(1-\bar{s})^{\sigma}}{\sigma} \tag{2}$$

and when γ_{ut} is a fraction, (2) holds with equality. The L.H.S. is the *conjectured* net benefit and the R.H.S. is the net utility cost from investment. Similarly, at any t, a skilled worker invests with probability γ_{st} if and only if

$$\delta\left[\frac{\Big[[1-\theta(1-\beta)]\cdot\big[(1-\beta)\gamma_{ut}+[1-\theta(1-\beta)]\gamma_{st}\big]^{-(1-\phi)}m_{st}+\theta(1-\beta)\Big]^{\sigma}}{\sigma}-\frac{1}{\sigma}\right]$$
$$\geq\frac{m_{st}^{\sigma}}{\sigma}-\frac{(m_{st}-\bar{s})^{\sigma}}{\sigma} \tag{3}$$

and when γ_{st} is a fraction, (3) holds with equality. The utility cost of investment is lower for a skilled worker. But, the *conjectured* benefits from investment cannot be ranked. This is because the under confident educated-unskilled workers underestimate the mass of future skilled workers and hence overestimate their income. The overconfident skilled workers do the opposite.[17] So unlike the benchmark case, the skilled workers do not necessarily invest with higher probability when the biased and educated-unskilled workers invest with a positive probability.[18]

Due to this observation, in order to characterize the equilibria, we need boundary conditions on the equilibrium probabilities of investment of both types of workers. Given a state variable m_{st}, the equilibrium probability of investment of one type of workers decreases with the probability of investment of the other type of workers. Accordingly for any m_{st}, we define the lower bound on equilibrium γ_{ut} when γ_{st} is one and similarly, the upper bound on equilibrium γ_{ut} when γ_{st} is zero. Similarly, the lower and upper bounds on γ_{st} are determined when γ_{ut} are one and zero respectively. For any state variable m_{st}, consider any equilibrium $\langle\gamma_{ut},\gamma_{st}\rangle$, then γ_{ut} is bounded below and above by $\underline{\gamma}_u(m_{st})$ and $\bar{\gamma}_u(m_{st})$, and γ_{st} is bounded below and above by $\underline{\gamma}_s(m_{st})$ and $\bar{\gamma}_s(m_{st})$ respectively.

[17]This captures the Contrast Effect introduced by Tversky and Griffin (2004)

[18]In the Supplementary Appendix, we show that at any equilibrium, if an educated-unskilled worker invests, then any biased skilled worker invests with a *positive* probability.

Definition 3. Upper and Lower Bounds on Equilibrium Probability of Investment

- $\underline{\gamma}_s(m_{st})$ *(or* $\bar{\gamma}_s(m_{st})$*) is the optimal probability of investment by the skilled workers when educated-unskilled workers invest with probability 1 (or 0 respectively) at* m_{st}*.* $\underline{\gamma}_s, \bar{\gamma}_s \in [0, 1]$ *are derived from eq.* (3)*.*

- $\underline{\gamma}_u(m_{st})$ *(or* $\bar{\gamma}_u(m_{st})$*) is the optimal probability of investment by the educated-unskilled workers when skilled workers invest with probability 1 (or 0 respectively) at* m_{st}*.* $\underline{\gamma}_u, \bar{\gamma}_u \in [0, 1]$ *are derived from eq* (2)*.*

The formal expressions and properties of these bounds can be found in the Supplementary Appendix.

Recall, we have confined our analysis to economies with not-low parental warm glow. We find in this range $\delta > \underline{\delta}$, at least one type of workers invest with a positive probability. Now, when the warm glow is *huge*, investing with probability one is a strictly dominating strategy for educated-unskilled workers. Accordingly, we define a new threshold of warm glow parameter. Further, we club the moderate and high-but-not-huge warm glow into *moderately high* as the equilibrium properties have similar features.

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

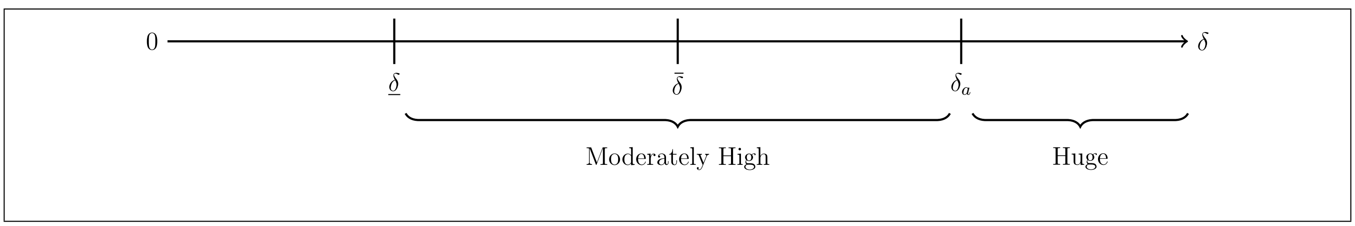


Figure 4: Relevant Ranges of Parental Warm Glow to Analyze the Implications of Biases

Finally, for a given warm glow parameter, we define two thresholds of the state variable which play a crucial role in characterizing the equilibria:

Definition 5. *For a given warm glow factor,*

- *suppose the educated-unskilled workers invest with probability 1, then* $\bar{a}_s(\delta)$ *is the minimum value of the state variable at which skilled workers invest with probability 1,*

- *suppose the skilled workers invest with probability 1, then* $\bar{a}_u(\delta)$ *is the minimum value of the state variable at which educated-unskilled workers invest with probability 1.*

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

Proposition 3. Characterization of the Equilibria

1. *Uneducated workers never invest.*

2. *When warm glow is huge, at any $m_{st} \geq 1$, the equilibrium $\langle \gamma_{ut}, \gamma_{st} \rangle$ is unique: The educated-unskilled workers invest with probability 1, and skilled workers invest with $\underline{\gamma}_s(m_{st})$ as in Definition 3 .*

3. *When the warm glow is moderately high, at least one type of educated workers invest with a positive probability. Further,*
   - *If $\bar{a}_u(\delta) \leq \bar{a}_s(\delta)$, then $\forall m_{st} > \bar{a}_u(\delta)$ $\gamma_{ut} = 1$ and $\gamma_{st} = \underline{\gamma}_s(m_{st})$ as in Definition 3 .*
   - *If $\bar{a}_u(\delta) > \bar{a}_s(\delta)$, then $\forall m_{st} > \bar{a}_s(\delta)$ $\gamma_{st} = 1$ and $\gamma_{ut} = \underline{\gamma}_u(m_{st})$ as in Definition 3.*
   - *At any $m_{st} < min\{\bar{a}_u(\delta), \bar{a}_s(\delta)\}$, γ_{ut} is bounded below and above by $\underline{\gamma}_u(m_{st})$ and $\bar{\gamma}_u(m_{st})$ respectively; and γ_{st} is bounded below and above by $\underline{\gamma}_s(m_{st})$ and $\bar{\gamma}_s(m_{st})$ respectively. Two other intersting characteristics of the equilibria: (a) if $\beta \geq \theta(1-\beta)$ then there could be multiple equilibria, (b) if $\beta < \theta(1-\beta)$ and $\bar{a}_u(\delta) < \bar{a}_s(\delta)$ then $\gamma_{st} < \gamma_{ut}$.*

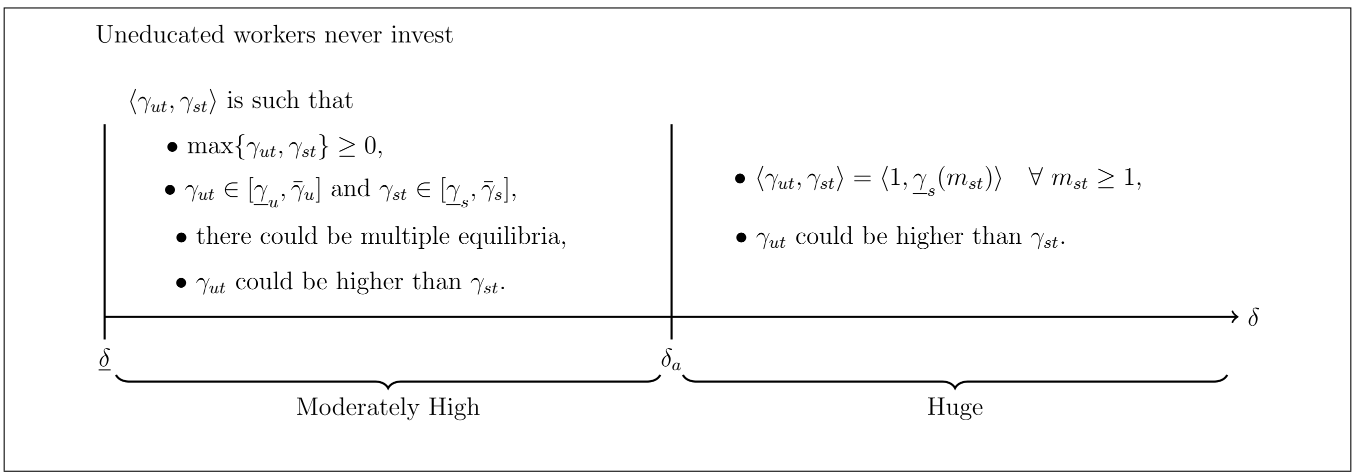


Figure 5: Characterization of the Equilibria with Behavioral Anomaly

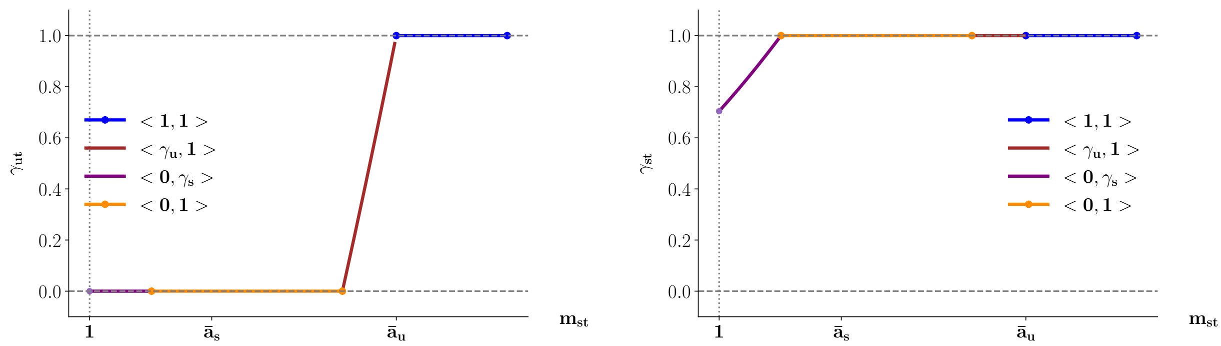
We prove this in Appendix 7.3 and Figure 5 depicts the equilibria at a glance.

Introduction of behavioral anomaly produces two key outcomes. First, educated-unskilled workers, in spite of their lower incomes and under confidence, can invest with a higher probability than skilled workers. Because agents are otherwise rational, under confident workers overestimate the rewards from education. Hence biases may embolden educated-unskilled workers. Second, there may exist multiple equilibria. This stems from the fact that the conjectured benefits from investment can not be ranked for the two types of educated workers.

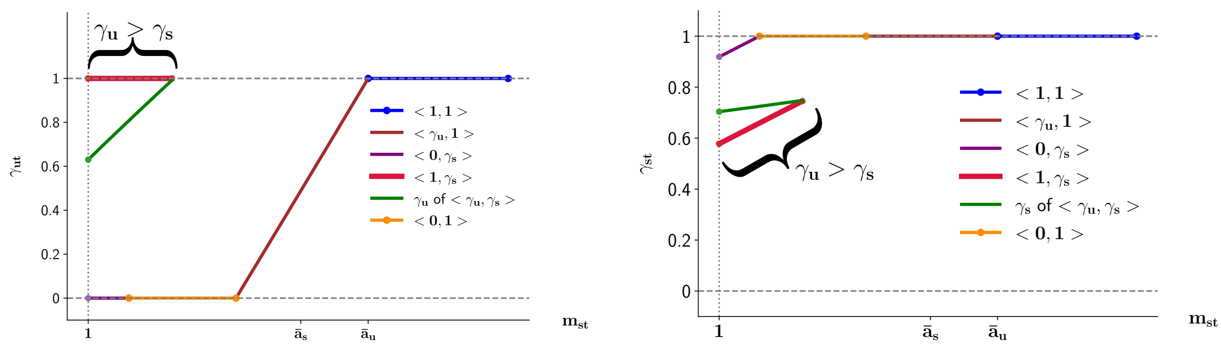
We depict numerical examples to show that when $\beta \geq \theta(1-\beta)$, depending on the parametric conditions, there can be unique or multiple equilibria.[19] We consider $\theta = 0.4$ and $\beta = 0.7$.

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

Depending on parental warm glow and returns to scale in the skilled sector ϕ, there could be unique equilibrium at every m_{st} (as in Figure 6a) or multiple equilibria for some m_{st} (as in Figure 6b).



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



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

Figure 6: Example: Unique and Multiple equilibria when $\beta > \theta(1 - \beta)$. [Colored Graphs]

Now the equilibrium properties are driven by two forces: extreme pessimism of uneducated workers and biases of educated workers. To understand the effects of extremely pessimistic beliefs, we assume $\theta = 1$ and find,

Corollary 1. Consequences of Extreme Pessimism Only ($\eta = 0$ and $\theta = 1$). *When uneducated workers are extremely pessimistic and educated workers are not biased then*

1. *uneducated workers never invest in their children's education,*
2. *educated-unskilled workers always invest with a (weakly) lower probability than skilled workers,*
3. *compared to the benchmark, at a given m_{st} the probability of investment of educated workers is (weakly) higher in the presence of extremely pessimistic uneducated workers.*

The first result follows from the extreme pessimism of uneducated workers. The second result is based on the fact that, at $\theta = 1$, the conjectured benefit for all educated workers is the same, and that the utility cost of investment of educated-unskilled workers is higher. The third result is interesting. Extreme pessimism affects the economy only where, to begin with, the uneducated invested with positive probability. Now, due their pessimism, uneducated do not invest. This reduces the mass of educated children and skilled workers, and increases future skilled incomes. This incentivises unbiased educated-unskilled to invest with a higher probability than the benchmark. Skilled, however, were already investing with probability one and hence do not get affected by the extreme pessimism of uneducated. Notice, even though only uneducated workers are pessimistic, it affects society at large.

Now we return our focus to the dynamics and steady state of the economy with all biased agents. We find that there are multiple steady states and they can be ranked in terms of inequality. The steady state where $a^*(\delta) = \max\{\bar{a}_u(\delta), \bar{a}_s(\delta)\}$, we call that the 'least unequal steady state'.

Proposition 4. Dynamics and the Steady States

1. *There is almost always a poverty trap in an economy.*
2. *Steady States: Any $m_{st} \geq a^*(\delta)$ is a steady state where all educated workers invest with probability 1. Steady-state income of a skilled worker is initial income $m_s^* = m_{st}$. Higher the steady-state income, larger is the steady-state income inequality.*
3. *Dynamics: If $m_{st} < a^*(\delta)$: at least one type of educated workers invest with probability between 0 and 1. The mass of educated individuals and that of skilled workers decrease over time. Skilled income increases over time and converges to some $m_s^* \geq a^*(\delta)$.*

We prove in Appendix 7.4.

The existence of poverty trap may seem driven by the assumption of extreme pessimism ($\eta = 0$). However, in Dasgupta and Saha (2022), we show that even when uneducated are slightly biased and educated are not at all biased, there could be poverty trap. Pessimism and warm glow factor interact to create poverty trap. In that paper, if the degree of pessimism is quite low or the warm glow factor is sufficiently large, only then can the economy escape the poverty trap.

Two interesting long-run implications of behavioral anomaly are as follows. First, there could be multiple equilibria, and therefore multiple paths to the steady state. This is because the beliefs about the return to investment are not aligned among various types of parents. Second, a mixed strategy cannot be played at any steady state. Due to the presence of extremely pessimistic uneducated workers, mixed strategy would decrease the mass of educated workers and would change the skilled incomes over time. Thus, at steady state, workers do not randomize their investment decisions and hence an individual's dynasty's education status does not change in the long run. The steady state has intergenerational education immobility, only job mobility among the educated workers.

5 Discussion

For a given parental warm glow, a parent's investment decision is contingent on three factors – (a) utility cost of investment which is independent of biases and depends on parental income, (b) the conjectured probability with which an educated child from her own group gets a skilled job, and (c) the conjectured income of a skilled worker at the subsequent period. Unbiased parents differ only in terms of (a) but have the same (b), and (c). Thus, due to their higher incomes, unbiased skilled workers always invest with higher probability than under-confident unskilled workers.

Biases affect (b) and (c). Uneducated workers are extremely pessimistic (i.e. (b) is zero), and hence, never invest. Educated-unskilled parents are under confident about (b) and overestimate (c). The opposite is true for skilled parents. When warm glow is not low, we classify three key implications of behavioral anomalies on equilibrium investment.

The spillover effect captures how investment incentives expand for unbiased-educated workers due to extreme pessimism of uneducated workers. Pessimism tends to lower the mass of educated children, and hence, increases skilled incomes in the future. This creates opportunities for unbiased-educated workers. Thus, behavioral anomalies result in higher investments by one group due to missing investments by the pessimistic group.[20]

The personal repercussions of over and under confidence on the educated workers is the deviation in investment decisions of the educated workers due to their own biases. While it is expected that overconfident skilled may over invest while the under-confident educated-unskilled worker may under invest, more interestingly, the opposite may also happen. Further, under-confident educated-unskilled worker, in spite of their lower incomes, may invest more than the overconfident rich (skilled) parents. This occurs when biased parents' estimation of factor (c) supersedes (b) and (a). Thus, biases may interact with socio-economic conditions (i.e. the warm glow parameter and the initial skilled income) to generate greater investments by the poor.

Peer effects emerge where over investment (or under investment) of one type of educated parent may crowd-out (or crowd-in respectively) investment by another type of educated parents. To abstract the peer effects from own confidence effects, we consider the situation when there are multiple equilibria. Consider two such equilibria. The equilibrium at which skilled workers invest with a lower probability is also where the educated-unskilled workers invest with a strictly higher probability relative to the other equilibrium. Thus, at the same degree of bias and same initial conditions, higher investments by educated-unskilled would lower the investments of the skilled, and vice versa. The equilibrium strategy depends on what the educated peers do, hence the term.

[20]Obviously, extreme pessimism affects the society only where, to begin with, unbiased-uneducated workers invested in education. Also, extreme pessimism can not increase the probability of investment for skilled as they were already investing with probability one.

	No one is Biased (Benchmark)	Extremely Pessimist Uneducated + Unbiased Educated	Extremely Pessimist Uneducated + Biased Educated
Moderately High Warm Glow $\delta \in (\underline{\delta},\ \delta_a)$	• Always unique equilibrium, • Poor invest with a lower probability than the rich.	• Spillover-effect: Non-investment of pessimists may induce higher investment by unbiased, • Overall investment cannot be higher than the benchmark.	• Possibly multiple equilibria, • Own bias may hinder or induce investment. • Peer Effect: Over or underinvestment of one type may crowd-out or crowd-in other type's investment, • There may be overall over- or underinvestment, as well as by educated-poor and educated-rich.
Huge Warm Glow $\delta \geq \delta_a$	• Always unique equilibrium, • All workers invest with probability 1.	• No spillover-effect as all educated continue to invest with probability 1, • Overall investment cannot be higher than the benchmark.	• Unique equilibria, • Own bias may hinder investments of skilled. • There may be overall underinvestment.

Figure 7: Implications of Behavioral Biases on Equilibria

While we have summarized the cross-sectional effects of behavioral biases, they also have long lasting effects on the steady states as well as the dynamic paths to the steady state. *First, a poverty trap* emerges where a mass of families (here, the uneducated) never invest and hence, never become rich, i.e. skilled workers. Without any biases, there is no poverty trap when the parental warm glow is not low. With extreme pessimism, there is always a poverty trap. More interestingly, in Dasgupta and Saha (2022), we show that even when only the uneducated are slightly biased there could be a poverty trap. *Second, inequality* may increase or decrease due to behavioral anomalies. Here, the strength of pessimism and optimism plays a role. It may be possible for the economy to have more skilled workers, and hence, less income inequality relative to the benchmark levels. *Third, multiple steady-states and multiple paths* are possible as biased parents do not have a consensus on the returns to investment. Thus, behavioral anomalies explain why two economies with similar socio-economic conditions may traverse different trajectories of economic development.

We formally state the steady-state comparisons and their proofs in the Supplementary Appendix.

No one is Biased (Benchmark)	Extremely Pessimist Uneducated + Unbiased Educated	Extremely Pessimist Uneducated Biased Educated
• Unique steady state, • At steady state, poor may randomize.	• Unique steady state, • At steady state, all educated invest with probability 1, • Steady-state inequality is weakly higher than the benchmark.	• Multiple steady states and multiple paths, • At steady state, all educated invest with probability 1, • Steady-state inequality may be lower or higher than the benchmark

Figure 8: Implications of Behavioral Biases on Steady States

6 Conclusion

The paper highlights that a community's personal beliefs and biases can have societal effects. There are some clear effects of behavioral anomalies. Extreme pessimism of uneducated induces higher investment by the educated. Over confidence of skilled workers lowers their own investment when warm glow is huge. We witness a behavioral trap which pushes uneducated into persistent poverty. Biases create multiple steady states and engender history dependence. There are some other ambiguous effects of behavioral anomalies which depend on existing socio-economic conditions. Own confidence may increase or decrease educational investments depending on the initial skilled income and warm glow parameter. Further, over investment of one type of parent may crowd out investment of other parent types. Finally, steady-state income inequality may be higher or lower relative to the benchmark.

Thus, biases bring complexity in any aggregate economy. The cross-sectional and inter-temporal effects of beliefs may explain economic disparities. Interestingly, in our model, a parent not only makes educational investments for their child, but these investments also partially contribute to the child's education-job network, and hence, the child's biases. Intergenerational beliefs are not identical, but inherited. Future work could incorporate more sophisticated dynamics for transmission of beliefs across generations.

As we expand our understanding of the behavioral sciences, we see that the point of intervention is not at the national level but should be decentralized. Our paper aligns with this finding. It aims to direct the focus of public policies towards behavioral impediments. Behavioral policies such as mentoring programs, improving social interaction, among others have to become a part of a policy maker's toolbox, and urgently so.

7 Appendix

In this Appendix, we provide the proofs of the Propositions. The detailed proofs of lemmas and observations are in the Supplementary Appendix.

7.1 Proof of Proposition 1

To prove this proposition, we use the fact that at any equilibrium if an unskilled worker invests with a positive probability, all skilled workers invest with certainty; i.e. $\lambda_{st} = 1$ whenever $\lambda_{ut} > 0$. We provide the formal statement and proof in the Supplementary Appendix.
The uniqueness follows directly from the fact that at any equilibrium, if $\lambda_{ut} > 0$ then $\lambda_{st} = 1$.

1. Given that at any equilibrium $\lambda_{st} = 1$ whenever $\lambda_{ut} > 0$, it is sufficient to show that for any $\delta > \bar{\delta}$, $\lambda_{ut} = 1$ even when $\lambda_{st} = 1$. From (1) it can be seen that $\lambda_{ut} = 1$ is the strictly dominating strategy for the unskilled workers. When $\delta = \bar{\delta}$, similarly, it can be seen that $\lambda_{ut} = 1$ is the weakly dominating strategy for an unskilled worker. Now, observe again from (1), if a positive mass of unskilled worker plays any strategy other than $\lambda_{ut} = 1$, then such an unskilled worker has an incentive to deviate and play $\lambda_{ut} = 1$. Therefore, $\langle 1, 1\rangle$ is a unique equilibrium $\forall\ \delta \geq \bar{\delta}$.

2. Consider $\delta \in [\underline{\delta}, \bar{\delta})$. As discussed above, in any equilibrium if $\lambda_{ut} > 0$ then $\lambda_{st} = 1$. Now, from (1), it can be 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. Similarly, it can be shown that for any $m_{st} \geq 1$, $\langle 0, 0\rangle$ cannot be an equilibrium.

   Further, from the definitions of $\underline{b}_u(\delta)$, $\bar{b}_s(\delta)$ we get the equilibrium strategies:

   - For $m_{st} > \underline{b}_u(\delta)$: unskilled workers invest with a probability such that (1) binds and skilled workers invest with probability 1.
   - For $m_{st} \in [\bar{b}_s(\delta), \underline{b}_u(\delta)]$: unskilled workers do not invest and skilled workers invest with probability 1.
   - For $m_{st} \in [1, \bar{b}_s(\delta))$: unskilled workers do not invest and skilled workers invest with a probability such that (1) binds.

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

   The skilled workers' equilibrium probability follows from the definitions of $\bar{b}_s(\delta)$ and $\underline{b}_s(\delta)$:

   - For $m_{st} \geq \bar{b}_s(\delta)$: skilled workers invest with probability 1. Hence, $\langle 0, 1\rangle$ is the equilibrium.

- For $m_{st} \in (\underline{b}_s(\delta), \bar{b}_s(\delta))$: skilled workers invest with a fractional probability λ_{st} such that (1) binds. Hence, $\langle 0, \lambda_{st} \rangle$ is the equilibrium.
- For $m_{st} \in [1, \underline{b}_s(\delta))$: no worker invests. Hence, $\langle 0, 0 \rangle$ is the equilibrium. $\square$

7.2 Proof of Proposition 2

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

At the steady state, the probability with which a family becomes rich, i.e. the probability with which an adult from that family works as a skilled worker is

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

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

1.b. When warm glow is moderate then at any equilibrium, the unskilled workers invest with a probability less than 1. Further, $\langle 0, 1 \rangle$ or $\langle 0, (0, 1) \rangle$ cannot be an equilibrium strategy *at any steady state* because in such a case, the mass of skilled workers decreases over time. And finally, $\langle 0, 0 \rangle$ cannot be an equilibrium, as the workers have incentives to deviate. So, the steady state equilibrium would be $\langle \lambda_u^*, 1 \rangle$, where λ_u^* is such that the mass of skilled workers remains constant over time. Let that be L_s^*. We now find out the value of λ_u^*.

The unskilled workers randomize, so corresponding (1) binds which can be written as[21]

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

Here, we use the formal expression for $\underline{b}_u(\delta)$

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

From (A.1) and (S.3), we find $L_s^* = \beta \left(\frac{\underline{b}_u(\delta)}{A} \right)^{-\frac{1}{1-\phi}}$.

And, the income of a skilled worker, at the steady state, would be $A{L_s^*}^{-(1-\phi)} = \beta^{-(1-\phi)} \underline{b}_u(\delta)$.

[21] Observe equation (1) can be written as $\delta \left[\frac{\left[\beta A L_{st+1}^{-(1-\phi)} + 1 - \beta\right]^\sigma}{\sigma} - \frac{1}{\sigma} \right] \geq \frac{1}{\sigma} - \frac{(1-\bar{s})^\sigma}{\sigma}$

2. From Proposition 1 (subpoint 3.), we have that when $\delta < \underline{\delta}$, then no unskilled workers invest at any m_{st}. Moreover, only when $m_{st} > \underline{b}_s(\delta)$, then skilled workers invest with a positive probability. So, the mass of educated workers and hence, the mass of skilled workers decrease over time and converge to zero, whereas the income of a skilled worker increases over time and tends to infinity.

   In addition, for $m_{st} \leq \underline{b}_s(\delta)$, no parents invest. So, the economy is in a steady state where no parent invests and all workers are unskilled. $\square$

7.3 Proof of Proposition 3

First, we introduce a Lemma which would be used to prove Proposition 3.

Lemma 7.1. *Suppose $\delta \in (\underline{\delta}, \delta_a]$ and $\beta < \theta(1-\beta)$. Let $\langle \gamma_{ut}, \gamma_{st} \rangle$ be an equilibrium at any $m_{st} \in [1, min\{\bar{a}_u(\delta), \bar{a}_s(\delta)\})$. If $\gamma_{st} \geq \gamma_{ut}$, then at all $\tilde{m}_{st} \in (m_{st}, max\{\bar{a}_u(\delta), \bar{a}_s(\delta)\})$ $\tilde{\gamma}_{st} > \tilde{\gamma}_{ut}$ where $\langle \tilde{\gamma}_{ut}, \tilde{\gamma}_{st} \rangle$ is an equilibrium at $\tilde{m}_{st}$.*

This Lemma is proved in the Supplementary Appendix. Now, we sequentially prove Proposition 3.

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

2. For $\delta > \delta_a$, we get that $\bar{a}_u(\delta) < 1$. So, by the definition of $\bar{a}_u(\delta)$, $\gamma_{ut} = 1 \; \forall \; m_{st} \geq 1$. Thus, from Definition 3, γ_{st} must be equal to $\underline{\gamma}_s(m_{st}) \; \forall \; m_{st} \geq 1$.

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

3 • If $\bar{a}_u(\delta) \leq \bar{a}_s(\delta)$, by the definition of $\bar{a}_u(\delta)$, for $m_{st} > \bar{a}_u(\delta)$ we have $\gamma_{ut} = 1 \; \forall \; \gamma_{st} \in [0, 1]$. Hence, due to Definition 3, γ_{st} must be equal to $\underline{\gamma}_s(m_{st})$ and that the equilibrium is unique is now immediate.

   • Similarly, if $\bar{a}_u(\delta) > \bar{a}_s(\delta)$, by the definition of $\bar{a}_s(\delta)$, for $m_{st} > \bar{a}_s(\delta)$ we must have $\gamma_{st} = 1$. Hence, due to Definition 3, γ_{ut} must be equal to $\underline{\gamma}_u(m_{st})$ and that the equilibrium is unique is now immediate.

   • The probabilities of investments bounded below and above can be directly observed from Definition 3.

   (a) Now, we show that if $\beta \geq \theta(1-\beta)$ then there could be multiple equilibria.

   Suppose at any $m_{st} \geq 1$, there are two equilibria $\langle \gamma_{ut}, \gamma_{st} \rangle$ and $\langle \tilde{\gamma}_{ut}, \tilde{\gamma}_{st} \rangle$. Then, from investment decisions of both types of parents given by (2) and (3), we get the following observation: 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 $\tilde{\gamma}_{kt} = 1$.

Using this along with the definitions of $\bar{a}_u(\delta)$, $\bar{a}_s(\delta)$, we get that at any $m_{st} \in [1, \min\{\bar{a}_u(\delta), \bar{a}_s(\delta)\})$, if there are two equilibria $\langle \gamma_{ut}, \gamma_{st} \rangle$ and $\langle \tilde{\gamma}_{ut}, \tilde{\gamma}_{st} \rangle$, then we must have

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

where at least one of the educated parents does not invest with certainty (as $m_{st} < \min\{\bar{a}_u(\delta), \bar{a}_s(\delta)\}$), and skilled workers must invest with a positive probability (as at least one worker must invest with positive probability for moderately high warm glow). Using (2) and (3), we get

$$\begin{aligned} & \theta(1-\beta)\tilde{\gamma}_{ut} + \beta\tilde{\gamma}_{st} \geq \theta(1-\beta)\gamma_{ut} + \beta\gamma_{st} \\ \Rightarrow \quad & \beta[\tilde{\gamma}_{st} - \gamma_{st}] \geq \theta(1-\beta)[\gamma_{ut} - \tilde{\gamma}_{ut}] \end{aligned} \tag{A.3}$$

$$\begin{aligned} \text{and, } & (1-\beta)\gamma_{ut} + [1-\theta(1-\beta)]\gamma_{st} \geq (1-\beta)\tilde{\gamma}_{ut} + [1-\theta(1-\beta)]\tilde{\gamma}_{st} \\ \Rightarrow \quad & (1-\beta)[\gamma_{ut} - \tilde{\gamma}_{ut}] \geq [1-\theta(1-\beta)][\tilde{\gamma}_{st} - \gamma_{st}]. \end{aligned} \tag{A.4}$$

Both conditions (A.3) and (A.4) hold, i.e. the necessary condition for the coexistence of $\langle \gamma_{ut}, \gamma_{st} \rangle$ and $\langle \tilde{\gamma}_{ut}, \tilde{\gamma}_{st} \rangle$ is $\beta \geq \theta[1-\theta(1-\beta)] \quad \Rightarrow \quad \beta \geq \theta(1-\beta)$.

(b) We now show if $\beta < \theta(1-\beta)$ and $\bar{a}_u < \bar{a}_s$, then $\gamma_{st} < \gamma_{ut}$. Suppose not and there exists $m_{st} \in [1, \min\{\bar{a}_u(\delta), \bar{a}_s(\delta)\})$ at which $\gamma_{st} \geq \gamma_{ut}$. It follows from Lemma 7.1 that at all $\tilde{m}_{st} \in (m_{st}, \bar{a}_s(\delta))$ we will have $\tilde{\gamma}_{st} > \tilde{\gamma}_{ut}$.

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

7.4 Proof of Proposition 4

1. Suppose an economy where warm glow parameter is not low starts with all educated adults, then $m_{st} = A\beta^{-(1-\phi)}$. So, if $A\beta^{-(1-\phi)} \geq \max\{\bar{a}_u(\delta), \bar{a}_s(\delta)\}$, then all parents would invest at all t and there would be no poverty trap.

   Suppose an economy where warm glow parameter is not low, but $A\beta^{-(1-\phi)} < \max\{\bar{a}_u(\delta), \bar{a}_s(\delta)\}$ 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, hence they would be in a poverty trap.

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

2. and 3. Follows directly from Proposition 3, so we skip the details. □

References

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

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.

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

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.

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

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

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

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.

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.

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.

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

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

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

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

Dasgupta, D. and Saha, A. (2022). Towards a theory of behavioral poverty traps. *Available at SSRN 4086552*.

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.

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.

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

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

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

Goel, D. and Deshpande, A. (2020). Social identity and perceived income adequacy. *Review of Development Economics*, 24(2):339–361.

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.

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

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

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

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.

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

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

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

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.

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.

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.

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.

SA Supplementary Appendix

SA.1 Proofs of claims made in Section 2

SA.1.1 Income Comparison of Skilled and Unskilled Workers

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

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

SA.2 Proofs of claims made in Section 3

SA.2.1 Ranking of Warm Glow Parameters

Observation 2. $0 < \underline{\delta} < \bar{\delta}$.

Proof. 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.2.2 Properties of the benchmark equilibrium, $\langle \lambda_{ut}, \lambda_{st} \rangle$.

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

1. *if an unskilled worker invests in her child's education with a positive probability ($\lambda_{ut} > 0$), then a skilled worker invests in her child's education with certainty ($\lambda_{st} = 1$),*
2. *at any period t, the probabilities of investment of both types of workers (weakly) increase with an increase in income of a skilled worker at that period.*

Proof. We spell the steps in detail.

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

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

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

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

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

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

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

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

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

2. Recall (1), a worker of type j, where $j \in \{u, s\}$, invests with probability λ_{jt} if

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

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

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

□

SA.2.3 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

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

$$\Rightarrow \quad \underline{b}_s(\delta): \quad \underline{b}_s^{\sigma} - (\underline{b}_s - \bar{s})^{\sigma} + \delta = 0. \tag{S.1}$$

- 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

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

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

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

SA.2.4 Properties of income thresholds

Lemma 2. *The thresholds of the state variable are such that,*

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

Proof. We derive the claims sequentially.

1. The ranking is immediate from comparing definitions, (S.1), (S.2), and (S.3).
   Differentiating these equations with respect to δ, we get that the thresholds are decreasing in δ.

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

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

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

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

   (ii) When $\delta \in [0, \underline{\delta})$, that $\underline{b}_u(\delta) = \infty$, follows directly from (S.3). $1 < \underline{b}_s(\delta)$ has already been shown above. Hence, proved.

□

SA.3 Proofs of claims made in Section 4

SA.3.1 Properties of the behavioral anomalies equilibrium, $\langle \gamma_{ut}, \gamma_{st} \rangle$.

Lemma 3. *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 $\gamma_{st} > 0$.*

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

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

$$\begin{aligned}
\delta\Big[1 - \big[\theta\beta[\theta(1-\beta)\gamma_{ut}]^{-(1-\phi)} m_{st} + 1 - \theta\beta\big]^{\sigma}\Big] &\geq (1-\bar{s})^{\sigma} - 1 > (m_{st} - \bar{s})^{\sigma} - m_{st}^{\sigma} \quad (\text{since } m_{st} > 1) \\
&> \delta\Big[1 - \big[[1-\theta(1-\beta)][(1-\beta)\gamma_{ut}]^{-(1-\phi)} m_{st} + \theta(1-\beta)\big]^{\sigma}\Big] \quad (\text{since } \gamma_{st} = 0) \\
\Rightarrow \quad (1-\theta) &\geq [(1-\beta)\gamma_{ut}]^{-(1-\phi)} \big[[1-\theta(1-\beta)] - \theta^{\phi}\beta\big] m_{st}. \qquad \text{(S.4)}
\end{aligned}$$

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

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

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

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

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

Hence, $[(1-\beta)\gamma_{ut}]^{-(1-\phi)} \big[[1 - \theta(1-\beta) - \theta^{\phi}\beta\big] m_{st} > (1-\beta)^{-(1-\phi)} \big[[1 - \theta(1-\beta) - \theta^{\phi}\beta\big] > 1 - \theta$ which contradicts (S.4). ☐

Observation 3. *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}.*

Proof. The above observation follows directly from the optimal investment decisions of educated-unskilled workers and skilled workers as stated in equations (2) and (3) respectively. ☐

SA.3.2 Additional Warm Glow Parameter

Observation 4. $0 < \underline{\delta} < \delta_a$.

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

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

SA.3.3 Formal Expressions for Definition 5

- At $\bar{a}_s(\delta)$ a skilled worker is indifferent between investing and not investing, when all other educated workers invest with certainty. So, from (3)

$$\begin{aligned} \bar{a}_s(\delta): \quad & \frac{\delta}{\sigma}\Big[\big[[1-\theta(1-\beta)][1+(1-\theta)(1-\beta)]^{-(1-\phi)}\bar{a}_s + \theta(1-\beta)\big]^{\sigma} - 1\Big] \\ = \quad & \frac{\bar{a}_s^{\sigma} - (\bar{a}_s - \bar{s})^{\sigma}}{\sigma}. \end{aligned} \tag{S.5}$$

- At $\bar{a}_u(\delta)$ an educated-unskilled worker is indifferent between investing and not investing, when all other educated workers invest with certainty. So, from (2)

$$\bar{a}_u(\delta): \quad \frac{\delta}{\sigma}\Big[\big[\theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}\bar{a}_u + 1 - \theta\beta\big]^{\sigma} - 1\Big] = \frac{1-(1-\bar{s})^{\sigma}}{\sigma}. \tag{S.6}$$

Further, there are other thresholds of state variable m_{st}, at which other equilibrium strategies exist.

- Given Lemma 3, when $\gamma_{st} = 0$, γ_{ut} is also zero. So, from (3) we have

$$\underline{b}_s(\delta): \quad \underline{b}_s^{\sigma} - (\underline{b}_s - \bar{s})^{\sigma} + \delta = 0. \tag{S.7}$$

  For a given degree of warm glow δ, at $\underline{b}_s(\delta)$ a skilled worker is indifferent between investing and not investing, when no other worker invests.

- For a given degree of warm glow δ, at $a'_s(\delta)$ a skilled worker is indifferent between investing and not investing, when all other skilled workers invest with certainty and no educated-unskilled worker invests. So, from (3)

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

- At $\tilde{a}_u(\delta)$ an educated-unskilled worker is indifferent between investing and not investing, when all skilled workers invest and no other educated-unskilled worker invests. From (2)

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

- Here at $a'_u(\delta)$ an educated-unskilled worker is indifferent between investing and not investing, when all other educated-unskilled workers invest with certainty and no skilled worker invests. So, from (2)

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

SA.3.4 Properties of the income thresholds

Lemma SA.1. 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) $1 < \bar{a}_s(\delta)$, (ii) $\underline{a}_s(\delta) < \tilde{a}_s(\delta) < \bar{a}_s(\delta)$, (iii) $\underline{a}_s(\delta) < a'_s(\delta) < \bar{a}_s(\delta)$, and (iv) if and only if $\theta(1-\beta) > \beta$, $\tilde{a}_s(\delta) > a'_s(\delta)$.*
3. *The thresholds related to the educated-unskilled workers' decisions are such that:*
   a. *If and only if $\delta > \underline{\delta}$, $\quad \bar{a}_u(\delta)$, $a'_u(\delta)$ and $\tilde{a}_u(\delta)$ are finite.*
   b. *If and only if $\delta < \delta_a$, $\quad 1 < \bar{a}_u(\delta)$.*
   c. *$\forall\ \delta > \underline{\delta}$, $\quad a'_u(\delta) > \tilde{a}_u(\delta)$ if and only if $\theta(1-\beta) > \beta$, and $\max\{\tilde{a}_u(\delta), a'_u(\delta)\} < \bar{a}_u(\delta)$.*
4. *$\forall\ \delta \leq \delta_a$, we have $\tilde{a}_s(\delta) \leq \bar{a}_u(\delta)$ and $\forall\ \delta > \delta_a$, we have $\tilde{a}_s(\delta) < 1$.*
5. *Cut-offs relative to the benchmark case: (i) $\underline{b}_u(\delta) < \tilde{a}_u(\delta)$.*

Proof of Lemma SA.1

1. We want to show that all the thresholds are decreasing in δ. Consider the income-cutoff $a'_s(\delta)$, which is determined by (S.8). The L.H.S. of (S.8) is increasing in a'_s and δ and the R.H.S is decreasing in a'_s. Hence, the claim. Following similar argument, this negative relationship can be shown for the other thresholds as well.

2. (i) Suppose not and $\bar{a}_s(\delta) \leq 1$. Since $y(x)$ is decreasing in x as in Section SA.2.2 we have

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

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

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

The following steps give us that

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

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

We show $\bar{a}_s(\delta)>\tilde{a}_s(\delta)$. Suppose not and $\bar{a}_s(\delta)\le\tilde{a}_s(\delta)$. Hence, from Lemma 1 it follows that the utility cost of investment at $\bar{a}_s(\delta)$ is no less than that at $\tilde{a}_s(\delta)$:

$$\frac{\tilde{a}_s(\delta)^{\sigma}-(\tilde{a}_s(\delta)-\bar{s})^{\sigma}}{\sigma}\le\frac{\bar{a}_s(\delta)^{\sigma}-(\bar{a}_s(\delta)-\bar{s})^{\sigma}}{\sigma}\quad\Rightarrow(\tilde{a}_s(\delta)-\bar{s})^{\sigma}-\tilde{a}_s(\delta)^{\sigma}\le(\bar{a}_s(\delta)-\bar{s})^{\sigma}-\bar{a}_s(\delta)^{\sigma}.$$

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

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

which contradicts the above.

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

(iv) We show by contradiction that when $\theta(1-\beta)>\beta$ then $\tilde{a}_s(\delta)>a'_s(\delta)$. The converse can be shown similarly which we skip.

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

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

or the benefit at the primitive of $\tilde{a}_s(\delta)$ is strictly lower than that of $a'_s(\delta)$. Hence, the statements of both the definitions of $\tilde{a}_s(\delta)$ and $a'_s(\delta)$ cannot simultaneously be true.

3. a. Let us write the threshold $\bar{a}_u(\delta)$ expression (S.6) as:

$$\left[\theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}\bar{a}_u + 1 - \theta\beta\right]^{\sigma} = \frac{1-(1-\bar{s})^{\sigma}+\delta}{\delta}.$$

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

Similarly using equations (S.9) and (S.10) the same can be shown for $\tilde{a}_u(\delta)$ and $a'_u(\delta)$.

b. Suppose $\bar{a}_u(\delta) \gtreqless 1$. Using this in (S.6) we get

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

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

<u>First we show $\tilde{a}_u(\delta) < \bar{a}_u(\delta)$</u> Comparing (S.6) and (S.9) we get,

$$\frac{\delta}{\sigma}\Big[\left[\theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}\bar{a}_u(\delta)+1-\theta\beta\right]^{\sigma}-1\Big] = \frac{\delta}{\sigma}\left[\left[\theta\beta^{\phi}\tilde{a}_u(\delta)+1-\theta\beta\right]^{\sigma}-1\right]$$
$$\Rightarrow \quad [\theta(1-\beta)+\beta]^{-(1-\phi)}\bar{a}_u(\delta) = \beta^{-(1-\phi)}\tilde{a}_u(\delta) \qquad \Rightarrow \quad \bar{a}_u(\delta) > \tilde{a}_u(\delta) \quad \text{as } \theta(1-\beta)+\beta > \beta$$

<u>Now, we show $a'_u(\delta) < \bar{a}_u(\delta)$</u> Comparing (S.6) and (S.10), we get

$$\frac{\delta}{\sigma}\Big[\left[\theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}\bar{a}_u(\delta)+1-\theta\beta\right]^{\sigma}-1\Big] = \frac{\delta}{\sigma}\left[\left[\theta\beta[\theta(1-\beta)]^{-(1-\phi)}a'_u(\delta)+1-\theta\beta\right]^{\sigma}-1\right]$$
$$\Rightarrow \quad \theta\beta[\theta(1-\beta)+\beta]^{-(1-\phi)}\bar{a}_u(\delta) = \theta\beta[\theta(1-\beta)]^{-(1-\phi)}a'_u(\delta)$$
$$\Rightarrow \quad \bar{a}_u(\delta) > a'_u(\delta) \quad \text{as } \theta(1-\beta)+\beta > \theta(1-\beta)$$

$\underline{a'_u(\delta) > \tilde{a}_u(\delta) \text{ if and only if } \theta(1-\beta) > \beta}$ Comparing (S.10) and (S.9) we get,

$$\frac{\delta}{\sigma}\left[\left[\theta\beta[\theta(1-\beta)]^{-(1-\phi)}a'_u(\delta) + 1 - \theta\beta\right]^{\sigma} - 1\right] = \frac{\delta}{\sigma}\left[\left[\theta\beta^{\phi}\tilde{a}_u(\delta) + 1 - \theta\beta\right]^{\sigma} - 1\right]$$
$$\Rightarrow \quad \theta\beta[\theta(1-\beta)]^{-(1-\phi)}a'_u(\delta) = \theta\beta^{\phi}\tilde{a}_u(\delta)$$

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

4. $\underline{\text{First, we show } \forall\ \delta \leq \delta_a,\ \tilde{a}_s(\delta) \leq \bar{a}_u(\delta).}$ Suppose not. $\exists\ \delta \leq \delta_a$ such that $\bar{a}_u(\delta) < \tilde{a}_s(\delta)$. Consider any $m_{st} \in [\bar{a}_u(\delta), \tilde{a}_s(\delta)]$. Since, $m_{st} \geq \bar{a}_u(\delta)$, $\gamma_{ut} = 1\ \forall\ \gamma_{st} = [0, 1]$. From the definition of $\tilde{a}_s(\delta)$, γ_{st} must be equal to zero for $m_{st} < \tilde{a}_s(\delta)$, which contradicts Lemma 3.

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

5. Comparing definitions (S.1) and (S.7) we get $\underline{b}_s(\delta) = \underline{a}_s(\delta)$.

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

$$\delta\left[\frac{[\theta\beta^{\phi}\underline{b}_u(\delta) + 1 - \theta\beta]^{\sigma}}{\sigma} - \frac{1}{\sigma}\right] < \delta\left[\frac{[\beta^{\phi}\underline{b}_u(\delta) + 1 - \beta]^{\sigma}}{\sigma} - \frac{1}{\sigma}\right]. \tag{S.11}$$

   As this implies

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

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

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

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

Observe if $\underline{b}_u = \beta^{1-\phi}$, the benefit from investment is zero and cost is positive, but $\underline{b}_u(\delta)$ should be such that the benefit is equal to cost. The benefit is increasing in m_{st}, hence, $\underline{b}_u(\delta)$ must be greater than $\beta^{1-\phi}$. Hence, the inequality S.11 is true $\forall\theta < 1$. $\square$

SA.3.5 Formal Definition of Boundary conditions

The boundary values are such that

$$\underline{\gamma}_s(m_{st}) \begin{cases} = 0 & \forall\ m_{st} \le \tilde{a}_s(\delta), \\ \in (0,1) & \forall\ m_{st} \in (\tilde{a}_s(\delta), \bar{a}_s(\delta)), \\ = 1 & \forall\ m_{st} \ge \bar{a}_s(\delta), \end{cases} \text{ and } \bar{\gamma}_s(m_{st}) \begin{cases} = 0 & \forall\ m_{st} \le \underline{a}_s(\delta), \\ \in (0,1) & \forall\ m_{st} \in (\underline{a}_s(\delta), a'_s(\delta)), \\ = 1 & \forall\ m_{st} \ge a'_s(\delta). \end{cases}$$

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

$$\underline{\gamma}_u(m_{st}) \begin{cases} = 0 & \forall\ m_{st} \le \tilde{a}_u(\delta), \\ \in (0,1) & \forall\ m_{st} \in (\tilde{a}_u(\delta), \bar{a}_u(\delta)), \\ = 1 & \forall\ m_{st} \ge \bar{a}_u(\delta), \end{cases} \text{ and } \bar{\gamma}_u(m_{st}) \begin{cases} = 1 & \forall\ m_{st} \ge a'_u(\delta), \\ < 1 & \text{at } m_{st} < a'_u(\delta). \end{cases}$$

We derive the conditions for $\underline{\gamma}_s(m_{st})$ and the other bounds are derived similarly. As stated in Definition 3, $\underline{\gamma}_s(m_{st})$ captures the optimal response of a skilled worker when $\gamma_{ut} = 1$. Thus, for $m_{st} \ge \tilde{a}_s(\delta)$, the formal expression for $\underline{\gamma}_s(m_{st})$ is

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

the inequality binds only when $m_{st} \in [\tilde{a}_s(\delta), \bar{a}_s(\delta)]$. Clearly from the definition of $\tilde{a}_s(\delta)$ and Observation 3 we can find that $\underline{\gamma}_s(m_{st}) = 0$ for $m_{st} \le \tilde{a}_s(\delta)$.
Similarly, including the definition of $\bar{a}_s(\delta)$, we get $\underline{\gamma}_s(m_{st})$. As γ_{ut} can atmost be one, now it is clear that at any equilibrium $\langle \gamma_{ut}, \gamma_{st} \rangle$, $\gamma_{st}(m_{st}) \ge \underline{\gamma}_s(m_{st})$.

Finally we show that if $m_{st} \in (\tilde{a}_s(\delta), \bar{a}_s(\delta))$ then $d\underline{\gamma}_s(m_{st})/dm_{st} > 0$.
Suppose not. $\tilde{a}_s(\delta) < m_{st}^1 < m_{st}^2 < \bar{a}_s(\delta)$ and $1 > \underline{\gamma}_s^1 \equiv \underline{\gamma}_s(m_{st}^1) \ge \underline{\gamma}_s(m_{st}^2) \equiv \underline{\gamma}_s^2 > 0$. Then, we

must have

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

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

Following similar arguments as above, it can be shown that at any equilibrium $\langle\gamma_{ut}, \gamma_{st}\rangle$, $\gamma_{st}(m_{st})$ is bounded above by $\bar{\gamma}_{st}(m_{st})$. Also, $\gamma_{ut}(m_{st})$ is bounded below and above by $\underline{\gamma}_u(m_{st})$ and $\bar{\gamma}_u(m_{st})$ respectively. □

SA.3.6 Proof of Lemma 7.1.

Proof. $\delta \in (\underline{\delta}, \delta_a]$. Then from Lemma 2 and Lemma SA.1 (subpoint 5.), we have $\underline{a}_s(\delta) < 1$.

Suppose not. $\gamma_{st} \geq \gamma_{ut}$ and $\exists\ \tilde{m}_{st} > m_{st}$ such that $\tilde{\gamma}_{ut} \geq \tilde{\gamma}_{st}$.
First observe from the previous claim that for $\beta < \theta(1-\beta)$, at any m_{st} there will always be a unique equilibrium $\langle\gamma_{ut}, \gamma_{st}\rangle$. Second, m_{st} less than less than $\min\{\bar{a}_u(\delta), \bar{a}_s(\delta)\}$ and $\tilde{m}_{st}$ less than $\max\{\bar{a}_u(\delta), \bar{a}_s(\delta)\}$ imply $\gamma_{ut} < 1$ and $\tilde{\gamma}_{st} < 1$.
Now from the investment decision of educated-unskilled workers, given by (2), 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 (3), 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

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

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

□

SA.3.7 Figure from Numerical Example

Figure 9 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.

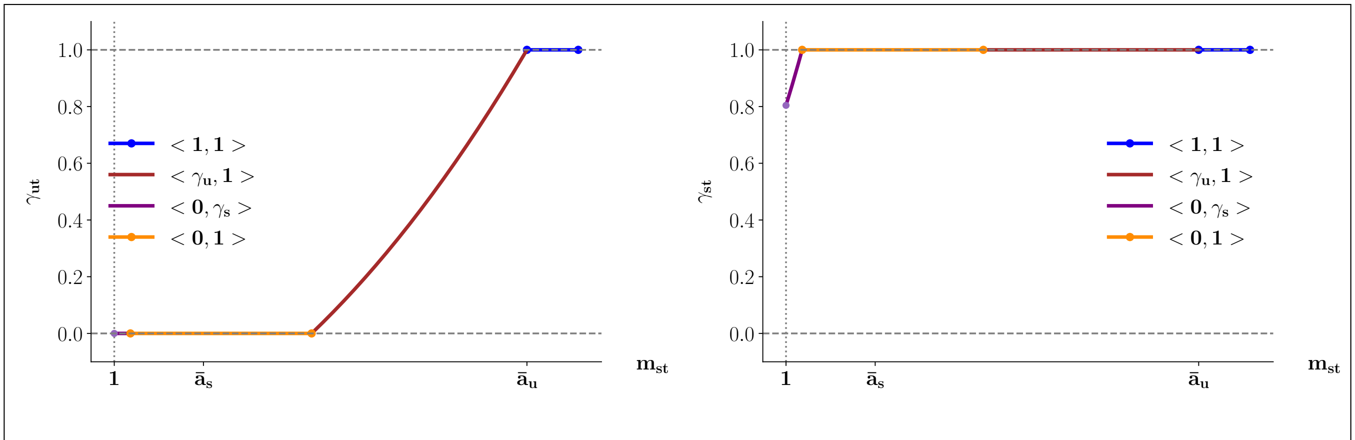


Figure 9: Example to depict equilibrium for $\theta(1-\beta) > \beta$. [Colored Graphs]

SA.4 Proofs of claims made in Section 5

SA.4.1 Comparison between Benchmark and Extremely Pessimism in the Uneducated.

Observation 5. *When the degree of warm glow is not low, the distortions in investment decisions are as follows:*

1. *at any $m_{st} \geq 1$ the skilled workers invest with the same probability in both the cases*

2. *under extreme pessimism, educated-unskilled workers invest with a strictly higher probability than that in the benchmark case when the degree of warm glow 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.*

Proof. We prove the claims sequentially.

1. When warm glow is not low and when educated workers are not biased, skilled workers invest with a positive probability. When the probability of investment is a fraction, it must be that the unskilled workers do not invest. Here, clearly, extreme pessimism in uneducated does not affect the probability of investment of the skilled workers. When the probability of investment is unity, unskilled workers invest with positive probability. Due to extreme pessimism in uneducated, they do not invest but educated-unskilled do. Thus, the behavioral anomaly of uneducated workers incentivises educated-unskilled to invest with a higher probability while the skilled workers continue to invest with certainty.

2. Follows from Propositions 1 and 3.

□

Observation 6. *When the degree of warm glow is not low, steady-state inequality is (weakly) higher under the presence of extremely pessimistic uneducated workers.*

Proof. If there is a positive mass of uneducated workers, due to extreme pessimism they would not invest in the steady state. The mass of parents who invest in the steady state is lower and hence steady-state skilled income, and hence income inequality, is higher. Only in the extreme case when there are no uneducated in an economy and the economy is already at the steady state, we find that the steady-state inequality in the benchmark and the economy with extremely pessimistic uneducated workers is the same. □

SA.4.2 Comparison between Benchmark and the case of all biased agents.

Observation 7. *The distortions in investment decisions are as follows:*

1. *When the degree of warm glow is huge,*
   a. *at any $m_{st} \geq 1$ the educated-unskilled workers with or without behavioral bias invest with the same probability,*
   b. *skilled workers with behavioral bias under invest when the state variable is lower than $\bar{a}_s(\delta)$.*
2. *When the degree of warm glow is high and the state variable is less than $\max\{\bar{a}_u(\delta), \bar{a}_s(\delta)\}$, then both types of educated workers may under invest.*
3. *When warm glow is moderate, then both types of educated workers may over or under invest.*

Proof. This observation follows from Propositions 1 and 3. □

Observation 8. *When all agents are biased:*

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

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

Proof. The aforementioned observation follows from Propositions 2 and 4. □