Erasmus
School of
Economics

Academic year 2021 – 2022

ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Master Thesis International Economics

Discrimination in the workplace: how taste-based discrimination affects a task assignment decision

Name student: Lennert Jacobs
Student ID number: 613267

Supervisor: Jurjen Kamphorst
Second assessor: Otto Swank

Date final version: 09/07/2022

Erasmus University Rotterdam

Abstract

Building on the existing economic literature on confidence management and pioneering theoretical work regarding taste-based discrimination, this study extends an existing model by Kamphorst and Swank (2016) to study how a manager's "taste for discrimination" affects a task assignment decision between two employees. The main building blocks of the model are the following: First, effort and ability are assumed to be complements. Second, the manager has superior and perfect information about the ability levels of the employees. Third, the employees' ability levels are independently drawn from uniform distributions, making them identical ex-ante. In this simplified setting, the model predicts discrimination to be a stable outcome. The first part of the analysis shows how a bias changes the number of equilibria and the severity of discrimination in those equilibria. Whereas Kamphorst and Swank (2016) find that discrimination can be avoided with a high enough level of major task importance, the extended model used here shows that a bias makes it impossible to reach the non-discriminatory equilibrium regardless of the level of major task importance. On the other hand, the most discriminatory equilibria in the model are always stable yet inefficient. The second part of the analysis suggests that a negative bias has the strongest effect in worsening discrimination, while a positive bias is stronger in ameliorating the level of discrimination. Finally, an equal and high positive bias towards the employees simultaneously gives rise to a unique and stable non-discriminatory equilibrium. However, it is highly unlikely for a manager's bias toward two individuals to be identical in reality. Discrimination remains a stable outcome whenever the biases are not identical.

Table of contents

List of figures and tables

1) Introduction

The phenomenon of discrimination in the workplace is widespread and well-documented. For instance, a study by the U.S. Equal Employment Opportunity Commission (EEOC) presents the total amount of charges of employment discrimination together with the type of discrimination for each complaint. *Figure 1* below shows how the total charges of employment discrimination remained largely stable over time, with a rise between 2007 and 2017 followed by a relative drop in more recent years. Regardless, there have been over two million charges of employment discrimination since 1997. Even though the number of cases has decreased in recent years, workplace discrimination is still a major issue that deserves attention. For example, a diversity and inclusion report by Glassdoor (2019) indicates how 61% of U.S. employees indicate to have encountered discrimination based on race, gender, age, or LGBTQ status. They further find this number to be lower in the U.K. (55%), France (43%), and Germany (37%). However, across all four of these countries, 50% of employees believe that their employer should take action to improve diversity and inclusion in the workplace. The question has to be asked: "Why is discrimination still this prevalent regardless of it being well-documented and opposed by employees?".

This study aims to answer that question by extending a model created by Kamphorst and Swank (2016) in which a manager has to assign a major and a minor task to two employees. Unlike their model, the one used in this paper allows the manager to have a "taste for discrimination" in the form of a bias towards either one of the two employees or both of them simultaneously. The model assumes that the manager has perfect and superior information on the employee ability levels, meaning that the employees infer their ability based on the managerial decision made by the manager. This is often referred to as the "looking-glass self" defined by

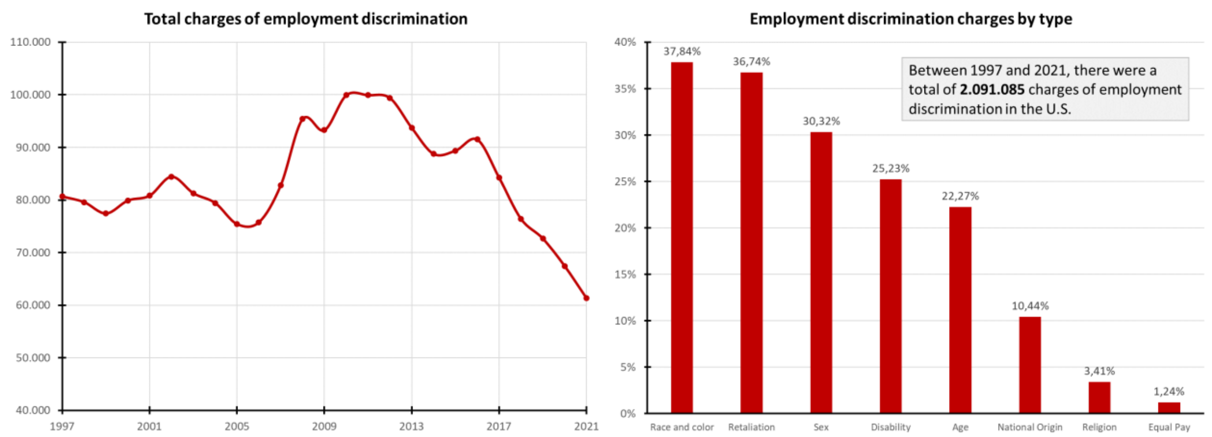


Figure 1: Total charges of employment discrimination together with the types of discrimination for which the charges were filed. Note that the percentages add up to more than 100% because some charges are filed for multiple types simultaneously.

Cooley (1902). Following the existing literature on confidence management, it is assumed that ability and effort are complements. Therefore, the task assignment decision made by the manager will directly impact the level of effort exerted by the employees through the inference process of their ability levels.

The results of the model suggest that, with no pre-existing bias, two stable discriminatory equilibria exist together with one unstable non-discriminatory equilibrium. In other words, even though the employees are identical ex-ante, the model predicts that discrimination is a stable outcome. Kamphorst and Swank (2016) suggest that discrimination is self-fulfilling. Whenever employees expect discrimination, it is optimal for the manager to follow these beliefs to not disproportionately demotivate the employee who expected discrimination in his/her favour. This is because not doing so would then send a signal to the employee that his/her ability level is considerably lower than the other, which would decrease motivation. Additionally, a high enough level of bias causes a unique discriminatory equilibrium to arise in which the manager discriminates heavily towards one of the employees. However, a bias does not entirely prevent an equilibrium from existing in favour of the employee towards whom the manager has a negative bias. The model further finds that a negative bias has the strongest effect in worsening discrimination and that there exists a maximum negative bias after which a manager perfectly discriminates in favour of the opposite employee. On the other hand, a positive bias is found to be the strongest in ameliorating the level of discrimination. In contrast to the work of Kamphorst and Swank (2013), this study finds that whenever the manager has a pre-existing bias towards one of the employees, a high enough major task importance can no longer fully eliminate discrimination. In other words, extending their model and modifying it to be closer to reality sheds light on why discrimination has proven to be persistent and difficult to counter.

The rest of this paper is organised as follows. *Section 2* provides a brief overview of the literature on discrimination and confidence management. It focuses on the most prominent theories regarding discrimination together with empirical evidence and discusses the literature on confidence management in more detail. Afterwards, *Section 3* sets up and explains the three theoretical models used in this paper to explain discrimination and study the effect of the manager's ”taste for discrimination”. Subsequently, the results for the three models are discussed and compared in *Section 4*, after which *Section 5* summarises the main findings and formulates a brief conclusion to the paper. Finally, *Section 6* sheds light on the study's main limitation and proposes directions for future research.

2) Literature review

In this section, insights from the existing economic literature will be discussed regarding discrimination in the workplace and confidence management. First, the work by Kamphorst and Swank (2016) will be considered as it served as the primary influence in setting up and extending the model used in this analysis. Second, several pioneering papers are discussed that formulated the most prominent theories regarding discrimination. In doing so, they generally assume that employers have imperfect information about an employee’s ability. However, the model used in this study will assume that the manager perfectly observes the ability level of the employees while the employees themselves do not. In other words, the manager is assumed to have superior and perfect information. This assumption is in line with a strain of economic literature on confidence management that strongly influenced parts of the model below. Finally, some empirical studies on discrimination in the workplace will be considered to frame the analysis of this paper and underline why it is essential to understand the occurrence of discrimination in the workplace and the challenges in countering it.

The paper titled “*Don’t Demotivate, Discriminate*” by Jurjen Kamphorst and Otto Swank (2016) served as the primary influence on the model used for the analysis in this study. In their research, the authors set up a similar model to the one used below in which a manager assigns a major and a minor task to two employees. They find that discrimination is a stable outcome in a simplified setting. However, they show that high major task importance can partially or entirely counter discrimination, resulting in a unique non-discriminatory equilibrium. Additionally, they study the effect of cheap-talk and how a common prejudice regarding employees' ability reduces discrimination towards one employee while aggravating it against the other. An important assumption in their model is that the manager has “no taste for discrimination”, meaning that the manager maximises his/her utility when total production is maximised. The analysis in this paper loosens that assumption and allows the manager to have a pre-existing bias towards either one of the employees or both of them simultaneously. Therefore, even though this study can be viewed as an extension of Kamphorst and Swank (2016), their model can be interpreted as a “special case” of the one used here.

In the economic literature on discrimination, there exist two opposing views on its occurrence in the workplace. For instance, the pioneering work by Becker (1957) with his book “*The Economics of Discrimination*” introduces the concept of taste-based discrimination. Becker argues that discrimination reflects an employer’s taste against a specific type of worker.

Therefore, these individuals receive lower wages to be accepted as employees. However, Arrow (1971) and Phelps (1972) proposed a different theory for discrimination, arguing that Becker's theory was flawed because discriminating firms would get pushed out of the market due to competition. Therefore, they introduced the theory of statistical discrimination. In other words, they suggested that discrimination arises from the manager having imperfect information about the employees' productivity level and thus basing decisions on statistical information about the average performance of the group to which an individual belongs. Additionally, Coate and Loury (1993) further contribute to this literature by setting up a model in which a manager observes an employee's group identity but, again, not his productivity level (or ability). They argue that employees against whom a manager has a negative bias are assigned to lower rewarding jobs within the firm even though their ex-ante productivity is identical. Therefore, the expected return on investment for these employees will be lower, making discrimination a self-fulfilling prophecy. Furthermore, they study whether a government policy forcing employers to assign employees of different groups to the more highly rewarding jobs at the same rate can eliminate negative prejudices. However, their results are mixed, finding only limited evidence for the effectiveness of affirmative action.

Whereas the theories on statistical discrimination rely on the assumption of a manager having imperfect information, the model used here assumes that a manager has superior and perfect information on the employees' ability levels. Therefore, the manager's "taste for discrimination" is taste-based (Becker, 1957) rather than statistical (Arrow, 1971; Phelps, 1972). The existing body of economic literature on taste-based discrimination primarily establishes theoretical models explaining discrimination in the credit market (Dymski, 1995; Han, 2001; Han, 2004) and housing market (Masson, 1973; Lee and Warren, 1977; Courant, 1978; Cronin, 1982; Yinger, 1975). Other studies provide empirical evidence on taste-based discrimination in the labour market (Carlsson and Rooth, 2012; Busetta et al., 2018) or summarise the main trends in research regarding discrimination (Dymski, 2006). This study contributes to that literature by providing a model applying the notion of taste-based discrimination to the workplace.

On the other hand, the assumption of the manager having perfect and superior information is in line with an existing body of economic literature on confidence management. For instance, Bénabou and Tirole (2003) study the effect of rewards (extrinsic motivation) on the employees' perception of the task and their own ability, affecting their intrinsic motivation. Fundamental in their model is the asymmetry of information between the principal (manager) and the agent (employee). Similarly to the model in this paper, the authors assume that the principal observes the agent's ability while the agent himself has imperfect self-knowledge. Therefore, the agent attempts to infer the principal's private information from a managerial decision. This inference process is also referred to as the "looking-glass self" (Cooley, 1902) and will also play a crucial role in this study. Furthermore, Bénabou and Tirole (2003) assume that effort and ability are complements. Therefore, as offering rewards to employees might negatively influence the inference process of their ability, providing extrinsic motivation can be counterproductive. Other studies use this same line of reasoning to explain how a promotion decision (Ishida, 2006), ordinary talk (Crutzen et al., 2013), the early selection of stars (Prendergast, 1992) or credible communication (Swank and Visser, 2007) can affect an employee's motivation. However, none of the studies above focus on discrimination driven by motivating employees. This study will combine the earlier theories on discrimination and the literature on confidence management by establishing a model in which the manager has a "taste for discrimination" (Becker, 1957) together with perfect and superior information on the employees' ability levels. It will be shown how a "taste for discrimination" combined with the complementarity of ability and effort will induce the manager to discriminate to keep employee morale high.

Next to these theoretical studies regarding discrimination and confidence management, there exists an extensive empirical literature on discrimination in the workplace. For instance, Ozeren (2014) provides a literature review of fifty-two papers concerning discrimination based on sexual orientation in the workplace. In doing so, he defines the major themes on which the current literature focuses and provides the key pillars through which discrimination occurs. Similarly, Ghumman et al. (2013) also set up a literature review but focus on religious discrimination. Other papers consider subtle discrimination against ethnic minorities both in the workplace (Van Laer and Janssens, 2011) and the labour market (Andriessen, 2019), subtle discrimination in general (Jones et al., 2017) or racial discrimination (Vassilopoulou, 2019). Additionally, Wooten and James (2004) use discrimination lawsuits and argue that such lawsuits reflect failures in discrimination management and learning by firms.

The papers above show how discrimination in the workplace is well-documented and widespread. This study does not provide empirical evidence on the occurrence of discrimination. However, it provides a framework for explaining why discrimination occurs in a simplified setting. Moreover, it sheds light on how taste-based discrimination can worsen the equilibria and make it more challenging to counter discrimination in practice.

3) The model

The different models used to uncover the relation between discrimination and demotivation will be set up and explained in this section. These models are heavily inspired by the work of Kamphorst and Swank (2016) and are extensions of their research. Consequently, the model constructed by Kamphorst and Swank (2016), together with their results, can be seen as a special case of the one used in this study. The first framework set up below describes a task assignment game in which the manager has a bias towards one out of two employees. A second model includes a bias towards the opposite employee to determine whether or not the effect mirrors the first model. Thirdly, both biases are included simultaneously to uncover whether they can be taken together and interpreted as a *net bias*.

3.1) Task assignment game with a bias towards employee 1

The model considers a firm with a manager, *M*, and two employees, *i* (with $i \in \{1,2\}$). The manager's goal is to optimally assign both a major and a minor task to the employees. The level of production resulting from both tasks differs according to a *task importance parameter* (η). This parameter is equal to or greater than one ($\eta \geq 1$). In other words, the major task is either equally as important as the minor task (which will be the case in the *basic model*) or more important. In addition to task importance, production is determined by the level of *effort* (e_i) exerted by the employees and their level of *ability* (a_i). The task assignment decision of the manager will be denoted by *m*. If the manager assigns the major task to employee 1, then *m* equals one and vice versa. The above can be combined into the following production function:

$$y_i = \begin{cases} \eta a_i e_i & for \;\; m = i \\ a_i e_i & for \;\; m \neq i \end{cases} \tag{1}$$

Equation (1) above shows that the level of production is increasing in both employee ability and effort. Moreover, if employee *i* is assigned to the major task ($m = i$), the level of production increases with major task importance (η). Important to note is that the model makes a number of simplifying assumptions regarding the employee ability levels. First, the ability levels of both employees are independently drawn from uniform distributions on the interval [0,1]. In other words, the employees are assumed to be identical ex-ante. Second, the model assumes that the manager has superior information on these ability levels. More specifically, the manager observes both a_1 and a_2 while the employees only know their distribution. The reasoning behind this is that, in the setting of a firm, the manager is more aware of what both tasks entail and has greater knowledge about the specific skills needed to perform them successfully. Moreover, the manager has more experience in assigning similar tasks while the employees are assumed to have only limited professional experience. While observing both employees, the manager knows what skills to look for and thus observes the ability levels while the employees themselves do not.

Employee preferences are described by *Equation (2)*. Again, utility differs according to whether the employee is assigned to the major ($m = i$) or the minor ($m \neq i$) task. The first part of the equations shows that employee utility is increasing in the level of production. This implies that utility would be higher when assigned to the major task and that this effect is increasing in the level of major task importance (η). However, as previously mentioned, employees do not observe their ability level. Consequently, in the employee utility function, ability takes the form of an expectation conditional on the manager's task assignment decision. In other words, as the employees only observe the task assignment, their expected level of ability will be based on the task they get assigned to. This inference process is what Cooley (1902) defined as the ”looking-glass self”. The second part of the equations shows that employees are effort averse regardless of the task assignment decision. On the one hand, employees benefit from choosing a higher level of effort by achieving a higher level of production. This is because employees have an intrinsic motivation to contribute to the firm. However, there has to be a limit to how much effort employees can provide before becoming overworked, which would decrease utility.

$$U_i(e_i) = \begin{cases} \eta E(a_i|m)e_i - \frac{1}{2}e_i^2 & if\ m = i \\ E(a_i|m)e_i - \frac{1}{2}e_i^2 & if\ m \neq i \end{cases} \tag{2}$$

Next to employee utility, *Equation (3)* represents the manager utility function. As explained at the beginning of this section, this first model assumes a *pre-existing manager bias* (ρ_1) towards employee 1. The first part of the equation depicts the joint output by the two employees, while the second part represents the manager's bias when employee 1 is assigned to the major task. The model used by Kamphorst and Swank (2016) assumes “no taste for discrimination”, which is why the manager has no pre-existing bias towards either of the employees ($\rho_1 = 0$). This makes their model a special case of the one used in this study. In a model without bias, the manager assigns the tasks to maximise total output. However, next to total production, a manager's utility might be influenced by a positive or negative pre-existing bias towards one of the two employees. The second part of *Equation (3)* shows that, in this model, the manager does have some “taste for discrimination” (Becker, 1957).

$$U_M(a_1, a_2, m) = \sum_{i=1}^{2} y_i + (2 - m)\rho_1 a_1 \tag{3}$$

For instance, if the manager has a pre-existing positive bias towards employee 1 ($\rho_1 > 0$), utility from giving the major task to employee 1 ($m = 1$) will increase proportionally with the severity of the bias (higher ρ_1). A positive bias implies that the manager prefers employee 1 to receive the major task, which is why the manager's utility increases whenever this is the case. However, as the manager observes both a_1 and a_2, the extent to which utility is affected by the bias can be tied to the employee's ability level too. If a positive bias towards employee 1 already exists, the positive effect on the manager's utility will increase more if a high ability level is observed. The rationale for this is that if a manager already prefers one of the two employees, observing that employee has a high level of skill would logically strengthen the manager in believing that employee 1 is well suited to assign the major task to. The manager is strengthened in his/her prejudice that employee 1 deserves to receive the major task, and the manager's utility would increase by more if the major task is assigned to employee 1 with a high level of ability. Note that the pre-existing manager bias is called a "taste for discrimination" because the bias is taste-based (Becker, 1957) rather than statistical (Arrow, 1971; Phelps, 1972). This is because statistical discrimination results from imperfect information whereas the manager in this model has superior and perfect information about the employees' ability levels.

On the other hand, a pre-existing negative bias against employee 1 ($\rho_1 < 0$) would decrease the manager's utility if the major task were to be assigned to employee 1. Again, this effect would only be present if employee 1 gets the major task ($m = 1$) and strengthens with the ability level of the employee. At first sight, a negative bias getting aggravated by a higher level of ability might seem counterintuitive. However, a study by Dietz et al. (2015) shows that recruiters are more biased against immigrant applicants the more skilled they are. The authors refer to this phenomenon as the *skill paradox*. In the task assignment game set up above, a negative manager bias ($\rho_1 < 0$) reflects the skill paradox described in their study, where a higher level of ability further aggravates an existing negative bias towards an employee.

Given the utility function described by *Equation (2)* above, employees choose a level of effort ($e_i > 0$) to maximise their utility. Taking the derivative of the employee utility function with respect to effort results in the following optimal effort levels:

$$e_i = \begin{cases} \eta E(a_i|m) & if\ m = i \\ E(a_i|m) & if\ m \neq i \end{cases} \tag{4}$$

The optimal level of effort exerted by an employee is larger when assigned to the major task. This difference increases the more important the major task becomes compared to the minor one. Additionally, the optimal effort level is a positive function of the expected ability given the manager's task assignment decision, meaning that effort and ability are complements in this model. This is because employees have a higher intrinsic motivation to contribute to the firm by exerting more effort whenever their expectations about their ability level are higher,. In other words, the task assignment decision directly influences the employees' perception of their ability, affecting their optimal effort level. The complementarity of effort and ability is in line with the models used in the economic literature on confidence management (Bénabou and Tirole, 2003; Ishida, 2006; Crutzen et al., 2013; Prendergast, 1992). Basically, through the task assignment decision, the manager influences the self-image of the employees as their perception of their ability is based on the task to which they get assigned. In turn, this will affect their optimal effort levels.

The table below summarizes the different stages of the task assignment game. The manager is the first mover because he/she observes the employee ability levels on which the task assignment decision will be based. Then, given the manager's task assignment decision, employees update their beliefs about their ability level and choose an optimal level of effort. Additionally, the manager assigns the tasks optimally given the employees' ability levels and expected effort levels based on their beliefs. Moreover, beliefs are updated in line with Bayes' rule. Therefore, the game results in perfect Bayesian-Nash equilibria where employees choose their effort levels optimally given their expected ability levels.

The task assignment game
1. Nature draws the employee ability levels (a_1 and a_2) from a uniform distribution on the interval [0,1].
2. The manager observes both ability levels while the employees only know the distribution of their ability levels.
3. The manager assigns tasks to maximise his/her utility
4. Based on the task assignment decision, the employees update their beliefs regarding their abilities according to Bayes' rule.
5. Given the beliefs about their ability levels, employees choose an optimal effort level.
6. Payoffs are realised.

3.2) Task assignment game with a bias towards employee 2

The second extension of the model by Kamphorst and Swank (2016) includes a pre-existing bias towards employee 2, while the model set up in the previous section included a bias towards employee 1. The goal of including a bias towards the opposite employee is to study whether a positive bias toward employee 1 has the same effect as an equal negative bias toward employee 2, and vice versa. In order to account for this new bias, the manager utility function is altered slightly. *Equation (3')* describes the modified manager preferences with a pre-existing bias towards employee 2.

$$U_M(m, a_1, a_2) = \sum_{i=1}^{2} y_i + (m-1)\rho_2 a_2 \quad (3')$$

The second part of the equation will now only be present when employee 2 is assigned to the major task ($m = 2$). As the bias is now directed to employee 2, the bias severity variable (ρ_2) is now interacted with employee 2's ability level. The underlying interpretation is identical to the one in the previous section.

3.3) Task assignment game with a simultaneous bias

Finally, *Equations (3)* and *(3')* can be combined into a single equation with a bias towards the two employees simultaneously. Combining the two gives rise to the following equation:

$$U_M(m, a_1, a_2) = \sum_{i=1}^{2} y_i + (2-m)\rho_1 a_1 + (m-1)\rho_2 a_2 \quad (3'')$$

In other words, the manager's level of utility might increase (or decrease) to different extents depending on which employee out of the two receives the major task. The aim of combining the two equations is to examine whether two simultaneous biases can be taken together and interpreted as a *net bias*. For example, if the manager has a strong positive bias towards employee 1 while also having an equally strong positive bias towards employee 2, will these two biases cancel out against each other and produce identical results to the case without bias? In the equation above, the manager now only has a bias toward employee 1 if the major task is assigned to that employee ($m = 1$), while the same goes for employee 2 whenever $m = 2$.

4) Results

In the following section, the framework set up above will be worked out and analysed to uncover the optimal task assignment strategy for the manager and the effect of a pre-existing bias on that strategy. The previous section first described a benchmark model, after which two alterations to that model were introduced. For clarity, the results will be split into three sub-sections discussing each model separately. Afterwards, the insights from all three models will be combined and discussed. The three different models will be labelled as follows:

- *Model (a)*: *The model with a pre-existing manager bias (positive or negative) <u>towards employee 1</u>.*
- *Model (b)*: *The model with a pre-existing manager bias (positive or negative) <u>towards employee 2</u>.*
- *Model (c)*: *The model with a pre-existing manager bias (positive or negative) <u>towards both employees simultaneously</u>.*

Furthermore, to make the interpretation of the results for *Model (a)* and *Model (b)* more convenient, solving these models will be done in two steps. First, the importance of both the major and the minor task will be equalised ($\eta = 1$) to simplify the model and solve it more easily. This simplified version of the model will be referred to as the *basic game*. Afterwards, the model will be solved with the possibility that the major task is more important than the minor one ($\eta \geq 1$), which will be referred to as the *extended game*. On the other hand, solving *Model (c)* will require a slightly different approach as it includes two different bias variables instead of one.

4.1) Analysis of the task assignment game using *Model (a)*

This sub-section focuses on the task assignment game in which the manager has a pre-existing bias towards employee 1. This bias can be both positive or negative and enters the manager utility function described by *Equation (3)* as an interaction term with the ability level of the employee. First, the equilibria will be determined for both the *basic* and *extended games*. Afterwards, the stability of the different possible equilibria will be discussed in more detail to conclude by providing insight into the efficiency of the different possible equilibria.

After observing the employees' ability levels, the manager aims to assign tasks to maximise his/her utility as described by *Equation (3)*. The optimal strategy considers that employees form beliefs about their ability given the task assignment. The manager is indifferent in this task assignment decision if the utility level from assigning the major task to employee 1 equals that of assigning the major task to employee 2. Solving this leads to the following expression:

$$a_1 = ta_2 \tag{5}$$

Where

$$t = \frac{\eta^2 E(a_2|2) - E(a_2|1)}{\eta^2 E(a_1|1) - E(a_1|2) + \rho_1}$$

Section A.1 in the appendix elaborates on how these expressions were derived. *Equation (5)* describes the manager's task assignment decision as a straight line through the origin. The slope is determined by the relative importance of the major task (η), the optimal employee effort levels ($e_i = E(a_i|m)$) and a possible pre-existing bias towards employee 1 (ρ_1). Like this, the optimal task assignment decision is determined for all possible combinations of employee ability levels. The manager assigns the major task to employee 1 ($m = 1$) whenever $a_1 \geq ta_2$ and assigns it to the other employee ($m = 2$) otherwise ($a_1 < ta_2$).

4.1.1) Solution for the basic game ($\eta = 1$)

4.1.1.1) Optimal task assignment strategy

In the *basic game*, the major and minor tasks are assumed to be equally important ($\eta = 1$). Therefore, the optimal employee effort levels described in *Equation (4)* can be generalised to $e_i = E(a_i|m)$ regardless of whether the employee is assigned to the major ($m = i$) or minor ($m \neq i$) task. After using Bayes' rule to determine the employees' beliefs about their ability levels given the manager's task assignment decision, t can be determined as a function of the pre-existing bias towards employee 1. The resulting equilibria of the model are summarised in *Proposition 1* below. Calculations are provided in *Section A.2* in the appendix.

Proposition 1

Depending on the level of bias towards employee 1, one to three possible equilibria exist in the *basic game* described above. The optimal values for the manager task assignment strategy (t^*) are determined for $t \in [0,1]$ and $t > 1$ separately. The equilibrium strategies are equal to $t^* = \frac{\sqrt{36\rho_1^2+24\rho_1+1}+6\rho_1+3}{6\rho_1+4}$ and $t^* = \frac{-\sqrt{36\rho_1^2+24\rho_1+1}+6\rho_1+3}{6\rho_1+4}$ when $t \in [0,1]$, while they are equal to $t^* = \frac{\sqrt{9\rho_1^2-30\rho_1+1}+3\rho_1+3}{12\rho_1+2}$ and $t^* = \frac{-\sqrt{9\rho_1^2-30\rho_1+1}+3\rho_1+3}{12\rho_1+2}$ for $t > 1$.

Using these equilibrium task assignment strategies, *Figure 2* depicts the optimal level of t as a function of the pre-existing manager bias towards employee 1 (ρ_1). The figure shows how the level of bias determines the number of equilibria in the game. Visually, the different levels of bias are depicted by horizontal lines so that the intersections depict the equilibria in the game for each level of manager bias towards employee 1. In what follows, the different possible equilibria will be explained in more detail.

4.1.1.2) Discussion of the equilibria without bias

First of all, without any bias ($\rho_1 = 0$), solving the task assignment game results in three equilibria: $t^* = \frac{1}{2}$, $t^* = 1$, and $t^* = 2$. In *Figure 2*, these equilibria are depicted as the intersection points with the x-axis. In the non-discriminatory equilibrium ($t^* = 1$), the manager bases the task assignment solely on the ability level of the employees. The employee with the highest ability level is assigned to the major task, while the other is assigned to the minor task. As depicted by *Equation (5)*, the only case in which the manager is indifferent between the two employees is when their ability levels are identical. On the other hand, there are two discriminatory equilibria in which the manager discriminates in favour of employee 1 $\left(t^* = \frac{1}{2}\right)$

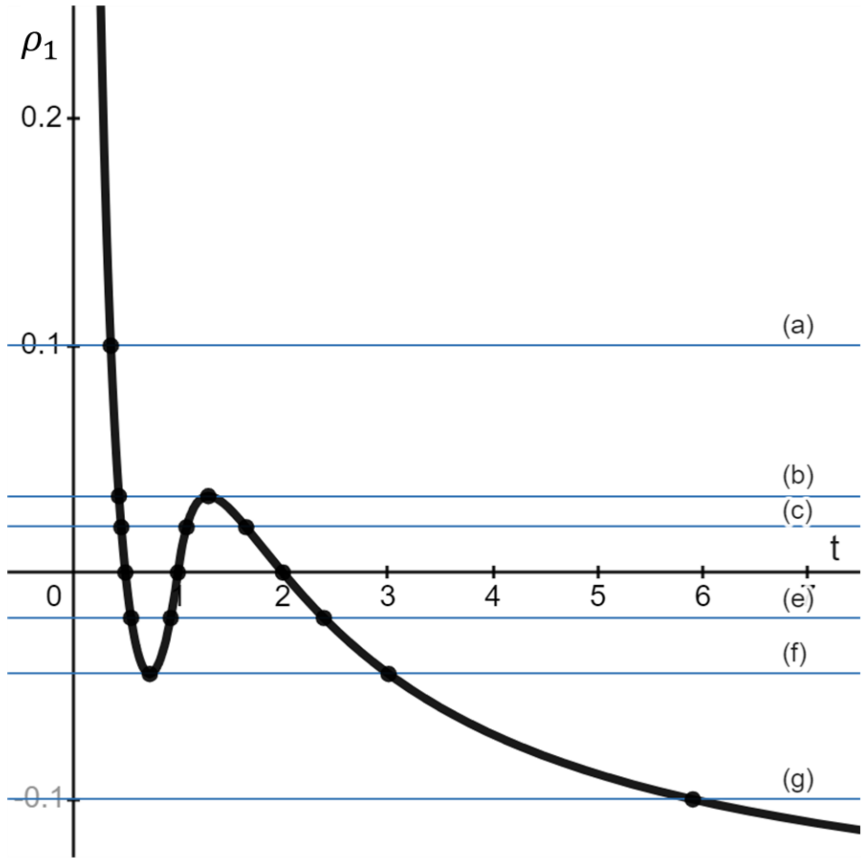


Figure 2: Graphical representation of the equilibria in the basic game for different levels of manager bias. The curve depicts the relationship between the bias (ρ_1) and the optimal task assignment strategy t.

or employee 2 ($t^* = 2$). For instance, in the case where $t^* = \frac{1}{2}$, the manager assigns the major task to employee 1 if his/her ability level is equal to or greater than half the ability level of employee 2. Remember that these ability levels are independently drawn from uniform distributions on the interval [0,1]. In other words, even if employee 2 has a maximum level of ability ($a_2 = 1$), the major task would still be assigned to employee 1 whenever he/she has an ability level equal to or greater than $\frac{1}{2}$. The opposite is true for the equilibrium $t^* = 2$.

These discriminatory equilibria result from the fact that optimal employee effort levels depend upon their beliefs about their ability. As depicted by *Equation (4)* above, effort and ability are assumed to be complements. Consequently, when employees expect discrimination in favour of employee 1 ($t^* = \frac{1}{2}$), assigning the major task to employee 2 will disproportionately demotivate employee 1. This is because such a task assignment decision would send a signal to employee 1 that his/her ability level is smaller than half the ability level of the other employee. On the other hand, if the manager discriminates as expected and assigns the major task to the expected employee (employee 1 in this case), the self-confidence of the disfavoured employee (employee 2) is less affected. This is because even with a maximum level of ability ($a_2 = 1$), the major task is assigned to employee 2 in only half the cases. The manager thus follows the employees' beliefs and discriminates to not disproportionately demotivate employee 1 in the minor task. This line of reasoning is drawn directly from Kamphorst and Swank (2013). It should be kept in mind that the *basic model* assumes the major and minor tasks to be equally important ($\eta = 1$). The *extended model* below will loosen this assumption and uncover how a higher level of major task importance affects discrimination. Generally, there is discrimination in favour of one of the two employees because the manager follows employees' expectations. This makes the discriminatory equilibria self-fulfilling.

4.1.1.3) Discussion of the equilibria with bias

Varying levels of pre-existing bias affect the equilibrium levels of t in the game, which in turn changes the slope of the linear function described by *Equation (5)*. *Figure 3* below uses a different graphical representation for the equilibria in the game. It depicts the function described by *Equation (5)* on a plane with the employee ability levels (a_1 and a_2) on the axes. The areas above each curve represent the a priori combinations of a_1 and a_2 for which employee 1 is assigned to the major task ($m = 1$). In contrast, the areas below depict the cases in which it is the other employee to whom it is given ($m = 2$).

For instance, panel d depicts the case discussed previously in which there is no pre-existing manager bias. In the three resulting equilibria, the a priori chances for each employee to get assigned to the major task are either equal ($t^* = 1$) or three times larger for employee 1 ($t^* = \frac{1}{2}$) or employee 2 ($t^* = 2$). However, with a substantial positive bias in favour of employee 1 as depicted in panel a, only one equilibrium remains in which employee 1 is over four times more likely to receive the major task. On the flip side, with a negative bias of the same magnitude against employee 1 shown in panel g, he/she is more than ten times less likely to get assigned to the major task. In other words, a negative bias against employee 1 worsens discrimination considerably more than an equally positive bias. Additionally, it is important to notice that the equilibrium moves to perfect discrimination in favour of employee 1 ($t^* = 0$) as the bias towards that employee gets closer and closer to infinity ($\rho_1 \to \infty$). However, it only takes a negative bias equal to $\rho_1 = -\frac{1}{6}$ for there to be perfect discrimination against employee 1 ($t^* \to \infty$). The optimal t resulting from a positive bias is not equal to the inverse of that resulting from a negative bias. This shows how the effect of a positive bias does not mirror that of a negative bias. *Section 4.2.2* will elaborate further on this comparison.

A manager with a positive bias towards employee 1 will always discriminate in favour of that employee given that the bias is sufficiently high. A high level of positive bias will worsen the level of discrimination but does not exclude that employee 2 might still receive the major task. Even with a positive bias towards employee 1, the manager will remain somewhat rational and consider the employees' ability levels. Whenever the favoured employee has a low ability level, the manager still optimally assigns the major task to the other (more capable) employee. However, due to the skill paradox described by Dietz et al. (2015), a negative bias is aggravated by a higher level of ability. On the one hand, an employee will need a relatively higher level of ability to receive the major task whenever the manager has a negative bias against him/her. On the other hand, a higher level of ability will worsen the negative bias even further. Therefore, there is a limit to a manager's negative bias before perfectly discriminating against that employee.

Panel b shows the possible equilibria in the game with a maximum positive bias towards employee 1 for which an equilibrium still exists with discrimination in favour of the other employee. Any positive bias greater than this maximum level would result in a single discriminatory equilibrium in favour of employee 1 depicted by panel a. In other words, some level of positive bias can result in an equilibrium with a lower level of discrimination than the

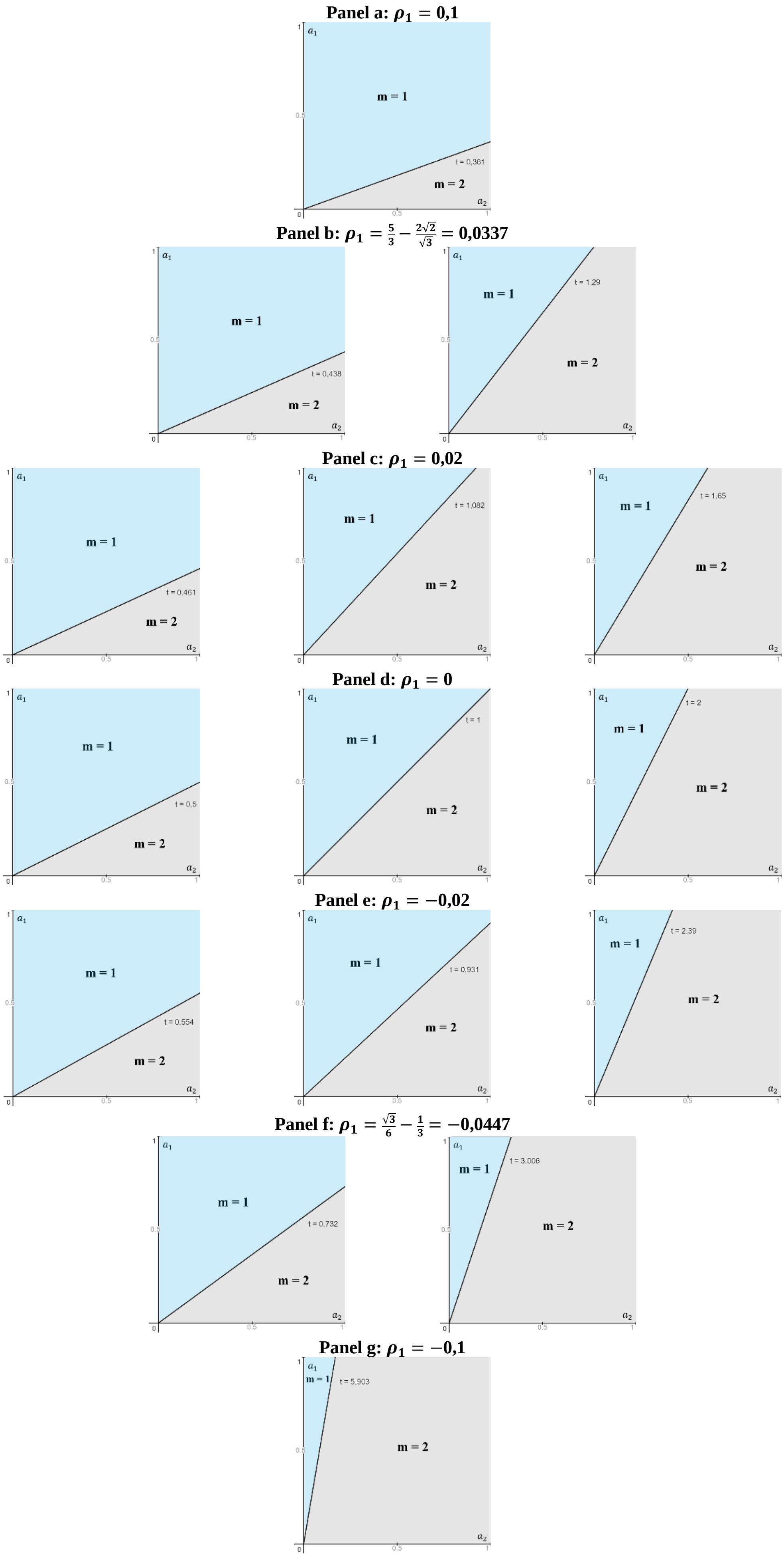


Figure 3: Graphical representation of the manager's task assignment strategy depicted by the slope of the linear relationship between a_1 and a_2. The different panels show the equilibria for varying levels of bias.

case without bias. A positive bias towards one employee might still incite the manager to discriminate in favour of the other. For example, a father (manager) with a positive bias towards his son (employee 1) might still, to some extent, discriminate in favour of the other employee because he opposes nepotism[1].

On the other hand, panel f depicts the case with a maximum negative bias against employee 1 for which an equilibrium still exists with discrimination in favour of that employee. A greater negative bias would lead to only one equilibrium in which the manager only discriminates in favour of employee 2. For example, a manager with a negative bias against a female (employee 1) might still discriminate in favour of that employee in a promotion decision because he aims to diversify the workplace or modernise the company's image. This shows that a negative bias also does not exclude the manager from discriminating in the opposite direction as long as the bias does not become too severe.

While the strategies resulting from greater positive and negative biases discussed above produce intuitive results, this is not the case for smaller biases at first sight. For instance, panel c shows the three equilibria in the game with a slight manager bias in favour of employee 1 ($\rho_1 = 0{,}02$). The two outside graphs of panel c remain intuitive. On the one hand, a slight positive bias towards employee 1 worsens the discriminatory equilibrium in favour of that employee, as the discriminatory equilibrium $t^* = 0{,}5$ moves towards more discrimination ($t^* = 0{,}461$). On the other hand, the discriminatory equilibrium in favour of employee 2 moves away from $t^* = 2$ closer to the non-discriminatory equilibrium ($t^* = 1{,}65$). However, a slight positive bias towards employee 1 makes the non-discriminatory equilibrium ($t^* = 1$) slightly discriminatory in favour of employee 2 ($t^* = 1{,}082$). To understand why a bias towards one employee causes discrimination in favour of the other, it should be noted that the following section will show how the discriminatory equilibria ($t^* = 0{,}5$ and $t^* = 2$) are stable while the non-discriminatory equilibrium ($t^* = 1$) is not. On the one hand, the manager discriminates to not disproportionately demotivate one of the employees. However, ability serves as a limit to the level of discrimination because the manager cares less about demotivating an employee whenever he/she has a low ability level (Kamphorst and Swank, 2016). These two mechanisms combined result in the discriminatory equilibria being stable in the model without bias.

Thus, stability works as a force making it optimal for the manager to discriminate in equilibrium. Given employee beliefs, the manager's optimal task assignment decision will

[1] Nepotism is defined as using your power or influence to benefit members of your family.

move to $t^* = 0{,}5$ or $t^* = 2$. On the other hand, a positive or negative pre-existing bias towards employee 1 works as a force pulling the non-discriminatory equilibrium towards a discriminatory equilibrium in favour of employee 1 or employee 2, respectively. For instance, a slight positive bias towards employee 1 makes the manager want to discriminate in favour of employee 1. Thus, it pulls the optimal task assignment decision from $t^* = 1$ towards $t^* = 0$. In order for an equilibrium to exist, the non-discriminatory equilibrium becomes equal to $t^* = 1{,}082$ because here the upward force of stability pulling the equilibrium towards $t^* = 2$ balances out the downward force of the bias. The equilibria in panel e mirror those in panel c and follow the same reasoning.

Summary of section 4.1.1

- No pre-existing manager bias results in one non-discriminatory and two discriminatory equilibria. One in favour of employee 1, the other in favour of employee 2.
- Discrimination is self-fulfilling: the manager follows the employees' beliefs about discrimination to not disproportionately demotivate the employee that expects to be favoured (Kamphorst and Swank, 2013).
- A high enough positive or negative bias will result in only one discriminatory equilibrium in favour of employee 1 or employee 2, respectively.
- There exists a maximum level of positive bias towards employee 1 for which the manager can still discriminate in favour of employee 2.
- There exists a maximum level of negative bias against employee 1 for which the manager can still discriminate in favour of employee 1.
- With a slight positive bias towards employee 1, there exists one equilibrium with discrimination in favour of employee 1 and two equilibria with discrimination in favour of employee 2.
- With a slight negative bias towards employee 1, there exists one equilibrium with discrimination in favour of employee 2 and two equilibria with discrimination in favour of employee 1.

4.1.2) Equilibrium stability

This section focuses on the stability of the equilibria found for the *basic game* described above to determine which ones are most likely to arise. The analysis below will first determine the stability of the equilibria without bias ($\rho_1 = 0$) to then examine the effect of the bias becoming either positive or negative. For each case, stability will be determined for $t \in [0,1]$ and $t > 1$ separately.

4.1.2.1) Equilibrium stability without bias

To uncover the stability of each equilibrium, it is assumed that employees have beliefs about the task assignment decision $(\hat{t})$ that differ from the equilibrium task assignment strategy (t). The manager will then assign tasks optimally taking into account the employees' beliefs. For instance, considering the cases for $t \in [0,1]$, the equilibrium $t = 1$ is considered to be stable if, given employees' beliefs $(\hat{t})$ smaller than 1, the manager chooses an optimal task assignment strategy (t) even closer to 1. In other words, an equilibrium is considered stable if $\hat{t} < t$. This being the case would imply that the equilibrium would converge towards $t = 1$ for any $\hat{t}$. On the other hand, an optimal task assignment strategy smaller than $\hat{t}$ $(\hat{t} > t)$ would imply that the manager discriminates more than anticipated. This is because a task assignment strategy closer to zero implies more discrimination. The equilibrium $t = 1$ would then be unstable as t moves further and further away from it. Using this rule, the stability of the equilibria without bias is summarised in *Proposition 2.1*. Calculations are provided in *Section A.3* in the appendix.

Proposition 2.1

As proven by Kamphorst and Swank (2013), only the discriminatory equilibria ($t^* = \frac{1}{2}$ and $t^* = 2$) are stable while the non-discriminatory equilibrium ($t^* = 1$) is unstable in the *basic game* without manager bias ($\rho_1 = 0$).

The fact that only the discriminatory equilibria are found to be stable is the result of two main forces. On the one hand, it was already discussed above that the manager follows the employees' beliefs and discriminates because one of them would otherwise be disproportionately demotivated by the task assignment decision. On the other hand, the manager takes the ability levels of the employees into account when assigning the tasks. This ensures that there is a limit to the level of discrimination. For instance, the manager would not want to discriminate in favour of an employee with a significantly low relative ability level. These two forces combined make the discriminatory equilibria $t^* = \frac{1}{2}$ and $t^* = 2$ stable.

4.1.2.1) Equilibrium stability with bias

The proposition above focuses on the stability of the equilibria in a game without bias. However, the previous section showed how a bias towards employee 1 affected the level of discrimination and the number of equilibria. *Section A.4* in the appendix provides calculations on the stability of the equilibria given the four possible ranges of bias in the *basic game*. Given the findings from *appendix A.4*, *Proposition 2.2* summarises the main findings regarding the stability of the equilibria in a game with a pre-existing manager bias towards employee 1.

Proposition 2.2

The stability of the equilibria in a *basic game* with manager bias towards employee 1 can be separated into six cases of bias (or four ranges):

Case 1: $\boldsymbol{\rho_1 > \frac{5}{3} - \frac{2\sqrt{2}}{\sqrt{3}}}$ (= **0,0337**)

Whenever the manager has a high positive bias towards employee 1, the only equilibrium that exists is discriminatory in favour of employee 1 and stable.

Case 2: $\boldsymbol{\rho_1 = \frac{5}{3} - \frac{2\sqrt{2}}{\sqrt{3}}}$ (= **0,0337**)

This is the maximum level of positive bias for which an equilibrium still exists with discrimination in favour of employee 2. In this case, there exists one stable discriminatory equilibrium in favour of employee 1 and one unstable discriminatory equilibrium in favour of employee 2.

Case 3: $\boldsymbol{0 < \rho_1 < \frac{5}{3} - \frac{2\sqrt{2}}{\sqrt{3}}}$ (= **0,0337**)

Whenever the manager has a slight positive bias towards employee 1, three discriminatory equilibria exist, two of which are stable. The only stable equilibria at this level of bias are the ones with the highest level of discrimination. The one equilibrium in the middle, although discriminatory, is unstable.

Case 4: $\boldsymbol{0 > \rho_1 > \frac{\sqrt{3}}{6} - \frac{1}{3}}$ (= **−0,0447**)

Whenever the manager has a slight negative bias towards employee 1, three discriminatory equilibria exist, two of which are stable. The only stable equilibria at this level of bias are the ones with the highest level of discrimination. The one equilibrium in the middle, although discriminatory, is unstable.

Case 5: $\boldsymbol{\rho_1 = \frac{\sqrt{3}}{6} - \frac{1}{3}}$ (= **−0,0447**)

This is the maximum level of negative bias for which an equilibrium still exists with discrimination in favour of employee 1. In this case, there exists one stable discriminatory equilibrium in favour of employee 2 and one unstable discriminatory equilibrium in favour of employee 1.

Case 6: $\boldsymbol{\rho_1 < \frac{\sqrt{3}}{6} - \frac{1}{3}}$ (= **−0,0447**)

Whenever the manager has a high negative bias towards employee 1, the only equilibrium that exists is discriminatory in favour of employee 2 and stable.

Figure 4 visualises the stability of the equilibria found in *Section 4.1.1.2*. Those marked by green[2] points are stable, while the red[3] equilibria are unstable. This shows how levels of t within the interval $[0{,}732\,;1{,}29]$ are never stable. The graph shows how certain levels of bias can result in equilibria with a lower level of discrimination compared to the equilibria without bias ($t^* = 0{,}5$ and $t^* = 2$). However, there is a limit to how far discrimination can be reduced before the equilibrium loses stability. Cases 3 and 4 in *Proposition 2.2* show how only the most discriminatory equilibria remain stable with slight levels of bias. A positive or negative bias alters the level of discrimination in the two stable equilibria, always making one equilibrium less and the other more discriminatory. In the stable discriminatory equilibria, the manager still discriminates to not disproportionately demotivate one employee while also taking into account the employees' ability levels.

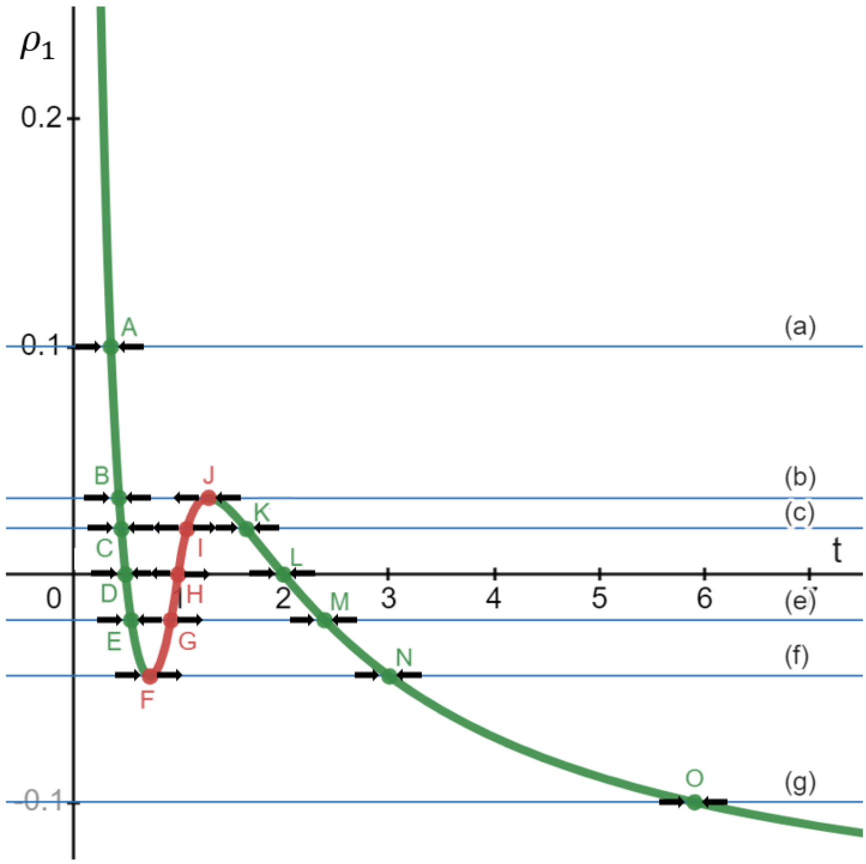


Figure 4: Graphical representation of the equilibria in the basic game for different levels of manager bias. The curve depicts the relationship between the bias (ρ_1) and the optimal task assignment strategy t. The arrows visualise the stability for each of the equilibria.

[2] The points labelled A, B, C, D, E, K, L, M, N, and O.
[3] The points labelled F, G, H, I, and J.

This section has shown that only the discriminatory equilibria are stable whenever there is no manager bias towards employee 1. Once the bias becomes positive or negative, the non-discriminatory equilibrium disappears while the discriminatory ones remain stable. The analysis also shows how all equilibria within a certain range of the non-discriminatory equilibrium are never stable. Whereas the results in the previous section showed how a bias worsens discrimination, this section underlines how it is always the most discriminatory equilibria that are stable.

4.1.3) Equilibrium efficiency

Whereas the previous two sections focused on determining the equilibria and the likeliness of each equilibrium to arise, the question should be asked which one is the most efficient. Several payoffs should be considered: First, the manager aims to maximise his/her utility as described by *Equation (3)*. Second, the employees want to maximise their utility depicted by *Equation (2)*. Third, the firm's payoff equals an unbiased manager's payoff. This is because the firm only intends to maximise output, whereas the manager is influenced by a pre-existing bias towards one of the employees. *Section A.5* in the appendix provides the calculations on the efficiency of the equilibria leading to *Proposition 3*.

Proposition 3

In the *basic game* with a manager bias towards employee 1:

- The firm maximises its output when $t = 1$.
- Both employees *individually* prefer some level of discrimination in their favour.
- The *combined* utility for the employees is maximised when $t = 1$.
- The manager's utility is maximised with some level of discrimination. The direction and severeness of this discrimination is determined by the sign and magnitude of the bias. Without a bias, the manager also prefers $t = 1$.

The proposition above shows how a bias towards one of the employees causes the manager's and firm's objectives to diverge. Without bias, the only goal for the manager and the firm is to maximise output. The way to do so is without discrimination $(t = 1)$ because only then does the task assignment strategy provide the most information about the ability levels of the employees. However, the manager now also cares whether employee 1 specifically is assigned to the major task. In other words, a bias causes the manager to deviate from what is objectively optimal for the firm due to being more concerned with his/her own preferences.

Whenever the manager has a positive bias towards employee 1, some level of discrimination in favour of that employee would be payoff maximising. Additionally, as the manager's utility is increasing in the bias, the payoff would be higher at this discriminatory equilibrium. On the other hand, a negative bias against employee 1 makes it payoff maximising to discriminate in favour of employee 2. However, even with discrimination in favour of employee 2, there is still a chance that employee 1 is assigned to the major task. This can only be the case if employee 1 has a sufficiently higher ability level than employee 2. Therefore, a negative bias causes the manager's payoff to be lower because it interacts with the ability level of employee 1 in the manager's utility function. The manager's payoff is lower because of those cases in which employee 1 is assigned to the major task regardless of the negative bias.

Just like for the firm, the combined payoff for the employees is maximised in the non-discriminatory equilibrium. Although, the employees individually prefer some level of discrimination in their own favour. If they could choose any task assignment strategy t for the manager to follow, what would be the first-best option? Interestingly, reaching the first-best level t for employee 1 would require a slight negative bias against him/herself, while employee 2 would prefer some positive bias in favour of employee 1. This raises the question of whether it would be truly optimal for the employees to have a bias against themselves or in favour of the other employee. For example, employee 1's payoff is maximised at $t = \frac{3}{2} - \frac{\sqrt{3}}{2} (\approx 0{,}634)$. Reaching this stable discriminatory equilibrium requires a manager bias of $\rho_1 = -0{,}0377$. However, it can also give rise to an equilibrium $t = 2{,}811$ in which the negative bias causes more discrimination in favour of employee 2, resulting in a payoff for employee 1 of around 11% lower than optimal. On the other hand, a positive bias of $\rho_1 = 0.04$ gives rise to a unique discriminatory equilibrium $t = 0{,}429$ in which the payoff is only 1,3% lower than optimal. Therefore, the question is whether the employee prefers a maximum payoff with the risk of being considerably worse off or certain discrimination in his/her favour with a slightly lower payoff. Answering this question is beyond the scope of this paper, as the model does not predict which of the discriminatory equilibria is most likely to emerge. Nonetheless, it is something worth considering when interpreting the results.

For the joint employees' interests together with that of the firm, it would be optimal for the manager to be unbiased and not discriminate. Whenever the manager becomes (either positively or negatively) biased towards one of the employees, the payoff maximising equilibrium ($t = 1$) disappears, making the firm unable to maximise production. Additionally, an increasingly biased manager is more and more concerned with his/her own utility instead of the firm's interests. In other words, Both for the firm and the employees jointly, the first-best solution would be a task assignment strategy of $t = 1$. This would also be optimal for an unbiased manager. However, any level of bias makes the manager want to diverge from this strategy and discriminate.

4.1.4) Solution for the extended game ($\eta \geq 1$)

4.1.4.1) Optimal task assignment strategy

The *basic game* in the previous section was solved by assuming that both the major and minor tasks are equally important. However, different tasks will have different levels of importance to a firm in reality. The model can be modified to account for different levels of task importance by allowing the parameter η in *Equation (5)* to be greater than one ($\eta \geq 1$). Again, the analysis below first focuses on a model with an unbiased manager to then uncover how a bias affects the results. *Section A.6* and *A.7* in the appendix provide calculations on how major task importance greater than one affects the equilibria found in the *basic game* and their stability. The main points for the game without bias are summarised in *Proposition 4.1*.

Proposition 4.1

In the *extended game* without a manager bias towards employee 1, there are three possible equilibria: $t^* = 1$, $t^* = \frac{\eta^2 - \sqrt{17\eta^4 - 26\eta^2 + 13} + 1}{2 - 2\eta^2}$ (for $t \in [0,1]$), and $t^* = \frac{\eta^2 + \sqrt{17\eta^4 - 26\eta^2 + 13} + 1}{8\eta^2 - 6}$ (for $t > 1$). The non-discriminatory equilibrium is unstable, while the two discriminatory equilibria are stable. Whenever $\eta \geq \sqrt{1{,}5}$, there is only one equilibrium in the game ($t^* = 1$). It is non-discriminatory and stable (Kamphorst and Swank, 2013).

Whereas the *basic game* resulted in two stable discriminatory equilibria, extending the model by increasing the major task importance shows that the game does not necessarily result in

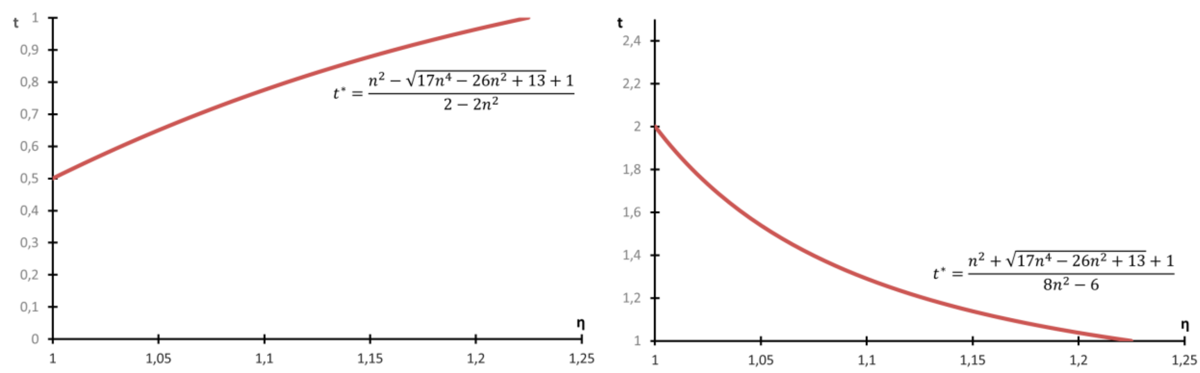


Figure 5: Graphical representation of the optimal task assignment strategy (t) as a function of the major task importance parameter (η).

discrimination. On the contrary, *Proposition 4.1* and *Figure 5* show that a high enough major task importance can be seen as a counter against it. As the major task becomes relatively more important, the equilibria in the game move closer and closer to the non-discriminatory one. Additionally, whenever η is sufficiently high ($\eta \geq \sqrt{1{,}5}$), the unique equilibrium ($t^* = 1$) is non-discriminatory and stable. A high enough major task importance eradicates discrimination.

The intuition behind this result is straightforward. The *basic game* discussed how the manager follows the employees' beliefs in discriminating to not disproportionately demotivate the employee that expects discrimination in his/her favour. If the manager goes against these beliefs, the employee assigned to the minor task is demotivated and exerts less effort. However, a higher level of η implies that the major task becomes relatively more important than the minor one. Therefore, a high level of major task importance implies that the manager cares less about the employee's motivation in the minor task. Consequently, if the major task is important enough, the manager will base the task assignment decision solely on the ability levels of the employees.

Including a manager bias towards employee 1 into the model drastically changes the results discussed above. *Proposition 4.2* summarises the main findings of such a model.

Proposition 4.2

In the *extended game* with a manager bias towards employee 1, the non-discriminatory equilibrium becomes unreachable. Any level of bias makes it impossible for a non-discriminatory equilibrium to exist, no matter how high the major task importance (η) becomes.

- Given that $t \in [0,1]$, a <u>positive bias</u> towards employee 1 makes it impossible to reach the non-discriminatory equilibrium ($t^* = 1$) no matter how important the major task becomes compared to the minor task. There will always be a **discriminatory equilibrium in favour of employee 1**.
- Given that $t > 1$, a <u>positive bias</u> towards employee 1 makes the equilibrium less discriminatory. However, an increasing positive bias in favour of employee 1 combined with increasing major task importance makes a **discriminatory equilibrium in favour of employee 2 less likely to exist.**
- Given that $t \in [0,1]$, a <u>negative bias</u> towards employee 1 makes the equilibrium less discriminatory. However, an increasing negative bias against employee 1 combined

with increasing major task importance makes a **discriminatory equilibrium in favour of employee 1 less likely to exist.**

- Given that $t > 1$, a negative bias towards employee 1 makes it impossible to reach the non-discriminatory equilibrium ($t^* = 1$) no matter how important the major task becomes compared to the minor task. There will always be a **discriminatory equilibrium in favour of employee 2**.

Higher major task importance in a model with either a positive or negative pre-existing manager bias makes the equilibrium less discriminatory but never non-discriminatory. The model without bias showed how a high enough major task importance could force the manager to choose a non-discriminatory task assignment rule. However, a pre-existing bias towards employee 1 causes the major task importance to lose power as a tool to counter discrimination. In the model with an unbiased manager, higher major task importance made the manager care less about demotivating the employee who receives the minor task. This is because the manager cares more about giving the (more important) major task to the most capable employee rather than keeping morale high for the employee who receives the relatively less important minor task. However, a positive bias makes the manager care about the employee towards whom the bias is directed regardless of how important the major task becomes. Therefore, even though the level of discrimination still decreases with a higher level of major task importance, there will always be some level of discrimination in favour of that employee.

Additionally, discrimination in the opposite direction of the bias is less likely to occur when combining it with a high level of major task importance. In other words, a positive bias towards employee 1 combined with high major task importance make it less likely for the manager to discriminate in favour of employee 2. There are two forces at work preventing such an equilibrium from arising. First, a positive bias towards employee 1 makes the manager want to discriminate in favour of that employee. Second, high major task importance counters discrimination in general as the manager cares less about confirming the employees' beliefs. These forces combined will prevent a discriminatory equilibrium in favour of employee 2 from arising. The same line of reasoning can be applied to the mirroring case.

Altogether, a model without bias suggests that a high enough task importance can counter discrimination and force a unique and stable non-discriminatory equilibrium to exist. However, introducing a positive or negative bias shows that countering discrimination is more challenging. Modifying the model to be closer to reality underlines why countering discrimination in the workplace has proven to be complicated. Including a bias allows this model to provide a more nuanced view on the challenges of eradicating discrimination.

4.2) Analysis of the task assignment game using *Model (b)*

This section will discuss an alteration of the previous model. Instead of a pre-existing manager bias towards employee 1, *Model (b)* incorporates a similar bias towards the other employee. In what follows, a brief discussion of the model will be provided first. Afterwards, *Models (a)* and *(b)* will be compared to determine how the effects of a positive bias differ from those of a negative one.

4.2.1) Solution and discussion of *Model (b)*

A pre-existing bias directed towards employee 2 instead of employee 1 directly changes the optimal task assignment decision for the manager. *Sections B.1* and *B.2* in the appendix set up and solve this new altered game. How the *basic game* is solved is identical to *Model (a)* except for the bias parameter. The optimal task assignment rule is now defined as follows:

$$a_1 = ta_2 \qquad (5')$$

Where

$$t = \frac{\eta^2 E(a_2|2) - E(a_2|1) + \rho_2}{\eta^2 E(a_1|1) - E(a_1|2)}$$

Using *Equation (5')*, the equilibria of the game using *Model (b)* are determined and summarised in *Proposition 5* below. *Figure 6* depicts the equilibria graphically as a function of the bias. Again, the stability of the equilibria is depicted by colouring the stable ones green[4] and the unstable ones red[5].

The figure below shows how a model with a manager bias towards employee 2 mirrors the previous one in which there was a bias towards employee 1. For each level of bias, the equilibrium t is equal to the inverse of that in *Model (a)*. Again, there are four possible ranges of bias with varying levels of discrimination and numbers of equilibria. Logically, a bias equal to zero gives rise to the same equilibria as before ($t^* = \frac{1}{2},\ t^* = 1,\ t^* = 2$). Additionally, an increasing positive bias leads to a continuously higher level of discrimination, while a negative

[4] The points labelled A, B, C, D, E, K, L, M, N, and O.
[5] The points labelled F, G, H, I, and J.

Proposition 5

Depending on the level of bias towards employee 2, one to three possible equilibria exist in the *basic game* described above. The optimal values for the manager task assignment strategy (t^*) are determined for $t \in [0,1]$ and $t > 1$ separately. The equilibrium strategies are equal to $t^* = \frac{\sqrt{9\rho_2^2-30\rho_2+1}+3\rho_2+3}{4}$ and $t^* = \frac{-\sqrt{9\rho_2^2-30\rho_2+1}+3\rho_2+3}{4}$ when $t \in [0,1]$, while they are equal to $t^* = \frac{\sqrt{36\rho_2^2+24\rho_2+1}+6\rho_2+3}{2}$ and $t^* = \frac{-\sqrt{36\rho_2^2+24\rho_2+1}+6\rho_2+3}{2}$ for $t > 1$.

bias equal to -0,1667 $\left(\rho_1 = -\frac{1}{6}\right)$ causes perfect discrimination in favour of employee 2. The ranges of bias for which discrimination can still occur in the opposite direction are identical to the cases described in *Proposition 2.2*. This section will not further elaborate on the stability nor efficiency of the equilibria as both perfectly mirror the findings for *Model (a)* explained previously. Moreover, the *extended model* with a bias towards employee 2 leads to identical conclusions as before. In short, a bias makes it impossible for a non-discriminatory equilibrium to exist in the *basic model*. Furthermore, the discriminatory equilibria that arise from a high enough level of bias are always stable. However, only the two most discriminatory equilibria

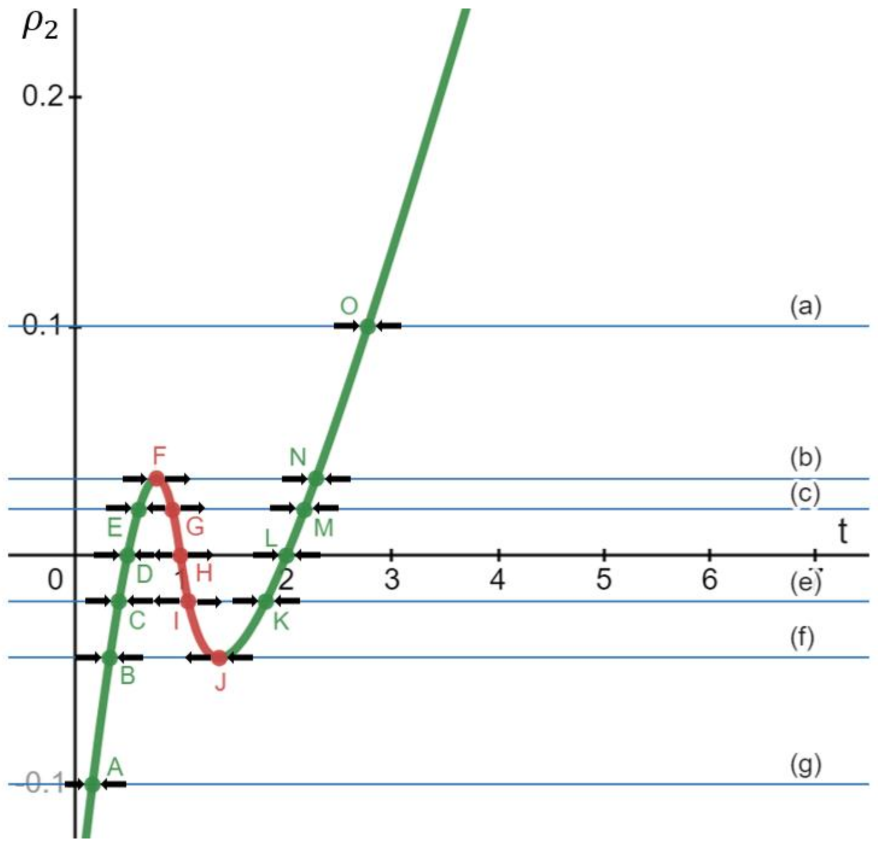


Figure 6: Graphical representation of the equilibria in the basic game for different levels of manager bias. The curve depicts the relationship between the bias (ρ_2) and the optimal task assignment strategy t. The arrows visualise the stability for each of the equilibria.

out of the three remain stable whenever there is only a slight bias. It is payoff maximising for the employees to have some level of discrimination in their favour, even though the non-discriminatory equilibrium is optimal for their joint interests and the firm. For a bias towards employee 1 and employee 2, the *extended game* shows that a bias prevents higher major task importance from countering discrimination. *Part B* of the appendix provides calculations on all of these findings. They are not repeated here, as the previous section already provided a thorough discussion.

4.2.2) *Model (a)* and *(b)* comparison

Even though *Model (b)* -with a bias towards employee 2- perfectly mirrors the results found previously for *Model (a)* -with a bias towards employee 1-, this does not necessarily have to be the case for the biases themselves. This sub-section explores how a positive bias in favour of one employee affects the equilibria compared to an identical negative bias against the other employee.

The first three columns of *Table 1* represent the optimal manager task assignment decisions for varying levels of bias using version *(a)* of the model. On the other hand, the three right-hand side columns were calculated using *Model (b)*. For each of the equilibria, the a priori chance that employee 1 is assigned to the major task ($m = 1$) is calculated. These probabilities are determined as the area above the curves in the different panels of *Figure 3*. Remember that the curve depicted in each graph is defined by *Equation (5)*. Given the employees' ability levels observed by the manager, it is optimal to assign the major task to employee 1 whenever $a_1 \geq ta_2$ while the manager optimally assigns it to employee 2 if $a_1 < ta_2$. Given that ability levels are independently drawn from a uniform distribution on the interval [0,1], the area above the curve contains all possible combinations of a_1 and a_2 resulting in employee 1 receiving the major task. The table below represents these chances as a percentage probability over the total graph area.

For instance, Panel a and Panel g in *Table 1* show how a positive and negative bias of the same magnitude differ in their effect on the optimal task assignment strategy of the manager. On the one hand, a relatively high positive bias towards employee 1 results in a unique discriminatory equilibrium equal to $t^* = 0{,}36$. In this equilibrium, employee 1 receives the major task in 81,95% of all possible combinations of employee ability levels. However, an identical negative bias against employee 2 increases the a priori chance for employee 1 by considerably more. The latter now receives the major task in 91,53% of the cases. On the other hand, Panel g shows

Table 1: Comparison of equilibria for opposite biases in Models (a) and (b).

Bias towards employee 1	Equilibrium t	A priori chance of m = 1	A priori chance of m = 1	Equilibrium t	Bias towards employee 2
		Panel a			
0,1	0,36	81,95%**	91,53%**	0,17	-0,1
		Panel b			
0,0337	0,44	78,08%**	81,53%**	0,37	-0,0337
	1,29	38,76%***	30,48%***	1,64	
		Panel c			
0,02	0,46	76,97%**	79,08%**	0,42	-0,02
	1,08	46,19%	46,53%	1,07	
	1,65	30,31%***	27,69%***	1,81	
		Panel d			
0	0,5	75,00%*	75,00%*	0,5	0
	1	50,00%*	50,00%*	1	
	2	25,00%*	25,00%*	2	
		Panel e			
-0,02	0,55	72,31%***	69,69%***	0,61	0,02
	0,93	53,47%	53,81%	0,92	
	2,39	20,92%**	23,03%**	2,17	
		Panel f			
-0,0447	0,73	63,42%	/	/	0,0447
	3,01	16,63%**	21,13%**	2,37	
		Panel g			
-0,1	5,90	8,47%**	18,05%**	2,77	0,1

Notes: the a priori chance of employee 1 being assigned to the major task is calculated as the area above $a_1 = ta_2$.

how a negative bias against employee 1 reduces the probability of employee 1 receiving the major task more so than an identical positive bias in favour of the other employee.

Panel d shows the results for the *basic game* without bias. As discussed previously, the game has three possible equilibria: two of which are stable and discriminatory, while the other is non-discriminatory but unstable. In the discriminatory equilibria, the table shows how employee 1 is either three times more likely ($t^* = \frac{1}{2}$) or three times less likely ($t^* = 2$) to be assigned to the major task. The outcomes marked in green[6] represent the baseline equilibria of the model. Those in the other panels can be seen as deviations from the baseline results due to an increasing positive or negative bias towards one of the employees.

The equilibria marked in blue[7] show how a negative bias towards one employee has a stronger effect on worsening discrimination in favour of the other employee compared to a positive bias

[6] The outcomes marked with *
[7] The outcomes marked with **

towards the latter. For example, Panel c shows how a slight positive bias in favour of employee 1 increases the probability for this employee to receive the major task from 75% to 76,97% in one of the equilibria. However, an identical negative bias against employee 2 increases this probability to 79,08%. In other words, a negative bias is stronger in making an equilibrium more discriminatory. If the manager is already discriminating in favour of one employee, a positive bias towards that employee will logically increase the level of discrimination. However, for the same equilibrium, a negative bias against the other employee will increase discrimination even more.

On the other hand, the equilibria marked in red[8] seem to show the opposite of what was discussed above. In these cases, a positive bias is shown to have a stronger effect in making an equilibrium less discriminatory. For instance, Panel c shows how a slight positive bias towards employee 1 reduces discrimination in favour of employee 2 from $t^* = 2$ to $t^* = 1{,}65$, increasing the probability for employee 1 to receive the major task from 25% to 30,31%. On the other hand, a negative bias against employee 2 also reduces discrimination in favour of that employee, be it to a lesser extent. The probability for employee 1 to receive the major task now increases from 25% to only 27,69%. This shows how a positive bias is stronger in making an equilibrium less discriminatory. Having a positive bias towards one employee has the strongest effect in making the manager discriminate less in favour of the other employee.

The reason why some of the equilibria in Panel c and Panel e were not marked nor discussed is twofold. First, these equilibria are similarly affected by a bias towards employee 1 and employee 2. The a priori probability of employee 1 receiving the major task is similar for both models. Second, these equilibria were shown to be unstable in *Proposition 2.2*. Taking together the main findings discussed in this section, *Proposition 6* provides a summary. Even though *Table 1* only shows the effect of some specific levels of bias on the equilibrium levels of t, *Section B.7* in the appendix proves that the results in *Proposition 6* hold true for all stable equilibria resulting from positive and negative bias combination across the two models.

Proposition 6

In the *basic game* with a pre-existing manager bias towards employee 1 or employee 2:

- A negative bias against employee A has a stronger effect than a positive bias towards employee B on making an equilibrium more discriminatory in favour of employee B.

[8] The outcomes marked with ***

- A positive bias towards employee A has a stronger effect than a negative bias against employee B on making an equilibrium less discriminatory in favour employee of B.

In other words, whether a positive or negative bias has the strongest effect on an equilibrium depends on whether the bias aggravates or reduces discrimination.

4.3) Analysis of the task assignment game using *Model (c)*

The final part of this study focuses on the analysis of *Model (c)*. Instead of assuming that the manager has a pre-existing bias towards one of the employees, this model attempts to resemble reality more closely by allowing the manager to have a bias towards the two employees simultaneously. Therefore, this model can be considered a combination of the previous ones.

4.3.1) Solution and discussion of *Model (c)*

As before, the first step in solving the model is determining the manager's optimal task assignment strategy. *Equation (5")* depicts the manager's optimal task assignment strategy in the *basic game*. Afterwards, given that employees update their beliefs using Bayes' rule, the equilibria of the game are determined. *Section C.1* and *C.2* in the Appendix provide more information on how *Model (c)* was solved.

$$a_1 = ta_2 \qquad (5")$$

Where

$$t = \frac{\eta^2 E(a_2|2) - E(a_2|1) + \rho_2}{\eta^2 E(a_1|1) - E(a_1|2) + \rho_1}$$

Proposition 7 summarises the equilibria of the *basic game* with a dual bias. The equation above shows how both biases will independently affect the manager's task assignment decision. In other words, the expression for t shows that the biases cannot be taken together into a net bias.

The equilibrium levels of t can thus be expressed as a function of two parameters (ρ_1 and ρ_2). This makes it challenging to depict them graphically in a transparent and interpretable way, as it would require a three-dimensional plane. Therefore, the effects of both biases on the equilibria in the game will be analysed by graphically illustrating the relation between t and one of the parameters while assigning a fixed value to the other. Using this method, the results of the model are presented in *Figure 7* below.

Proposition 7

In the *basic game* with a dual manager bias towards both employees simultaneously, one to three possible equilibria exist. The optimal values for the manager task assignment strategy (t^*) are determined for $t \in [0,1]$ and $t > 1$ separately. The equilibrium strategies are:

$t^* = \frac{\sqrt{36\rho_1^2 - 36\rho_1\rho_2 + 24\rho_1 + 9\rho_2^2 - 30\rho_2 + 1} + 6\rho_1 + 3\rho_2 + 3}{6\rho_1 + 4}$ and $t^* = \frac{-\sqrt{36\rho_1^2 - 36\rho_1\rho_2 + 24\rho_1 + 9\rho_2^2 - 30\rho_2 + 1} + 6\rho_1 + 3\rho_2 + 3}{6\rho_1 + 4}$

when $t \in [0,1]$, while they are equal to

$t^* = \frac{\sqrt{9\rho_1^2 - 36\rho_1\rho_2 - 30\rho_1 + 36\rho_2^2 + 24\rho_2 + 1} + 3\rho_1 + 6\rho_2 + 3}{12\rho_1 + 2}$ and $t^* = \frac{-\sqrt{9\rho_1^2 - 36\rho_1\rho_2 - 30\rho_1 + 36\rho_2^2 + 24\rho_2 + 1} + 3\rho_1 + 6\rho_2 + 3}{12\rho_1 + 2}$

when $t > 1$.

First, note that panels e and f are identical to *Figure 2* and *Figure 6*, respectively. Whenever one of the two bias parameters is equalised to zero, the equilibrium strategies from *Proposition 7* revert back to those found for *Models (a)* and *(b)* discussed in the previous sections.

On the other hand, panels a through d show how a positive bias shifts up and flattens the curves. For instance, a bias towards employee 2 equal to 0,1 results in a unique and stable discriminatory equilibrium in favour of employee 2 as long as the manager has a low or negative bias towards the other employee (in this case $0{,}0818 > \rho_1 > -\frac{1}{6}$). Logically, the level of discrimination in favour of employee 2 increases when the bias towards employee 1 becomes negative while decreasing when the bias is positive. Additionally, whenever the bias towards the two employees are equal and positive ($\rho_1 = \rho_2 \geq 0$), there exist three equilibria. Again, two stable discriminatory equilibria and one unstable non-discriminatory equilibrium. However, as the positive biases towards both the employees increase, the three equilibria at equal levels of bias move closer and closer together and the two discriminatory equilibria become less discriminatory. Even though discrimination still exists due to the risk of demotivation, an equal positive bias towards both the employees will make the manager want to discriminate less. This is because the manager would like to assign the major task to both employees equally. Whenever the manager discriminates in favour of one of them, the level of discrimination will be lower so as not to excessively disadvantage the other employee for whom he/she also has a positive bias. The manager is torn between the two employees in making a task assignment decision. Moreover, whenever the bias towards one of the employees exceeds $\frac{1}{3}$, only one equilibrium remains for all bias combinations. This unique equilibrium is stable and non-discriminatory $(t = 1)$ whenever the manager has an equal bias towards the employees. Otherwise, there is some discrimination in favour of the employee towards whom the manager

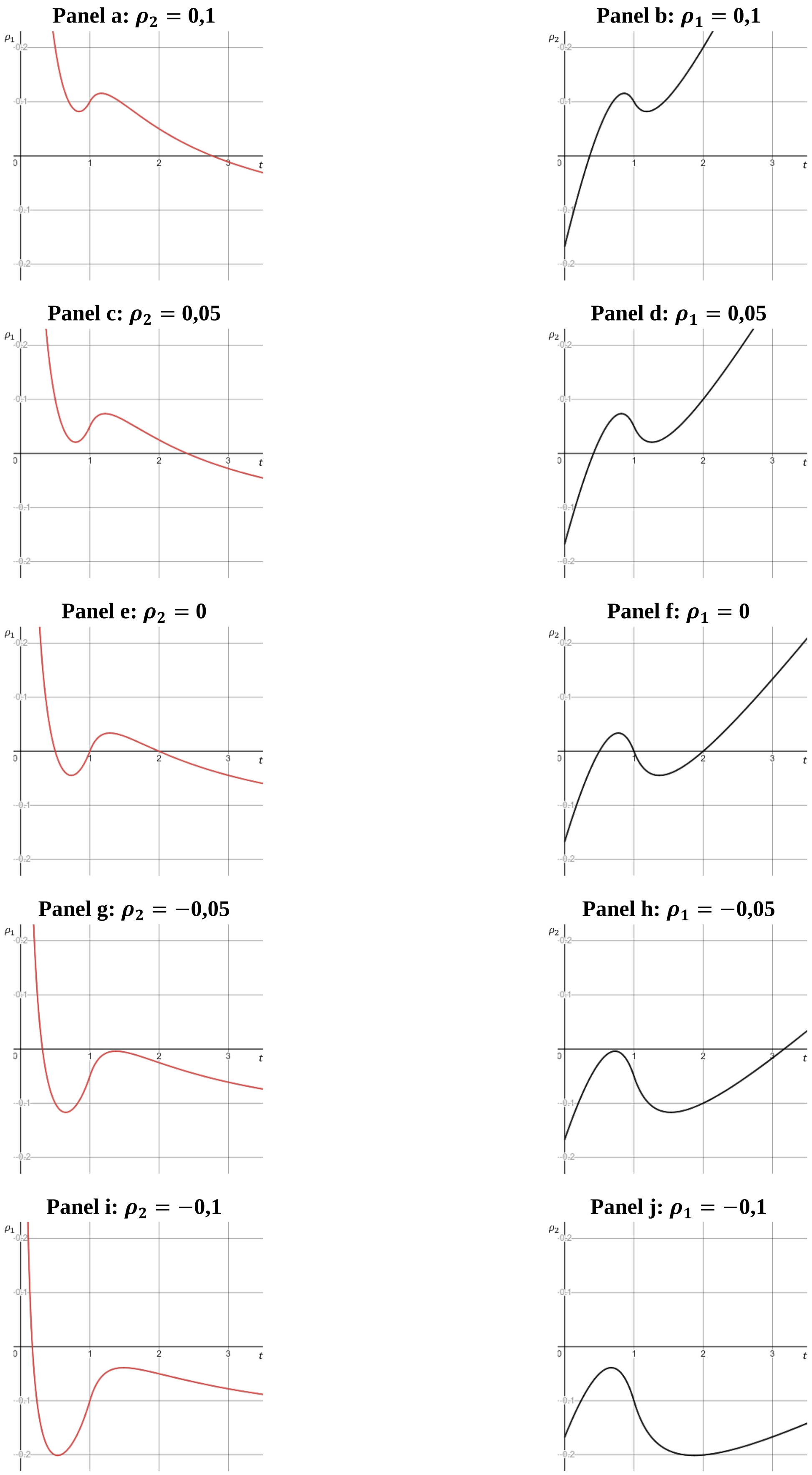


Figure 7: Graphical representation of the equilibria in the basic game for different levels of manager bias towards both employees. The curves depict the relationship between the bias and the task assignment strategy.

has the highest positive bias. In this case, the unique discriminatory equilibrium is stable as well. The calculations determining the stability of the equilibria can be found in *Section C.3* in the appendix.

In other words, whenever the manager has an equal and high positive bias towards the two employees simultaneously, a unique and non-discriminatory equilibrium arises. At this point, the manager favours the employees equally and wants to assign the major task to both of them. Therefore, it is optimal to base the task assignment decision fully on their ability levels. The most capable employee will receive the major task in this case. Naturally, it should be noted that a bias is hard to define in reality. It is hard to say whether pre-existing biases towards two employees can be identical. Whenever one bias is greater than the other, discrimination becomes a unique and stable equilibrium once more.

On the other hand, Panels g through j illustrate how a dual negative manager bias affects the possible equilibria in the game. The graphs show how such a bias shifts the curves downward while increasing the level of discrimination in the discriminatory equilibria. Again, at equal levels of negative bias, there still exists an unstable non-discriminatory equilibrium. However, whereas simultaneous positive biases eventually gave rise to a unique and stable equilibrium, simultaneous negative biases result in equilibria with more and more discrimination. For instance, a negative bias equal to $-\frac{1}{6}$ results in two stable equilibria with perfect discrimination ($t = 0$ and $t = \infty$) and one unstable non-discriminatory equilibrium ($t = 1$). In other words, the model predicts that, instead of basing the task assignment decision on the employee ability levels, an equal simultaneous negative bias makes the manager discriminate more towards one of the employees regardless of the negative bias against that employee.

Figure 7 shows how a positive bias decreases the level of discrimination and could even give rise to a unique and stable non-discriminatory equilibrium, whereas a negative bias has the opposite effect. A possible explanation for this is the interaction between bias an ability in *Equation (3'')*. On the one hand, when the manager has an equal positive bias towards both employees, the one with the highest ability level receives the major task ($t = 1$). When the biases are equal, the ability level will determine in favour of which employee the decision tilts, because a higher ability level aggravates the bias. Additionally, the positive bias being interacted with the ability level of the other employee causes the level of discrimination to be lower when the biases are close but not equal. Even though the manager has a larger bias towards one of the employees, he/she will discriminate less because the positive bias also causes utility to increase if the other employee receives the major task. On the other hand, an equal negative bias causes the manager to perfectly discriminate in favour of one of the employees. In this case, the manager will discriminate in favour of the employee with the lowest ability level due to the skill paradox embedded in the model and formulated by Dietz et al. (2015). A high and equal negative bias will cause the manager to perfectly discriminate in favour of the employee with the lowest level of ability in order to minimise the negative effect of the bias on his/her utility.

The intuition above is heavily based on the interaction between ability and bias due to the fact that a model without this assumption would produce considerably different results. Namely, an equal positive or negative bias towards both employees simultaneously in a model without an interaction between ability and bias would result in the same equilibria as the model without bias. In this case, equal biases would cancel out against each other. Therefore, if a negative bias is not aggravated by the ability level of the employee, an equal negative bias will not result in more discrimination. Although the manager will have a lower level of utility due to the negative bias, he/she will still assign tasks optimally as before instead of giving the task to the least capable employee due to the skill paradox.

The analysis in this section provides an insightful addition to the manager task assignment game set up and discussed in this study. Previously, it was established how a bias towards one of the employees prevented a higher level of major task importance from countering discrimination. However, this section shows how a high enough equal positive bias towards both employees simultaneously results in a unique and stable non-discriminatory equilibrium. Even if the bias is not identical, there will be a lower level of discrimination whenever the positive biases are close. On the other hand, it was established in *Proposition 6* that a negative bias has the strongest effect in worsening discrimination. This section strengthens that finding by showing that a simultaneous negative bias has a similar impact on the equilibria in the game.

5) Conclusion

This paper studies how a “taste for discrimination” affects the level of discrimination in a task assignment game similar to Kamphorst and Swank (2013). In doing so, it extends their model and combines elements from the existing economic literature on confidence management (Bénabou and Tirole, 2003; Ishida, 2006; Crutzen et al., 2013; Prendergast, 1992) with pioneering theoretical work on discrimination (Becker, 1957; Arrow, 1971; Phelps, 1972). Similarly to the literature on confidence management, the model used in this study assumes that a manager has superior information on employee ability levels. Furthermore, ability and effort are assumed to be complementary. On the other hand, the manager is assumed to have a "taste for discrimination" (Becker, 1957). However, the type of discrimination (race, gender, age, et cetera) is not specified.

In the *basic game* with an unbiased manager, the model predicts three equilibria: two of which are discriminatory, stable, and inefficient, while the other is non-discriminatory and efficient but unstable. Kamphorst and Swank (2013) suggest that discrimination is self-fulfilling in this setting. Whenever employees expect a manager to discriminate, it is optimal for the manager to follow these beliefs to not disproportionately demotivate the employee who expects to be favoured. These findings result from a manager having superior information on the employees' ability levels. Therefore, employees form expectations about their ability based on the manager's task assignment decision. This assumption, combined with the complementarity of effort and ability, results in the manager discriminating to keep motivation high.

However, a "taste for discrimination" in the form of a pre-existing manager bias alters the number of equilibria in the game as well as the level of discrimination. A high enough positive or negative bias towards one employee results in a unique discriminatory equilibrium either in favour or against that employee. For smaller levels of bias, there can still exist an equilibrium in favour of the employee who is disadvantaged by the bias (meaning either a positive bias in favour of the other or a negative bias against him/herself). However, a bias prevents a non-discriminatory equilibrium from existing. Additionally, the most discriminatory equilibria are always found to be stable. Depending on the level and sign of the bias, it can either worsen or slightly ameliorate discrimination. The results of the model also indicate that a negative bias has a stronger effect in worsening discrimination, while a positive bias is stronger in ameliorating discrimination.

Moreover, the study by Kamphorst and Swank (2016) showed that discrimination could be avoided by a high enough major task importance. The more important the major task becomes relative to the minor one, the less the manager cares about demotivating the employee in the minor task. However, a "taste for discrimination" makes it impossible to reach the non-discriminatory equilibrium regardless of the major task's importance. A positive bias towards an employee makes the manager still care about that employee regardless of how important the major task is. Therefore, while higher major task importance will still decrease the level of discrimination, it will never make the equilibrium non-discriminatory.

The final part of the analysis showed how an equal and high simultaneous bias towards both employees results in a unique and stable non-discriminatory equilibrium. However, a bias is an abstract concept in reality. A manager having an identical positive bias towards two employees is unlikely. Whenever the manager slightly prefers one over the other, the equilibrium immediately reverts back to discrimination.

This paper adds to the existing literature on taste-based discrimination and strengthens the alarming result found by Kamphorst and Swank (2013) that discrimination is found to be a stable outcome in a simplified setting. Results of the *extended game* indicate that a "taste for discrimination" makes countering it more challenging. This was shown by the fact that higher major task importance no longer gives rise to a unique non-discriminatory equilibrium. This could explain why workplace discrimination has proven to be widespread and persistent.

6) Limitations and future research

Finally, it is important to consider some possible limitations to the methodology used in this study. First of all, the analysis performed uses a theoretical model to study the occurrence of workplace discrimination rather than carrying out an empirical analysis. Therefore, the model makes a number of simplifying assumptions which do not necessarily hold for all real-life situations. For instance, certain behavioural actions could be taken by the employees that are not accounted for in this model. For example, an employee could threaten to leave the company when not receiving the major task. Naturally, this could influence the task assignment decision taken by the manager unless the firm has an up-or-out policy. In such a setting, receiving the minor task would be equal to getting fired from the job. Therefore, the employee's motivation in the minor task is of no importance to the firm, and there would be no discrimination. Moreover, the model assumes that the task assignment is the only signal the employees receive regarding their ability, which is not necessarily the case in reality. For example, a manager might send verbal signals about the employees' job performance, which would also influence how employees perceive their ability level. This idea was already explored by Kamphorst and Swank (2013). In their study, they dedicate a section to extending their model by allowing the manager to send verbal signals to employees in the form of cheap-talk messages. They conclude that cheap-talk aggravates discrimination, while it may also allow the manager to convey more information to the employees regarding their ability levels.

Additionally, the results found in this study heavily rely on the manager having superior information about the employees' ability levels. It should be noted that this would only be the case in specific situations. For instance, when the employees are young, their self-assessment of their ability level is based on limited professional experience (Swank and Visser, 2007). On the other hand, it is assumed that the manager has more experience and has made a similar task assignment decision for other employees before. Therefore, the results found in this model do not necessarily apply to any possible task assignment situation.

A third general limitation to this study is the type of taste-based discrimination assumed when setting up the model. More specifically, the bias is assumed to interact with the employee's ability level in the manager's utility function described by *Equation (3)*. However, a positive or negative bias is not necessarily aggravated by the ability level of the employee. In reality, the manager could very well be biased towards one of the employees regardless of their abilities. In the analysis of the results from *Model (c)*, it was discussed how relaxing this assumption would considerably alter the results found in the model. Although, solving the model algebraically without interacting the ability level with the bias makes it considerably more complicated compared to the current model.

Regarding future research, a possibly interesting refinement to the model could be some way to determine the likelihood of the different equilibria to occur. For instance, the equilibria predictions could be refined by applying the intuitive criterion developed by Cho and Kreps (1987). The intuitive criterion is a way of ruling out less reasonable equilibria in a game and thus to refine the equilibrium predictions made by the model. If an equilibrium does not satisfy this criterion, the manager can make a deviation and send a credible message about doing so. Therefore, such an equilibrium would not be reasonable. Refining the model this way could help in determining the optimal level of bias an employee would prefer to maximise his/her payoff. In *Section 4.1.3*, it was briefly touched upon how determining the first-best level of bias is complicated due to the uncertainty surrounding which equilibrium will manifest. The intuitive criterion could help in ruling out certain unreasonable equilibria.

References

Andriessen, I. (2019). Ethnic Discrimination in the Labour Market: The Dutch Case. In *Race Discrimination and Management of Ethnic Diversity and Migration at Work*. Emerald Publishing Limited.

Arrow, K. J. (1971). Some models of racial discrimination in the labor market. RAND CORP SANTA MONICA CA.

Becker, G. S. (1957). *The economics of discrimination*. Chicago: University of Chicago Press.

Bénabou, R., & Tirole, J. (2003). Intrinsic and extrinsic motivation. *The review of economic studies*, *70*(3), 489-520.

Busetta, G., Campolo, M. G., & Panarello, D. (2018). Immigrants and Italian labor market: statistical or taste-based discrimination?. *Genus*, *74*(1), 1-20.

Carlsson, M., & Rooth, D. O. (2012). Revealing taste-based discrimination in hiring: a correspondence testing experiment with geographic variation. *Applied Economics Letters, 19*(18), 1861-1864.

Cho, I. K., & Kreps, D. M. (1987). Signaling games and stable equilibria. *The Quarterly Journal of Economics*, *102*(2), 179-221.

Coate, S., & Loury, G. C. (1993). Will affirmative-action policies eliminate negative stereotypes?. *The American Economic Review*, 1220-1240.

Cooley, C. H. (1902). Looking-glass self. The production of reality: Essays and readings on social interaction, 6, 126-128.

Courant, P. N. (1978). Racial prejudice in a search model of the urban housing market. *Journal of Urban Economics*, *5*(3), 329-345.

Cronin, F. J. (2021). Racial differences in the search for housing. In *Modelling housing market search* (pp. 81-105). Routledge.

Crutzen, B. S., Swank, O. H., & Visser, B. (2013). Confidence management: on interpersonal comparisons in teams. *Journal of Economics & Management Strategy*, *22*(4), 744-767.

Dietz, J., Joshi, C., Esses, V. M., Hamilton, L. K., & Gabarrot, F. (2015). The skill paradox: Explaining and reducing employment discrimination against skilled immigrants. *The International Journal of Human Resource Management*, *26*(10), 1318-1334.

Dymski, G. A. (1995). The theory of bank redlining and discrimination: An exploration. *The Review of Black Political Economy*, *23*(3), 37-74.

Dymski, G. A. (2006). Discrimination in the credit and housing markets: findings and challenges. *Handbook on the Economics of Discrimination*, *215*, 220.

Ghumman, S., Ryan, A. M., Barclay, L. A., & Markel, K. S. (2013). Religious discrimination in the workplace: A review and examination of current and future trends. *Journal of Business and Psychology*, *28*(4), 439-454.

Glassdoor. (2019, augustus). *Diversity & Inclusion Study 2019*. https://www.glassdoor.com/about-us/app/uploads/sites/2/2019/10/Glassdoor-Diversity-Survey-Supplement-1.pdf

Han, S. (2001). On the Economics of Discrimination in Credit Markets. *Available at SSRN 298356*.

Han, S. (2004). Discrimination in lending: Theory and evidence. *The Journal of Real Estate Finance and Economics*, *29*(1), 5-46.

Ishida, J. (2006). Optimal Promotion Policies with the Looking-Glass Effect. *Journal of Labor Economics*, *24*(4), 857–877.

Jones, K. P., Arena, D. F., Nittrouer, C. L., Alonso, N. M., & Lindsey, A. P. (2017). Subtle discrimination in the workplace: A vicious cycle. *Industrial and Organizational Psychology*, *10*(1), 51-76.

Kamphorst, J. J., & Swank, O. H. (2016). Don’t demotivate, discriminate. *American Economic Journal: Microeconomics*, *8*(1), 140-65.

Lee, C. H., & Warren, E. H. (1976). Rationing by seller's preference and pacial price discrimination. *Economic Inquiry*, *14*(1), 36-44.

Masson, R. T. (1973). Costs of search and racial price discrimination. *Economic Inquiry*, *11*(2), 167-86.

Ozeren, E. (2014). Sexual orientation discrimination in the workplace: A systematic review of literature. *Procedia-Social and Behavioral Sciences*, *109*, 1203-1215.

Prendergast, C. (1992). Career development and specific human capital collection. *Journal of the Japanese and international Economies*, *6*(3), 207-227.

Phelps, E. S. (1972). The statistical theory of racism and sexism. *The American economic review*, *62*(4), 659-661.

Swank, O. H., & Visser, B. (2007). Motivating through delegating tasks or giving attention. *The Journal of Law, Economics, & Organization*, *23*(3), 731-742.

U.S. Equal Employment Opportunity Commission. (2021). *Charge Statistics (Charges filed with EEOC) FY 1997 Through FY 2021* [Dataset]. https://www.eeoc.gov/statistics/charge-statistics-charges-filed-eeoc-fy-1997-through-fy-2021

Van Laer, K., & Janssens, M. (2011). Ethnic minority professionals' experiences with subtle discrimination in the workplace. *Human Relations*, *64*(9), 1203-1227.

Vassilopoulou, J., Brabet, J., & Showunmi, V. (Eds.). (2019). Race Discrimination and Management of Ethnic Diversity and Migration at Work: European Countries' Perspectives. Emerald Group Publishing.

Wooten, L. P., & James, E. H. (2004). When firms fail to learn: The perpetuation of discrimination in the workplace. *Journal of Management Inquiry*, *13*(1), 23-33.

Yinger, J. (1975). A Model of Discrimination by Landlords.

Appendix

A) Analysis of the task assignment game using *Model (a)*

A.1) Optimal task assignment for the basic game

The manager is indifferent between assigning the major task to employee 1 and employee 2 if:

$$U_M(m=1) = U_M(m=2)$$

Filling in using *Equations (1)* and *(3)* yields:

$$\eta a_1 e_1 + a_2 e_2 + \rho_1 a_1 = a_1 e_1 + \eta a_2 e_2$$

Given the optimal effort levels $e_i = \eta E(a_i|m)$ and $e_i = E(a_i|m)$, this becomes:

$$\eta a_1 \eta E(a_1|m=1) + a_2 E(a_2|m=1) + \rho_1 a_1 = a_1 E(a_1|m=2) + \eta a_2 \eta E(a_2|m=2)$$

Rearranging gives:

$$[\eta^2 E(a_1|1) - E(a_1|2) + \rho_1]a_1 = [\eta^2 E(a_2|2) - E(a_2|1)]a_2$$

Resulting in:

$$a_1 = ta_2$$
$$where \quad t = \frac{\eta^2 E(a_2|2) - E(a_2|1)}{\eta^2 E(a_1|1) - E(a_1|2) + \rho_1}$$

A.2) Equilibria of the basic game

To determine the optimal task assignment strategy, the expected employee ability levels given the task assignment decision have to be determined. Employees update their beliefs according to Bayes' rule. First, consider the case where $t \in [0,1]$. Expected employee ability levels can then be calculated as follows:

	m = 1	**m = 2**
$\mathbf{a_1}$	$E(a_1\|1) = \frac{\int_0^1 \int_{ta_2}^1 a_1 da_1 da_2}{\int_0^1 \int_{ta_2}^1 da_1 da_2} = \frac{3-t^2}{6-3t}$	$E(a_1\|2) = \frac{\int_0^1 \int_0^{ta_2} a_1 da_1 da_2}{\int_0^1 \int_0^{ta_2} da_1 da_2} = \frac{1}{3}t$
$\mathbf{a_2}$	$E(a_2\|1) = \frac{\int_0^1 \int_{ta_2}^1 a_2 da_1 da_2}{\int_0^1 \int_{ta_2}^1 da_1 da_2} = \frac{3-2t}{6-3t}$	$E(a_2\|2) = \frac{\int_0^1 \int_0^{ta_2} a_2 da_1 da_2}{\int_0^1 \int_0^{ta_2} da_1 da_2} = \frac{2}{3}$

Filling in these beliefs into the expression for t found in *Section A.1* yields:

$$\Rightarrow t = \frac{E(a_2|2) - E(a_2|1)}{E(a_1|1) - E(a_1|2) + \rho_1}$$

$$\Leftrightarrow t = \frac{\frac{2}{3} - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t} - \frac{1}{3}t + \rho_1}$$

$$\Leftrightarrow t = \frac{\frac{2(2-t)}{3(2-t)} - \frac{3-2t}{3(2-t)}}{\frac{3-t^2}{3(2-t)} - \frac{(2-t)t}{3(2-t)} + \frac{3(2-t)\rho_1}{3(2-t)}}$$

$$\Leftrightarrow t = \frac{1}{3 - 2t + 6\rho_1 - 3\rho_1 t}$$

$$\Leftrightarrow 3t - 2t^2 + 6\rho_1 t - 3\rho_1 t^2 - 1 = 0$$

$$\Leftrightarrow (2 + 3\rho_1)t^2 + (-3 - 6\rho_1)t + 1 = 0$$

$$\Leftrightarrow \boldsymbol{t^* = \frac{\sqrt{36\rho_1^2 + 24\rho_1 + 1} + 6\rho_1 + 3}{6\rho_1 + 4}} \quad \textbf{\textit{or}} \quad \boldsymbol{t^* = \frac{-\sqrt{36\rho_1^2 + 24\rho_1 + 1} + 6\rho_1 + 3}{6\rho_1 + 4}}$$

On the other hand, in the case where <u>$t > 1$</u>, employees' beliefs about their level of ability are updated as follows:

	$\boldsymbol{m = 1}$	$\boldsymbol{m = 2}$
$\boldsymbol{a_1}$	$E(a_1\|1) = \frac{\int_0^1 \int_0^{\frac{a_1}{t}} a_1 da_2 da_1}{\int_0^1 \int_0^{\frac{a_1}{t}} da_2 da_1} = \frac{2}{3}$	$E(a_1\|2) = \frac{\int_0^1 \int_{\frac{a_1}{t}}^1 a_1 da_2 da_1}{\int_0^1 \int_{\frac{a_1}{t}}^1 da_2 da_1} = \frac{2-3t}{3-6t}$
$\boldsymbol{a_2}$	$E(a_2\|1) = \frac{\int_0^1 \int_0^{\frac{a_1}{t}} a_2 da_2 da_1}{\int_0^1 \int_0^{\frac{a_1}{t}} da_2 da_1} = \frac{1}{3t}$	$E(a_2\|2) = \frac{\int_0^1 \int_{\frac{a_1}{t}}^1 a_2 da_2 da_1}{\int_0^1 \int_{\frac{a_1}{t}}^1 da_2 da_1} = \frac{1-3t^2}{3t-6t^2}$

Filling in these beliefs into the expression for *t* found in *Section A.1* yields:

$$\Rightarrow t = \frac{E(a_2|2) - E(a_2|1)}{E(a_1|1) - E(a_1|2) + \rho_1}$$

$$\Leftrightarrow t = \frac{\frac{1-3t^2}{3t-6t^2} - \frac{1}{3t}}{\frac{2}{3} - \frac{2-3t}{3-6t} + \rho_1}$$

$$\Leftrightarrow t = \frac{\frac{1-3t^2}{3t(1-2t)} - \frac{1-2t}{3t(1-2t)}}{\frac{2(1-2t)}{3(1-2t)} - \frac{2-3t}{3(1-2t)} + \frac{3(1-2t)\rho_1}{3(1-2t)}}$$

$$\Leftrightarrow t = \frac{2t - 3t^2}{-t^2 + 3\rho_1 t - 6\rho_1 t^2}$$

$$\Leftrightarrow t^3 - 3\rho_1 t^2 + 6\rho_1 t^3 + 2t - 3t^2 = 0$$

$$\Leftrightarrow (1 + 6\rho_1)t^2 + (-3 - 3\rho_1)t + 2 = 0$$

$$\Leftrightarrow \boldsymbol{t^* = \frac{\sqrt{9\rho_1^2 - 30\rho_1 + 1} + 3\rho_1 + 3}{12\rho_1 + 2}} \quad \textbf{or} \quad \boldsymbol{t^* = \frac{-\sqrt{9\rho_1^2 - 30\rho_1 + 1} + 3\rho_1 + 3}{12\rho_1 + 2}}$$

A.3) Stability of the basic game (without bias)

<u>$t \in [0,1]$</u>

The beliefs about the task assignment probabilities held by the employees ($\hat{t}$) differ from the equilibrium beliefs (t). For instance, the equilibrium $t = 1$ is stable if the manager's optimal response for $\hat{t}$ close but smaller than 1 is even closer to 1. Put differently: $\hat{t} < t < 1$. Otherwise, if $t < \hat{t}$, the manager's optimal response is further away from 1 than anticipated by the employees. In other words, the manager discriminates in favour of employee 1 even more than anticipated.

First, calculate the best response of the manager given the employee expectations updated using Bayes' rule:

$$t = \frac{E(a_2|2) - E(a_2|1)}{E(a_1|1) - E(a_1|2) + \rho_1} = \frac{\frac{2}{3} - \frac{3 - 2\hat{t}}{6 - 3\hat{t}}}{\frac{3 - \hat{t}^2}{6 - 3\hat{t}} - \frac{1}{3}\hat{t} + \rho_1} = \frac{1}{6\rho_1 - 3\rho_1\hat{t} - 2\hat{t} + 3}$$

Then, check when $t > \hat{t}$ (in other words, when is the equilibrium stable?):

$$\frac{1}{6\rho_1 - 3\rho_1\hat{t} - 2\hat{t} + 3} > \hat{t}$$

$$3\rho_1\hat{t}^2 + 2\hat{t}^2 - 6\rho_1\hat{t} - 3\hat{t} + 1 > 0$$

$$(3\rho_1 + 2)\hat{t}^2 - (6\rho_1 + 3)\hat{t} + 1 > 0$$

For $\rho_1 = 0$, the above only holds on the interval $\hat{t} \in \left[0, \frac{1}{2}\right[$. In other words, whenever the employee beliefs about the task assignment decision ($\hat{t}$) lie in this interval, it is optimal for the manager to choose a t greater than $\hat{t}$. Therefore, the equilibrium t moves towards $t = \frac{1}{2}$. On the other hand, $t < \hat{t}$ whenever $\hat{t}$ lies in the interval $\left]\frac{1}{2}, 1\right[$. This implies that, whenever the employees hold a belief about the task assignment decision within this interval, the equilibrium manager task assignment strategy t moves towards $t = \frac{1}{2}$. Combining both the cases above, it can be concluded that the discriminatory equilibrium $t = \frac{1}{2}$ is stable while the non-discriminatory equilibrium $t = 1$ is unstable.

$\underline{t > 1}$

The beliefs about the task assignment probabilities held by the employees ($\hat{t}$) differ from the equilibrium beliefs (t). For instance, the equilibrium $t = 1$ is stable if the manager's optimal response for $\hat{t}$ close but greater than 1 is even closer to 1. Put differently: $1 < t < \hat{t}$. Otherwise, if $t > \hat{t}$, the manager's optimal response is further away from 1 than anticipated by the employees. In other words, the manager discriminates in favour of employee 2 even more than anticipated.

First, calculate the best response of the manager given the employee expectations updated using Bayes' rule:

$$t = \frac{E(a_2|2) - E(a_2|1)}{E(a_1|1) - E(a_1|2) + \rho_1} = \frac{\frac{1 - 3\hat{t}^2}{3\hat{t} - 6\hat{t}^2} - \frac{1}{3\hat{t}}}{\frac{2}{3} - \frac{2 - 3\hat{t}}{3 - 6\hat{t}} + \rho_1} = \frac{3\hat{t} - 2}{6\rho_1\hat{t} - 3\rho_1 + \hat{t}}$$

Then, check when $t < \hat{t}$ (in other words, when is the equilibrium stable?):

$$\frac{3\hat{t} - 2}{6\rho_1\hat{t} - 3\rho_1 + \hat{t}} < \hat{t}$$

$$6\rho_1\hat{t}^2 + \hat{t}^2 - 3\rho_1\hat{t} - 3\hat{t} + 2 > 0$$

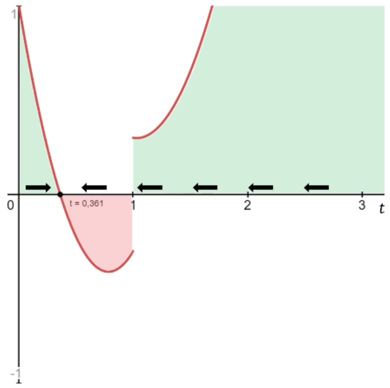
$$(6\rho_1 + 1)\hat{t}^2 - (3\rho_1 + 3)\hat{t} + 2 > 0$$

For $\rho_1 = 0$, the above only holds on the interval $\hat{t} \in]2, \infty[$. In other words, whenever the employee beliefs about the task assignment decision ($\hat{t}$) lie in this interval, it is optimal for the manager to choose a t smaller than $\hat{t}$. Therefore, the equilibrium t moves towards $t = 2$. On the other hand, $t > \hat{t}$ whenever $\hat{t}$ lies in the interval $]1,2[$. This implies that, whenever the employees hold a belief about the task assignment decision within this interval, the equilibrium manager task assignment strategy t moves towards $t = 2$. Combining both the cases above, it can be concluded that the discriminatory equilibrium $t = 2$ is stable while the non-discriminatory equilibrium $t = 1$ is unstable.

A.4) Stability of the basic game (with employee 1 bias)

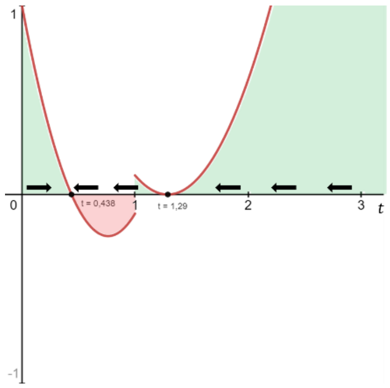
The stability of the equilibria in the game with bias will be determined by examining six separate cases with varying levels of bias. The reasoning is that different levels of bias determine the number of equilibria in the game and which are stable. In what follows, each case will be briefly discussed and graphically illustrated to make the interpretation more convenient.

<u>Case 1</u>: $\rho_1 > \frac{5}{3} - \frac{2\sqrt{2}}{\sqrt{3}}$ (= 0,0337)



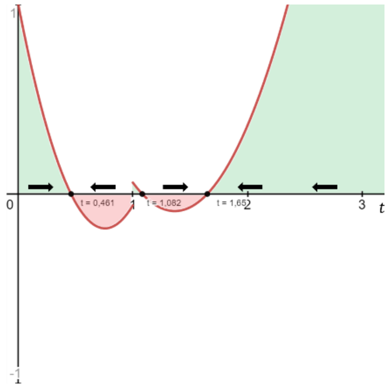
With a high positive bias towards employee 1 (in this case $\rho_1 = 0{,}1$) there exists only one equilibrium: $t = 0{,}361$. A high positive bias towards employee 1 leads to a unique discriminatory equilibrium in favour of that employee. To see this, note that for $t \in [0{,}1]$, the area indicated in green implies that $t > \hat{t}$ while the area in red implies that $t < \hat{t}$. Thus, the graph to the right shows that the optimal task assignment strategy converges to the single discriminatory equilibrium. For this level of bias, there is always one stable discriminatory equilibrium.

<u>Case 2</u>: $\rho_1 = \frac{5}{3} - \frac{2\sqrt{2}}{\sqrt{3}}$ (= 0,0337)



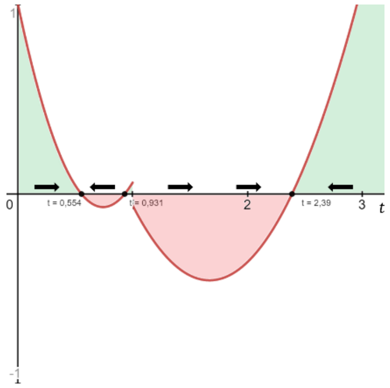
At the maximum positive bias towards employee 1 for which an equilibrium still exists with discrimination towards employee 2, there exist two equilibria: $t = 0{,}438$ and $t = 1{,}29$. The graph on the right-hand side shows that the most discriminatory equilibrium ($t = 0{,}438$) is stable because the optimal task assignment strategy converges towards it regardless of the employee beliefs $\hat{t}$. However, note that whenever the employees hold a belief on the interval $]1; 1{,}29[$, it is optimal for the manager to converge to the equilibrium in which he/she discriminates in the opposite direction ($t = 0{,}438$). The equilibrium $t = 1{,}29$ is unstable. For this level of bias, there is only one stable discriminatory equilibrium.

<u>Case 3</u>: $0 < \rho_1 < \frac{5}{3} - \frac{2\sqrt{2}}{\sqrt{3}}$ (= 0,0337)



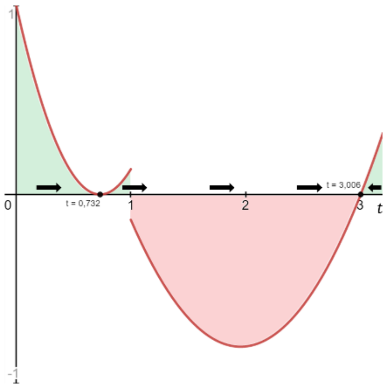
With a slight positive bias towards employee 1 (in this case $\rho_1 = 0{,}02$), there exist three equilibria: $t = 0{,}461$, $t = 1{,}082$ and $t = 1{,}65$. Note that for $t > 1$, the red area shows the values for the employee beliefs $\hat{t}$ for which $t > \hat{t}$. The graph shows how the optimal level of t converges toward the most discriminatory equilibria. The equilibrium in the middle ($t = 1{,}082$), although slightly discriminatory, is unstable. For this level of bias, out of three discriminatory equilibria, only the most discriminatory ones are stable.

<u>Case 4</u>: $0 > \rho_1 > \frac{\sqrt{3}}{6} - \frac{1}{3}$ (= −0,0447)



With a slight negative bias towards employee 1 (in this case $\rho_1 = -0{,}02$), there exist three equilibria: $t = 0{,}554$, $t = 0{,}931$ and $t = 2{,}39$. Similarly to case 3, the optimal level of t converges towards the most discriminatory equilibria. The equilibrium in the middle ($t = 0{,}931$), although slightly discriminatory, is unstable. For this level of bias, out of three discriminatory equilibria, only the most discriminatory ones are stable.

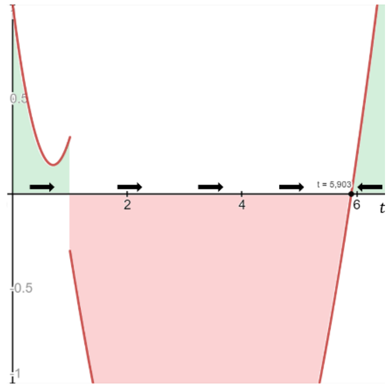
<u>Case 5</u>: $\rho_1 = \frac{\sqrt{3}}{6} - \frac{1}{3}$ (= −0,0447)



At the maximum negative bias towards employee 1 for which an equilibrium still exists with discrimination in favour of employee 1, there exist two equilibria: $t = 0{,}732$ and $t = 3{,}01$. The graph on the right-hand side shows that the most discriminatory equilibrium ($t = 3{,}01$) is stable because the optimal task assignment strategy converges towards it regardless of the employee beliefs $\hat{t}$. However, note that whenever the employees hold a belief on the interval $]0{,}732; 1[$, it is optimal for the manager to converge to the equilibrium in which he/she discriminates in the opposite direction ($t = 3{,}01$). The equilibrium $t = 0{,}732$ is unstable. For this level of bias, there is only one stable discriminatory equilibrium.

<u>Case 6</u>: $\rho_1 < \frac{\sqrt{3}}{6} - \frac{1}{3} (= -0{,}0447)$

With a high negative bias towards employee 1 (in this case $\rho_1 = -0{,}1$) there exists only one equilibrium: $t = 5{,}903$. A high negative bias towards employee 1 leads to a discriminatory equilibrium favouring the other employee. The graph to the right shows that the optimal task assignment strategy converges to the single discriminatory equilibrium. For this level of bias, there is always one stable discriminatory equilibrium.



A.5) Efficiency of the basic game

<u>Manager payoff:</u>

The manager's payoff is equal to the sum of the utilities resulting from the two possible task assignment decisions. Note that the manager's utility is increasing in both the total level of production and the bias towards employee 1.

For $t \in [0,1]$:

$$\underbrace{\int_0^1 \int_{ta_2}^1 \left(a_1 \frac{3-t^2}{6-3t} + a_2 \frac{3-2t}{6-3t} + a_1 \rho_1 \right) da_1 da_2}_{m=1} + \underbrace{\int_0^1 \int_0^{ta_2} \left(a_1 \frac{t}{3} + a_2 \frac{2}{3} \right) da_1 da_2}_{m=2} = \frac{2t^3 + 3\rho t^3 - 6t^2 - 6\rho t^2 - 4t - 9\rho t + 18 + 18\rho}{18(2-t)}$$

For $t > 1$:

$$\underbrace{\int_0^1 \int_0^{\frac{a_1}{t}} \left(a_1 \frac{2}{3} + a_2 \frac{1}{3t} + a_1 \rho \right) da_2 da_1}_{m=1} + \underbrace{\int_0^1 \int_{\frac{a_1}{t}}^1 \left(a_1 \frac{2-3t}{3-6t} + a_2 \frac{1-3t^2}{3t-6t^2} \right) da_2 da_1}_{m=2} = \frac{9t^3 - 2t^2 + 6\rho t^2 - 3t - 3\rho t + 1}{9t^2(2t-1)}$$

Combining the two functions above leads to three general insights:

1. Without a bias, the manager's payoff is maximised whenever $t = 1$.
2. An increasing positive bias moves the optimal level of t closer and closer to zero. The manager's payoff is maximised with more discrimination in favour of employee 1 and is higher.
3. An increasing negative bias moves the optimal level of t further and further away from 1 ($t > 1$). The manager's payoff is maximised with more discrimination in favour of employee 2. However, the payoff is lower.

The reason why a positive bias increases the payoff at the optimal level of t is that the bias enters the manager's utility function as a positive term. On the other hand, whenever the bias is negative, utility logically decreases because there is still a possibility that employee 1 is assigned to the major task regardless of the negative bias against him/her.

Employee 1 payoff

Employee 1's payoff is equal to the sum of the utilities resulting from the two possible task assignment decisions.

For $t \in [0,1]$:

$$\underbrace{\int_0^1 \int_{ta_2}^1 \left(a_1 \frac{3-t^2}{6-3t} - \frac{1}{2}\left(\frac{3-t^2}{6-3t}\right)^2 \right) da_1 da_2}_{m=1} + \underbrace{\int_0^1 \int_0^{ta_2} \left(a_1 \frac{t}{3} - \frac{1}{2}\left(\frac{t}{3}\right)^2 \right) da_1 da_2}_{m=2} = \frac{2t^3 - 6t^2 + 9}{36(2-t)}$$

For $t > 1$:

$$\underbrace{\int_0^1 \int_0^{\frac{a_1}{t}} \left(a_1 \frac{2}{3} - \frac{1}{2}\left(\frac{2}{3}\right)^2 \right) da_2 da_1}_{m=1} + \underbrace{\int_0^1 \int_{\frac{a_1}{t}}^1 \left(a_1 \frac{2-3t}{3-6t} - \frac{1}{2}\left(\frac{2-3t}{3-6t}\right)^2 \right) da_2 da_1}_{m=2} = \frac{9t-4}{36(2t-1)}$$

Combining the two functions above leads to two general insights:

1. The bias only affects the payoff of employee 1 indirectly through the optimal task assignment decision t.
2. Employee 1's payoff is maximised at $t = \frac{3}{2} - \frac{\sqrt{3}}{2} \approx 0{,}634$. Employee 1 prefers some level of discrimination in his/her favour. To reach this equilibrium, the manager would need to have a slight negative bias ($\rho_1 = -0{,}0377$) against employee 1.

Employee 2 payoff

Employee 2's payoff is equal to the sum of the utilities resulting from the two possible task assignment decisions.

For $t \in [0,1]$:

$$\underbrace{\int_0^1 \int_{ta_2}^1 \left(a_2 \frac{3-2t}{6-3t} - \frac{1}{2}\left(\frac{3-2t}{6-3t}\right)^2 \right) da_1 da_2}_{m=1} + \underbrace{\int_0^1 \int_0^{ta_2} \left(a_2 \frac{2}{3} - \frac{1}{2}\left(\frac{2}{3}\right)^2 \right) da_1 da_2}_{m=2} = \frac{9-4t}{36(2-t)}$$

For $t > 1$:

$$\underbrace{\int_0^1 \int_0^{\frac{a_1}{t}} \left(a_2 \frac{1}{3t} - \frac{1}{2}\left(\frac{1}{3t}\right)^2 \right) da_2 da_1}_{m=1} + \underbrace{\int_0^1 \int_{\frac{a_1}{t}}^1 \left(a_2 \frac{1-3t^2}{3t-6t^2} - \frac{1}{2}\left(\frac{1-3t^2}{3t-6t^2}\right)^2 \right) da_2 da_1}_{m=2} = \frac{9t^3 - 6t + 2}{36t^2(2t-1)}$$

Combining the two functions above leads to two general insights:

1. The bias only affects the payoff of employee 2 indirectly through the optimal task assignment decision t.

2. Employee 2's payoff is maximised at $t = 1 + \frac{1}{\sqrt{3}} \approx 1{,}577$. Employee 2 prefers some level of discrimination in his/her favour. To reach this equilibrium, the manager would need to have a slight positive bias ($\rho_1 = 0{,}024$) towards employee 1.

Combined employee payoff

For $t \in [0,1]$:

$$\underbrace{\frac{2t^3 - 6t^2 + 9}{36(2-t)}}_{Employee\ 1} + \underbrace{\frac{9 - 4t}{36(2-t)}}_{Employee\ 2} = \frac{t^3 - 3t^2 - 2t + 9}{36 - 18t}$$

For $t > 1$:

$$\underbrace{\frac{9t - 4}{36(2t-1)}}_{Employee\ 1} + \underbrace{\frac{9t^3 - 6t + 2}{36t^2(2t-1)}}_{Employee\ 2} = \frac{9t^3 - 2t^2 - 3t + 1}{36t^3 - 18t^2}$$

The combined utility for both employees is maximised only when $t = 1$. Even though the employees individually prefer some discrimination in their own favour, the combined payoff is maximised without discrimination.

Firm payoff:

The firm payoff is equal the manager payoff without bias. The only goal for the firm is to maximise the output generated by the two employees.

For $t \in [0,1]$:

$$\underbrace{\int_0^1 \int_{ta_2}^1 \left(a_1 \frac{3 - t^2}{6 - 3t} + a_2 \frac{3 - 2t}{6 - 3t}\right) da_1 da_2}_{m=1} + \underbrace{\int_0^1 \int_0^{ta_2} \left(a_1 \frac{t}{3} + a_2 \frac{2}{3}\right) da_1 da_2}_{m=2} = \frac{t^3 - 3t^2 - 2t + 9}{9(2-t)}$$

For $t > 1$:

$$\underbrace{\int_0^1 \int_0^{\frac{a_1}{t}} \left(a_1 \frac{2}{3} + a_2 \frac{1}{3t}\right) da_2 da_1}_{m=1} + \underbrace{\int_0^1 \int_{\frac{a_1}{t}}^1 \left(a_1 \frac{2 - 3t}{3 - 6t} + a_2 \frac{1 - 3t^2}{3t - 6t^2}\right) da_2 da_1}_{m=2} = \frac{9t^3 - 2t^2 - 3t + 1}{9t^2(2t-1)}$$

The payoff for the firm is maximised only when $t = 1$. The manager prefers some level of discrimination depending on whether he/she has a positive or negative bias towards employee 1. However, from a productivity point of view, it would be optimal for the manager to be unbiased and not to discriminate.

A.6) Equilibria of the extended game

To determine the optimal task assignment strategy, the expected employee ability levels given the task assignment decision have to be determined. The approach will be largely similar to the one in *Section A.2*. However, the task importance parameter η will now be equal to or greater than 1. Again, employees update their beliefs according to Bayes' rule. First, consider the case where $t \in [0,1]$. Expected employee ability levels can then be calculated the same as before:

	$\mathbf{m = 1}$	$\mathbf{m = 2}$
$\mathbf{a_1}$	$E(a_1\|1) = \frac{\int_0^1 \int_{ta_2}^1 a_1 da_1 da_2}{\int_0^1 \int_{ta_2}^1 da_1 da_2} = \frac{3 - t^2}{6 - 3t}$	$E(a_1\|2) = \frac{\int_0^1 \int_0^{ta_2} a_1 da_1 da_2}{\int_0^1 \int_0^{ta_2} da_1 da_2} = \frac{1}{3}t$
$\mathbf{a_2}$	$E(a_2\|1) = \frac{\int_0^1 \int_{ta_2}^1 a_2 da_1 da_2}{\int_0^1 \int_{ta_2}^1 da_1 da_2} = \frac{3 - 2t}{6 - 3t}$	$E(a_2\|2) = \frac{\int_0^1 \int_0^{ta_2} a_2 da_1 da_2}{\int_0^1 \int_0^{ta_2} da_1 da_2} = \frac{2}{3}$

Filling in these beliefs into the expression for t found in section A.1 yields:

$$\Rightarrow t = \frac{\eta^2 E(a_2|2) - E(a_2|1)}{\eta^2 E(a_1|1) - E(a_1|2) + \rho_1}$$

$$\Leftrightarrow t = \frac{\frac{2}{3}\eta^2 - \frac{3 - 2t}{6 - 3t}}{\frac{3 - t^2}{6 - 3t}\eta^2 - \frac{1}{3}t + \rho_1}$$

$$\Leftrightarrow t = \frac{\frac{2(2 - t)\eta^2}{3(2 - t)} - \frac{3 - 2t}{3(2 - t)}}{\frac{(3 - t^2)\eta^2}{3(2 - t)} - \frac{(2 - t)t}{3(2 - t)} + \frac{3(2 - t)\rho_1}{3(2 - t)}}$$

$$\Leftrightarrow t = \frac{4\eta^2 - 2\eta^2 t - 3 + 2t}{3\eta^2 - \eta^2 t^2 - 2t + t^2 + 6\rho_1 - 3\rho_1 t}$$

$$\boxed{\Leftrightarrow (\boldsymbol{\eta}^2 - 1)\boldsymbol{t}^3 + (2 + 3\boldsymbol{\rho}_1)\boldsymbol{t}^2 + (2 - 5\boldsymbol{\eta}^2 - 6\boldsymbol{\rho}_1)\boldsymbol{t} + 4\boldsymbol{\eta}^2 - 3 = 0}$$

On the other hand, in the case where $t > 1$, employees' beliefs about their level of ability are updated as follows:

	$\boldsymbol{m = 1}$	$\boldsymbol{m = 2}$
$\boldsymbol{a_1}$	$E(a_1\|1) = \frac{\int_0^1 \int_0^{\frac{a_1}{t}} a_1 da_2 da_1}{\int_0^1 \int_0^{\frac{a_1}{t}} da_2 da_1} = \frac{2}{3}$	$E(a_1\|2) = \frac{\int_0^1 \int_{\frac{a_1}{t}}^1 a_1 da_2 da_1}{\int_0^1 \int_{\frac{a_1}{t}}^1 da_2 da_1} = \frac{2 - 3t}{3 - 6t}$
$\boldsymbol{a_2}$	$E(a_2\|1) = \frac{\int_0^1 \int_0^{\frac{a_1}{t}} a_2 da_2 da_1}{\int_0^1 \int_0^{\frac{a_1}{t}} da_2 da_1} = \frac{1}{3t}$	$E(a_2\|2) = \frac{\int_0^1 \int_{\frac{a_1}{t}}^1 a_2 da_2 da_1}{\int_0^1 \int_{\frac{a_1}{t}}^1 da_2 da_1} = \frac{1 - 3t^2}{3t - 6t^2}$

Filling in these beliefs into the expression for t found in section A.1 yields:

$$\Rightarrow t = \frac{\eta^2 E(a_2|2) - E(a_2|1)}{\eta^2 E(a_1|1) - E(a_1|2) + \rho_1}$$

$$\Leftrightarrow t = \frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t} + \rho_1}$$

$$\Leftrightarrow t = \frac{\frac{(1-3t^2)\eta^2}{3t(1-2t)} - \frac{1-2t}{3t(1-2t)}}{\frac{2(1-2t)\eta^2}{3(1-2t)} - \frac{2-3t}{3(1-2t)} + \frac{3(1-2t)\rho_1}{3(1-2t)}}$$

$$\Leftrightarrow t = \frac{\eta^2 - 3\eta^2 t^2 - 1 + 2t}{2\eta^2 t - 4\eta^2 t^2 - 2t + 3t^2 + 3\rho_1 t - 6\rho_1 t^2}$$

$$\boxed{\Leftrightarrow \boldsymbol{(3 - 4\eta^2 - 6\rho_1)t^3 - (2 - 5\eta^2 - 3\rho_1)t^2 - 2t - \eta^2 + 1 = 0}}$$

<u>Case 1</u>: No manager bias ($\boldsymbol{\rho_1 = 0}$)

Without a manager bias towards employee 1, the equations found above can be solved for t as follows:

For $t \in [0,1]$:

$$t^* = \frac{\eta^2 - \sqrt{17\eta^4 - 26\eta^2 + 13} + 1}{2 - 2\eta^2} \qquad or \qquad t^* = 1$$

For $t > 1$:

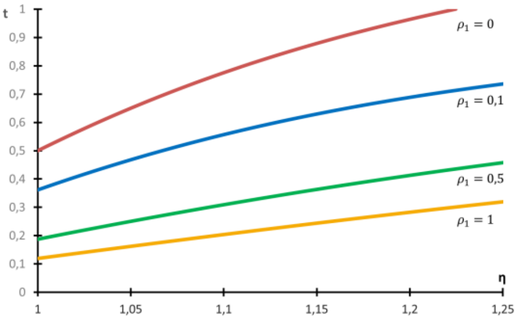
$$t^* = 1 \qquad or \qquad t^* = \frac{\eta^2 + \sqrt{17\eta^4 - 26\eta^2 + 13} + 1}{8\eta^2 - 6}$$

In other words, there are two discriminatory equilibria and one non-discriminatory equilibrium. However, whenever $\eta \geq \sqrt{1{,}5}$, $t^* = 1$ becomes the unique equilibrium.

<u>Case 2</u>: Positive manager bias ($\boldsymbol{\rho_1 > 0}$)

The following cases in which $\rho_1 \neq 0$ will be solved graphically as the third power in the equation makes it unnecessarily difficult to solve them algebraically. However, the most important results can be smoothly derived this way.

Plugging in different values for the bias makes it possible to graphically depict the relationship between the optimal task assignment strategy t and the relative major task importance η. For $t \in [0,1]$, the graph to the right shows how an increasing bias requires a higher level of η to move closer to the non-discriminatory equilibrium. Additionally, reaching the non-discriminatory equilibrium becomes impossible with any level of bias.



Algebraically:

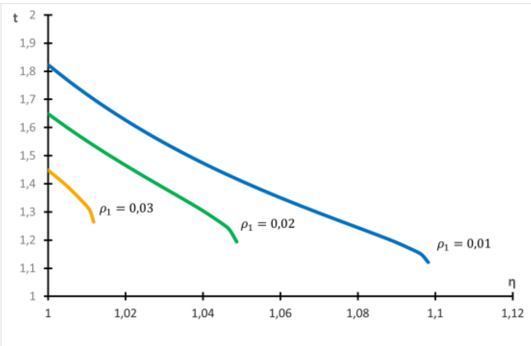
$(\eta^2 - 1)t^3 + (2 + 3\rho_1)t^2 + (2 - 5\eta^2 - 6\rho_1)t + 4\eta^2 - 3 = 0$

$\Leftrightarrow \eta^2 - 1 + 2 + 3\rho_1 + 2 - 5\eta^2 - 6\rho_1 + 4\eta^2 - 3 = 0$ whenever $t = 1$

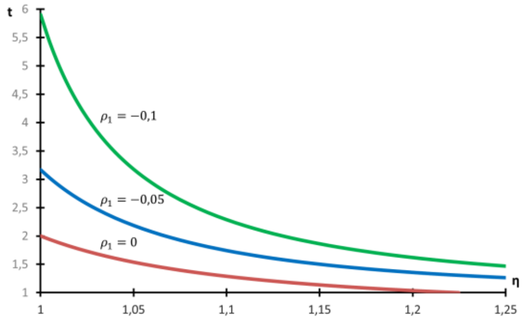
$\Leftrightarrow -3\rho_1 = 0$

The equality above can only hold if $\rho_1 = 0$. Any level of bias means that $t = 1$ cannot be an equilibrium regardless of the major task importance.

On the other hand, for $t > 1$, a positive bias towards employee 1 also moves the equilibrium task assignment decision t closer to 1 (no discrimination). However, for every level of bias, there is a maximum level of major task importance after which an equilibrium with $t > 1$ ceases to exist. This is because there are two forces at play in determining the equilibria: (1) the bias causes the manager to want to discriminate in favour of employee 1, while (2) the major task importance causes the manager to discriminate less. This is why an increasing positive bias combined with a high level of major task importance makes it impossible for an equilibrium to exist with discrimination in favour of employee 2.



<u>Case 3</u>: Negative manager bias ($\rho_1 < 0$)



Plugging in different values for the bias makes it possible to graphically depict the relationship between the optimal task assignment strategy t and the relative major task importance η. For $t > 1$, the graph to the right shows how an increasing bias requires a higher level of η to move closer to the non-discriminatory equilibrium. Additionally, reaching the non-discriminatory equilibrium becomes impossible with any level of bias.

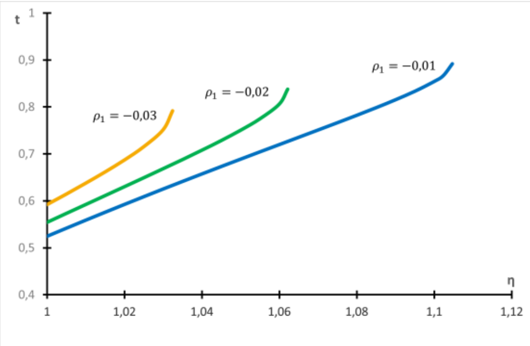
Algebraically:

$$(3 - 4\eta^2 - 6\rho_1)t^3 - (2 - 5\eta^2 - 3\rho_1)t^2 - 2t - \eta^2 + 1 = 0$$

$$\Leftrightarrow 3 - 4\eta^2 - 6\rho_1 - 2 + 5\eta^2 + 3\rho_1 - 2 - \eta^2 + 1 = 0 \qquad \text{whenever } t = 1$$

$$\Leftrightarrow -3\rho_1 = 0$$

The equality above can only hold if $\rho_1 = 0$. Any level of bias means that $t = 1$ cannot be an equilibrium regardless of the major task importance.



On the other hand, for $t \in [0,1]$, a negative bias towards employee 1 also moves the equilibrium task assignment decision t closer to 1 (no discrimination). However, for every level of bias, there is a maximum level of major task importance after which an equilibrium for $t \in [0,1]$ ceases to exist. This is because here too, there are two forces at play in determining the equilibria: (1) the negative bias causes the manager to want to discriminate in favour of employee 2, while (2) the major task importance causes the manager to discriminate less. This is why an increasing positive bias combined with a high level of major task importance makes it impossible for an equilibrium to exist with discrimination in favour of employee 1.

A.7) Stability of the extended game

<u>Stability for $t \in [0,1]$</u> ($\rho_1 = 0$)

Just like in the basic model, the equilibrium values for t are found as follows:

$$t = \frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t}$$

In order to check the stability of the equilibria, the following rule is used:

The equilibrium is stable if:

$$\frac{\partial}{\partial t}t > \frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} \Leftrightarrow \frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} < 1$$

The equilibrium is unstable if:

$$\frac{\partial}{\partial t}t < \frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} \Leftrightarrow \frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} > 1$$

Solving the derivative gives:

$$\frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} = \frac{-2[\eta^4(t^2-4t+3) + \eta^2(-2t^2+7t-7) + t^2 - 3t + 3]}{[(t-2)t - \eta^2(t^2-3)]^2}$$

<u>For $t = 1$</u>, this term becomes:

$$\frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} = \frac{-2[-2\eta^2+1]}{[2\eta^2-1]^2} = \frac{2[2\eta^2-1]}{[2\eta^2-1]^2} = \frac{2}{2\eta^2-1}$$

The equilibrium is stable if:

$$\frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} < 1 \Leftrightarrow \frac{2}{2\eta^2-1} < 1 \Leftrightarrow 2 < 2\eta^2 - 1 \Leftrightarrow \eta > \sqrt{\frac{3}{2}}$$

The equilibrium is unstable if:

$$\frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} > 1 \Leftrightarrow \frac{2}{2\eta^2-1} > 1 \Leftrightarrow 2 > 2\eta^2 - 1 \Leftrightarrow \eta < \sqrt{\frac{3}{2}}$$

In other words, this equilibrium is only stable if the discriminatory equilibrium does not exist, as $\eta = \sqrt{1.5}$ was the cutoff major task importance level after which there was no discrimination. The equilibrium without discrimination ($t = 1$) is only stable if the major task has a high enough relative importance compared to the minor one. If this is not the case, $t = 1$ is an unstable equilibrium. The non-discriminatory equilibrium is only stable if it is the unique equilibrium.

For $t = t^*$, first re-write t:

$$t = \frac{\eta^2 E(a_2|2) - E(a_2|1)}{\eta^2 E(a_1|1) - E(a_1|2)}$$
$$\Leftrightarrow t[\eta^2 E(a_1|1) - E(a_1|2)] = \eta^2 E(a_2|2) - E(a_2|1)$$
$$\Leftrightarrow t\eta^2 E(a_1|1) - \eta^2 E(a_2|2) = tE(a_1|2) - E(a_2|1)$$
$$\Leftrightarrow \eta = \sqrt{\frac{tE(a_1|2) - E(a_2|1)}{tE(a_1|1) - E(a_2|2)}} = \sqrt{\frac{\frac{1}{3}t^2 - \frac{3-2t}{6-3t}}{\frac{3t-t^3}{6-3t} - \frac{2}{3}}} = \sqrt{\frac{t^2 - t - 3}{t^2 + t - 4}}$$

Filling in η into the derivative found previously gives:

$$\frac{\partial}{\partial t} \frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t}$$
$$= \frac{-2\left[\left(\sqrt{\frac{t^2-t-3}{t^2+t-4}}\right)^4 (t^2 - 4t + 3) + \left(\sqrt{\frac{t^2-t-3}{t^2+t-4}}\right)^2 (-2t^2 + 7t - 7) + t^2 - 3t + 3\right]}{\left[(t-2)t - \left(\sqrt{\frac{t^2-t-3}{t^2+t-4}}\right)^2 (t^2 - 3)\right]^2}$$
$$= \frac{2t^3 - 4t^2 + 4t + 2}{9 - 5t}$$

If $t = t^*$, this becomes:

$$\frac{\partial}{\partial t} \frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} = \frac{2(t^*)^3 - 4(t^*)^2 + 4t^* + 2}{9 - 5t^*}$$

The equilibrium is stable if:

$$\frac{\partial}{\partial t} \frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} < 1$$

$$\Leftrightarrow \frac{2(t^*)^3 - 4(t^*)^2 + 4t^* + 2}{9 - 5t^*} < 1$$

$$\Leftrightarrow 2(t^*)^3 - 4(t^*)^2 + 4t^* + 2 < 9 - 5t^*$$

$$\Leftrightarrow 2(t^*)^3 - 4(t^*)^2 + 9t^* - 7 < 0$$

The inequality above holds for any t^* between 0 and 1. In other words, $t = t^*$ is a stable equilibrium. Again, this equilibrium is a discriminatory equilibrium. Taking everything together: if the major task importance is not high enough, the only stable equilibrium is the discriminatory one. Once the major task importance exceeds the cutoff level of $\sqrt{1.5}$, the non-discriminatory equilibrium becomes unique and stable.

Stability for $t > 1$ ($\rho_1 = 0$)

Just like in the basic model, the equilibrium values for t are found as follows:

$$t = \frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t}}$$

In order to check the stability of the equilibria, the following rule is used:

The equilibrium is stable if:

$$\frac{\partial}{\partial t}t > \frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t}} \Leftrightarrow \frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t}} < 1$$

The equilibrium is unstable if:

$$\frac{\partial}{\partial t}t < \frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t}} \Leftrightarrow \frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t}} > 1$$

Solving the derivative gives:

$$\frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t}} = \frac{-2[\eta^4(3t^2 - 4t + 1) + \eta^2(-7t^2 + 7t - 2) + 3t^2 - 3t + 1]}{t^2[\eta^2(4t-2) - 3t + 2]^2}$$

<u>For $t = 1$</u>, this term becomes:

$$\frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t}} = \frac{-2(-2\eta^2 + 1)}{[2\eta^2 - 1]^2} = \frac{2(2\eta^2 - 1)}{[2\eta^2 - 1]^2} = \frac{2}{2\eta^2 - 1}$$

The equilibrium is stable if:

$$\frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2-\frac{1}{3t}}{\frac{2}{3}\eta^2-\frac{2-3t}{3-6t}} < 1 \Leftrightarrow \frac{2}{2\eta^2-1} < 1 \Leftrightarrow 2 < 2\eta^2-1 \Leftrightarrow \eta > \sqrt{\frac{3}{2}}$$

The equilibrium is unstable if:

$$\frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2-\frac{1}{3t}}{\frac{2}{3}\eta^2-\frac{2-3t}{3-6t}} > 1 \Leftrightarrow \frac{2}{2\eta^2-1} > 1 \Leftrightarrow 2 > 2\eta^2-1 \Leftrightarrow \eta < \sqrt{\frac{3}{2}}$$

Again, A non-discriminatory equilibrium is only stable if it is the unique equilibrium.

<u>For $t = t^*$</u>, first re-write t:

$$t = \frac{\eta^2 E(a_2|2) - E(a_2|1)}{\eta^2 E(a_1|1) - E(a_1|2)}$$

$$\Leftrightarrow t[\eta^2 E(a_1|1) - E(a_1|2)] = \eta^2 E(a_2|2) - E(a_2|1)$$

$$\Leftrightarrow t\eta^2 E(a_1|1) - \eta^2 E(a_2|2) = tE(a_1|2) - E(a_2|1)$$

$$\Leftrightarrow \eta = \sqrt{\frac{tE(a_1|2) - E(a_2|1)}{tE(a_1|1) - E(a_2|2)}} = \sqrt{\frac{\frac{2-3t}{3-6t}t - \frac{1}{3t}}{\frac{2}{3}t - \frac{1-3t^2}{3t-6t^2}}} = \sqrt{\frac{3t^2+t-1}{4t^2-t-1}}$$

Filling in η into the derivative found above gives:

$$\frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2-\frac{1}{3t}}{\frac{2}{3}\eta^2-\frac{2-3t}{3-6t}}$$

$$= \frac{-2\left[\left(\sqrt{\frac{3t^2+t-1}{4t^2-t-1}}\right)^4(3t^2-4t+1) + \left(\sqrt{\frac{3t^2+t-1}{4t^2-t-1}}\right)^2(-7t^2+7t-2) + 3t^2-3t+1\right]}{t^2\left[\left(\sqrt{\frac{3t^2+t-1}{4t^2-t-1}}\right)^2(4t-2) - 3t + 2\right]^2}$$

$$= \frac{2t^3+4t^2-4t+2}{9t^3-5t^2}$$

If $t = t^*$, this becomes:

$$\frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2-\frac{1}{3t}}{\frac{2}{3}\eta^2-\frac{2-3t}{3-6t}} = \frac{2(t^*)^3+4(t^*)^2-4t^*+2}{9(t^*)^3-5(t^*)^2}$$

The equilibrium is stable if:

$$\frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2-\frac{1}{3t}}{\frac{2}{3}\eta^2-\frac{2-3t}{3-6t}}<1$$

$$\Leftrightarrow \frac{2(t^*)^3+4(t^*)^2-4t^*+2}{9(t^*)^3-5(t^*)^2}<1$$

$$\Leftrightarrow 2(t^*)^3+4(t^*)^2-4t^*+2<9(t^*)^3-5(t^*)^2$$

$$\Leftrightarrow -7(t^*)^3+9(t^*)^2-4t^*+2<0$$

The inequality above holds for any t^* greater than 1. In other words, $t=t^*$ is a stable equilibrium. If the major task importance is not high enough, the only stable equilibrium is the discriminatory one. Once the major task importance exceeds the cutoff level of $\sqrt{1.5}$, the non-discriminatory equilibrium becomes unique and stable.

Stability for $t \in [0,1]$ ($\rho_1 \neq 0$)

Just like in the basic model, the equilibrium values for t are found as follows:

$$t=\frac{\frac{2}{3}\eta^2-\frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2-\frac{1}{3}t+\rho_1}$$

In order to check the stability of the equilibria, the following rule is used:

The equilibrium is stable if:

$$\frac{\partial}{\partial t}t>\frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2-\frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2-\frac{1}{3}t+\rho_1}\Leftrightarrow\frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2-\frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2-\frac{1}{3}t+\rho_1}<1$$

The equilibrium is unstable if:

$$\frac{\partial}{\partial t}t<\frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2-\frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2-\frac{1}{3}t+\rho_1}\Leftrightarrow\frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2-\frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2-\frac{1}{3}t+\rho_1}>1$$

Solving the derivative gives:

$$\frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2-\frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2-\frac{1}{3}t+\rho_1}=\frac{-2\eta^4(t^2-4t+3)+2\eta^2(2t^2-7t+7)+3\rho_1-2t^2+6t-6}{[\eta^2(t^2-3)+(t-2)(3\rho_1-t)]^2}$$

Every possible combination of η and ρ_1 leading to a value of t will result in a derivative smaller than 1. All possible equilibria in the game with bias are discriminatory and stable.

<u>Stability for $t > 1$</u> ($\rho_1 \neq 0$)

Just like in the basic model, the equilibrium values for t are found as follows:

$$t = \frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t} + \rho_1}$$

In order to check the stability of the equilibria, the following rule is used:

The equilibrium is stable if:

$$\frac{\partial}{\partial t}t > \frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t} + \rho_1} \Leftrightarrow \frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t} + \rho_1} < 1$$

The equilibrium is unstable if:

$$\frac{\partial}{\partial t}t < \frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t} + \rho_1} \Leftrightarrow \frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t} + \rho_1} > 1$$

Solving the derivative gives:

$$\frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t}}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t} + \rho_1} = \frac{\left[\eta^4(-6t^2+8t-2) + \eta^2(14t^2-14t+4) + 3\rho_1 t^2 - 6t^2 + 6t - 2\right]}{t^2\left[\eta^2(4t-2) + 6\rho_1 t - 3\rho_1 - 3t + 2\right]^2}$$

Every possible combination of η and ρ_1 leading to a value of t will result in a derivative smaller than 1. All possible equilibria in the game with bias are discriminatory and stable.

B) Analysis of the task assignment game using *Model (b)*

B.1) Optimal task assignment for the basic game

The manager is indifferent between assigning the major task to employee 1 and employee 2 if:

$$U_M(m=1) = U_M(m=2)$$

Filling in using *Equations (1)* and *(3')* yields:

$$\eta a_1 e_1 + a_2 e_2 = a_1 e_1 + \eta a_2 e_2 + \rho_2 a_2$$

Given the optimal effort levels $e_i = \eta E(a_i|m)$ and $e_i = E(a_i|m)$, this becomes:

$$\eta a_1 \eta E(a_1|m=1) + a_2 E(a_2|m=1) = a_1 E(a_1|m=2) + \eta a_2 \eta E(a_2|m=2) + \rho_2 a_2$$

Rearranging gives:

$$[\eta^2 E(a_1|1) - E(a_1|2)]a_1 = [\eta^2 E(a_2|2) - E(a_2|1) + \rho_2]a_2$$

Resulting in:

$$a_1 = ta_2$$
$$where \quad t = \frac{\eta^2 E(a_2|2) - E(a_2|1) + \rho_2}{\eta^2 E(a_1|1) - E(a_1|2)}$$

B.2) Equilibria of the basic game

To determine the optimal task assignment strategy, the expected employee ability levels given the task assignment decision have to be determined. Employees update their beliefs according to Bayes' rule. First, consider the case where $t \epsilon [0,1]$. Expected employee ability levels can then be calculated as follows:

	$\mathbf{m = 1}$	$\mathbf{m = 2}$
$\mathbf{a_1}$	$E(a_1\|1) = \frac{\int_0^1 \int_{ta_2}^1 a_1 da_1 da_2}{\int_0^1 \int_{ta_2}^1 da_1 da_2} = \frac{3 - t^2}{6 - 3t}$	$E(a_1\|2) = \frac{\int_0^1 \int_0^{ta_2} a_1 da_1 da_2}{\int_0^1 \int_0^{ta_2} da_1 da_2} = \frac{1}{3}t$
$\mathbf{a_2}$	$E(a_2\|1) = \frac{\int_0^1 \int_{ta_2}^1 a_2 da_1 da_2}{\int_0^1 \int_{ta_2}^1 da_1 da_2} = \frac{3 - 2t}{6 - 3t}$	$E(a_2\|2) = \frac{\int_0^1 \int_0^{ta_2} a_2 da_1 da_2}{\int_0^1 \int_0^{ta_2} da_1 da_2} = \frac{2}{3}$

Filling in these beliefs into the expression for t found in section A.1 yields:

$$\Rightarrow t = \frac{E(a_2|2) - E(a_2|1) + \rho_2}{E(a_1|1) - E(a_1|2)}$$

$$\Leftrightarrow t = \frac{\frac{2}{3} - \frac{3 - 2t}{6 - 3t} + \rho_2}{\frac{3 - t^2}{6 - 3t} - \frac{1}{3}t}$$

$$\Leftrightarrow t = \frac{\frac{2(2 - t)}{3(2 - t)} - \frac{3 - 2t}{3(2 - t)} + \frac{3(2 - t)\rho_2}{3(2 - t)}}{\frac{3 - t^2}{3(2 - t)} - \frac{(2 - t)t}{3(2 - t)}}$$

$$\Leftrightarrow t = \frac{1 + 6\rho_2 - 3\rho_2 t}{3 - 2t}$$

$$\Leftrightarrow 2t^2 - 3t + 6\rho_2 - 3\rho_2 t + 1 = 0$$

$$\Leftrightarrow 2t^2 - (3\rho_2 + 3)t + 6\rho_2 + 1 = 0$$

$$\Leftrightarrow \boldsymbol{t}^* = \frac{\sqrt{9\boldsymbol{\rho}_2^2 - 30\boldsymbol{\rho}_2 + 1} + 3\boldsymbol{\rho}_2 + 3}{4} \qquad \boldsymbol{or} \qquad \boldsymbol{t}^* = \frac{-\sqrt{9\boldsymbol{\rho}_2^2 - 30\boldsymbol{\rho}_2 + 1} + 3\boldsymbol{\rho}_2 + 3}{4}$$

On the other hand, in the case where $\underline{t > 1}$, employees' beliefs about their level of ability are updated as follows:

	$\boldsymbol{m = 1}$	$\boldsymbol{m = 2}$
$\boldsymbol{a_1}$	$E(a_1\|1) = \frac{\int_0^1 \int_0^{\frac{a_1}{t}} a_1 da_2 da_1}{\int_0^1 \int_0^{\frac{a_1}{t}} da_2 da_1} = \frac{2}{3}$	$E(a_1\|2) = \frac{\int_0^1 \int_{\frac{a_1}{t}}^1 a_1 da_2 da_1}{\int_0^1 \int_{\frac{a_1}{t}}^1 da_2 da_1} = \frac{2-3t}{3-6t}$
$\boldsymbol{a_2}$	$E(a_2\|1) = \frac{\int_0^1 \int_0^{\frac{a_1}{t}} a_2 da_2 da_1}{\int_0^1 \int_0^{\frac{a_1}{t}} da_2 da_1} = \frac{1}{3t}$	$E(a_2\|2) = \frac{\int_0^1 \int_{\frac{a_1}{t}}^1 a_2 da_2 da_1}{\int_0^1 \int_{\frac{a_1}{t}}^1 da_2 da_1} = \frac{1-3t^2}{3t-6t^2}$

Filling in these beliefs into the expression for *t* found in section A.1 yields:

$$\Rightarrow t = \frac{E(a_2|2) - E(a_2|1) + \rho_2}{E(a_1|1) - E(a_1|2)}$$

$$\Leftrightarrow t = \frac{\frac{1-3t^2}{3t-6t^2} - \frac{1}{3t} + \rho_2}{\frac{2}{3} - \frac{2-3t}{3-6t}}$$

$$\Leftrightarrow t = \frac{\frac{1-3t^2}{3t(1-2t)} - \frac{1-2t}{3t(1-2t)} + \frac{3t(1-2t)\rho_2}{3t(1-2t)}}{\frac{2(1-2t)}{3(1-2t)} - \frac{2-3t}{3(1-2t)}}$$

$$\Leftrightarrow t = \frac{2t - 3t^2 + 3\rho_2 t - 6\rho_2 t^2}{-t^2}$$

$$\Leftrightarrow t^3 + 2t - 3t^2 + 3\rho_2 t - 6\rho_2 t^2 = 0$$

$$\Leftrightarrow t^2 - (3 + 6\rho_2)t + 3\rho_2 + 2 = 0$$

$$\Leftrightarrow \boldsymbol{t}^* = \frac{\sqrt{36\boldsymbol{\rho}_2^2 + 24\boldsymbol{\rho}_2 + 1} + 6\boldsymbol{\rho}_2 + 3}{2} \qquad \boldsymbol{or} \qquad \boldsymbol{t}^* = \frac{-\sqrt{36\boldsymbol{\rho}_2^2 + 24\boldsymbol{\rho}_2 + 1} + 6\boldsymbol{\rho}_2 + 3}{2}$$

B.3) Stability of the basic game (with employee 2 bias)

$\underline{t \in [0,1]}$

The beliefs about the task assignment probabilities held by the employees ($\hat{t}$) differ from the equilibrium beliefs (t). For instance, the equilibrium $t = 1$ is stable if the manager's optimal response for $\hat{t}$ close but smaller than 1 is even closer to 1. Put differently: $\hat{t} < t < 1$. Otherwise, if $t < \hat{t}$, the manager's optimal response is further away from 1 than anticipated by the employees. In other words, the manager discriminates in favour of employee 1 even more than anticipated.

First, calculate the best response of the manager given the employee expectations updated using Bayes' rule:

$$t = \frac{E(a_2|2) - E(a_2|1) + \rho_2}{E(a_1|1) - E(a_1|2)} = \frac{\frac{2}{3} - \frac{3 - 2\hat{t}}{6 - 3\hat{t}} + \rho_2}{\frac{3 - \hat{t}^2}{6 - 3\hat{t}} - \frac{1}{3}\hat{t}} = \frac{6\rho_2 - 3\rho_2\hat{t} + 1}{3 - 2\hat{t}}$$

Then, check when $t > \hat{t}$ (in other words, when is the equilibrium stable?):

$$\frac{6\rho_2 - 3\rho_2\hat{t} + 1}{3 - 2\hat{t}} > \hat{t}$$

$$2\hat{t}^2 - 3\rho_2\hat{t} - 3\hat{t} + 6\rho_2 + 1 > 0$$

$$2\hat{t}^2 - (3\rho_2 + 3)\hat{t} + 6\rho_2 + 1 > 0$$

<u>$t > 1$</u>

The beliefs about the task assignment probabilities held by the employees ($\hat{t}$) differ from the equilibrium beliefs (t). For instance, the equilibrium $t = 1$ is stable if the manager's optimal response for $\hat{t}$ close but greater than 1 is even closer to 1. Put differently: $1 < t < \hat{t}$. Otherwise, if $t > \hat{t}$, the manager's optimal response is further away from 1 than anticipated by the employees. In other words, the manager discriminates in favour of employee 2 even more than anticipated.

First, calculate the best response of the manager given the employee expectations updated using Bayes' rule:

$$t = \frac{E(a_2|2) - E(a_2|1) + \rho_2}{E(a_1|1) - E(a_1|2)} = \frac{\frac{1 - 3\hat{t}^2}{3\hat{t} - 6\hat{t}^2} - \frac{1}{3\hat{t}} + \rho_2}{\frac{2}{3} - \frac{2 - 3\hat{t}}{3 - 6\hat{t}}} = \frac{6\rho_2\hat{t} - 3\rho_2 + 3\hat{t} - 2}{\hat{t}}$$

Then, check when $t < \hat{t}$ (in other words, when is the equilibrium stable?):

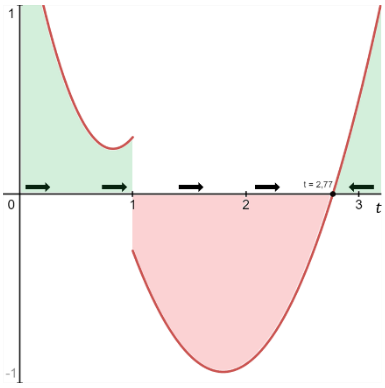
$$\frac{6\rho_2\hat{t} - 3\rho_2 + 3\hat{t} - 2}{\hat{t}} < \hat{t}$$

$$\hat{t}^2 - 6\rho_2\hat{t} + 3\rho_2 - 3\hat{t} + 2 > 0$$

$$\hat{t}^2 - (6\rho_1 + 3)\hat{t} + 3\rho_2 + 2 > 0$$

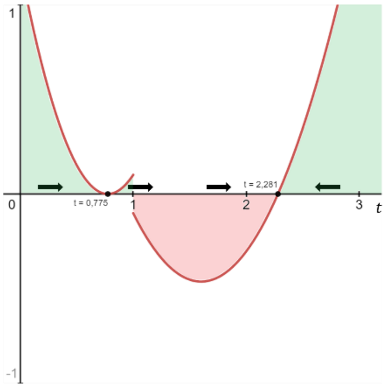
The stability of the equilibria in the game with bias will be determined by examining six separate cases with varying levels of bias. The reasoning is that different levels of bias determine the number of equilibria in the game and which are stable. In what follows, each case will be briefly discussed and graphically illustrated to make the interpretation more convenient.

<u>Case 1</u>: $\boldsymbol{\rho_2 > \frac{5}{3} - \frac{2\sqrt{2}}{\sqrt{3}}}$ (= 0,0337)



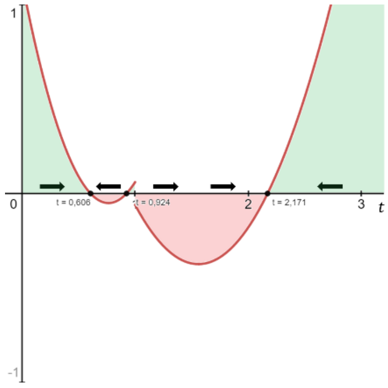
With a high positive bias towards employee 2 (in this case $\rho_2 = 0{,}1$) there exists only one equilibrium: $t = 2{,}77$. A high positive bias towards employee 2 leads to a discriminatory equilibrium in favour of that employee. To see this, note that for $t > 1$, the area indicated in green implies that $t < \hat{t}$ while the area in red implies that $t > \hat{t}$. On the other hand, for $t \in [0{,}1]$, the green area shows the values of the employees' beliefs $\hat{t}$ for which $t > \hat{t}$. Thus, the graph to the right shows that the optimal task assignment strategy converges to the single discriminatory equilibrium. For this level of bias, there is always one stable discriminatory equilibrium.

<u>Case 2</u>: $\boldsymbol{\rho_2 = \frac{5}{3} - \frac{2\sqrt{2}}{\sqrt{3}}}$ (= 0,0337)



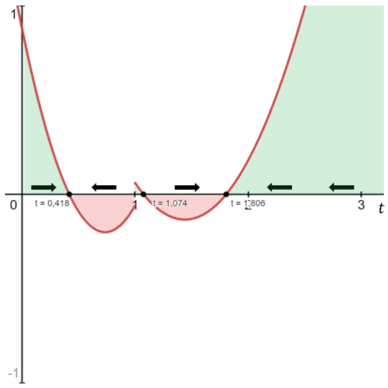
At the maximum positive bias towards employee 2 for which an equilibrium still exists with discrimination towards employee 1, there exist two equilibria: $t = 0{,}775$ and $t = 2{,}281$. The graph on the right-hand side shows that only $t = 2{,}281$ is stable because the optimal task assignment strategy converges towards it regardless of the employee beliefs $\hat{t}$. However, note that whenever the employees hold a belief on the interval $]0{,}775; 1[$, it is optimal for the manager to converge to the equilibrium in which he/she discriminates in the opposite direction ($t = 2{,}281$). The equilibrium $t = 0{,}775$ is unstable. For this level of bias, there is only one stable discriminatory equilibrium.

<u>Case 3</u>: $0 < \rho_2 < \frac{5}{3} - \frac{2\sqrt{2}}{\sqrt{3}}$ (= 0,0337)



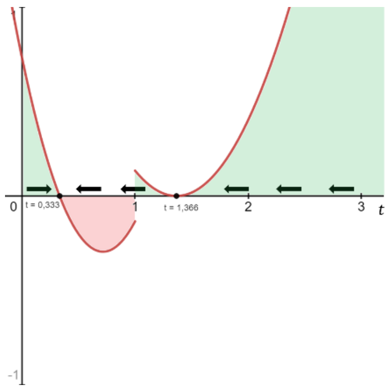
With a slight positive bias towards employee 2 (in this case $\rho_2 = 0{,}02$), there exist three equilibria: $t = 0{,}606$, $t = 0{,}924$ and $t = 2{,}171$. Note that for $t \in [0{,}1]$, the red area shows the values for the employee beliefs $\hat{t}$ for which $t < \hat{t}$. The graph shows how the optimal level of t converges toward the most discriminatory equilibria. The equilibrium in the middle ($t = 0{,}924$), although slightly discriminatory, is unstable. For this level of bias, out of three discriminatory equilibria, only the most discriminatory ones are stable.

<u>Case 4</u>: $0 > \rho_2 > \frac{\sqrt{3}}{6} - \frac{1}{3}$ (= −0,0447)



With a slight negative bias towards employee 2 (in this case $\rho_2 = -0{,}02$), there exist three equilibria: $t = 0{,}418$, $t = 1{,}074$ and $t = 1{,}806$. Similarly to case 3, the optimal level of t converges towards the most discriminatory equilibria. The equilibrium in the middle ($t = 1{,}074$), although slightly discriminatory, is unstable. For this level of bias, out of three discriminatory equilibria, only the most discriminatory ones are stable.

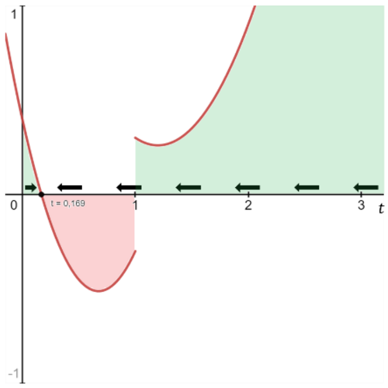
<u>Case 5</u>: $\rho_2 = \frac{\sqrt{3}}{6} - \frac{1}{3}$ (= −0,0447)



At the maximum negative bias towards employee 2 for which an equilibrium still exists with discrimination in favour of employee 2, there exist two equilibria: $t = 0{,}333$ and $t = 1{,}366$. The graph on the right-hand side shows that only $t = 0{,}333$ is stable because the optimal task assignment strategy converges towards it regardless of the employee beliefs $\hat{t}$. However, note that whenever the employees hold a belief on the interval $]1; 1{,}366[$, it is optimal for the manager to converge to the equilibrium in which he/she discriminates in the opposite direction ($t = 0{,}333$). The equilibrium $t = 1{,}366$ is unstable. For this level of bias, there is only one stable discriminatory equilibrium.

<u>Case 6</u>: $\boldsymbol{\rho_2 < \frac{\sqrt{3}}{6} - \frac{1}{3} (= -0{,}0447)}$

With a high negative bias towards employee 2 (in this case $\rho_2 = -0{,}1$) there exists only one equilibrium: $t = 0{,}169$. A high negative bias towards employee 2 leads to a discriminatory equilibrium favouring the other employee. The graph to the right shows that the optimal task assignment strategy converges to the single discriminatory equilibrium. For this level of bias, there is always one stable discriminatory equilibrium.



B.4) Efficiency of the basic game

<u>Manager payoff:</u>

The manager payoff is equal to the sum of the utilities resulting from the two possible task assignment decisions. Note that the manager's utility is increasing in both the total level of production and the bias towards employee 2.

For $t \in [0,1]$:

$$\underbrace{\int_0^1 \int_{ta_2}^1 \left(a_1 \frac{3-t^2}{6-3t} + a_2 \frac{3-2t}{6-3t} \right) da_1 da_2}_{m=1} + \underbrace{\int_0^1 \int_0^{ta_2} \left(a_1 \frac{t}{3} + a_2 \frac{2}{3} + a_2 \rho_2 \right) da_1 da_2}_{m=2} = \frac{t^3 - 3t^2 - 3\rho_2 t^2 - 2t + 6\rho_2 t + 9}{9(2-t)}$$

For $t > 1$:

$$\underbrace{\int_0^1 \int_0^{\frac{a_1}{t}} \left(a_1 \frac{2}{3} + a_2 \frac{1}{3t} \right) da_2 da_1}_{m=1} + \underbrace{\int_0^1 \int_{\frac{a_1}{t}}^1 \left(a_1 \frac{2-3t}{3-6t} + a_2 \frac{1-3t^2}{3t-6t^2} + a_2 \rho_2 \right) da_2 da_1}_{m=2} = \frac{18t^3 + 18\rho_2 t^3 - 4t^2 - 9\rho_2 t^2 - 6t - 6\rho_2 t + 2 + 3\rho_2}{18t^2(2t-1)}$$

Combining the two functions above leads to three general insights:

1. Without a bias, the manager's payoff is maximised whenever $t = 1$.
2. An increasing positive bias moves the optimal level of t further and further away from 1 ($t > 1$). The manager's payoff is maximised with more discrimination in favour of employee 2 and is higher.
3. An increasing negative bias moves the optimal level of t closer and closer to zero. The manager's payoff is maximised with more discrimination in favour of employee 1. However, the payoff is lower.

The reason why a positive bias increases the payoff at the optimal level of t is that the bias enters the manager's utility function as a positive term. On the other hand, whenever the bias is negative, utility logically decreases because there is still a possibility that employee 2 is assigned to the major task regardless of the negative bias against him/her.

Employee 1 payoff

Employee 1's payoff is equal to the sum of the utilities resulting from the two possible task assignment decisions.

For $t \in [0,1]$:

$$\underbrace{\int_0^1 \int_{ta_2}^1 \left(a_1 \frac{3-t^2}{6-3t} - \frac{1}{2}\left(\frac{3-t^2}{6-3t}\right)^2\right) da_1 da_2}_{m=1} + \underbrace{\int_0^1 \int_0^{ta_2} \left(a_1 \frac{t}{3} - \frac{1}{2}\left(\frac{t}{3}\right)^2\right) da_1 da_2}_{m=2} = \frac{2t^3 - 6t^2 + 9}{36(2-t)}$$

For $t > 1$:

$$\underbrace{\int_0^1 \int_0^{\frac{a_1}{t}} \left(a_1 \frac{2}{3} - \frac{1}{2}\left(\frac{2}{3}\right)^2\right) da_2 da_1}_{m=1} + \underbrace{\int_0^1 \int_{\frac{a_1}{t}}^1 \left(a_1 \frac{2-3t}{3-6t} - \frac{1}{2}\left(\frac{2-3t}{3-6t}\right)^2\right) da_2 da_1}_{m=2} = \frac{9t-4}{36(2t-1)}$$

Combining the two functions above leads to two general insights:

1. The bias only affects the payoff of employee 1 indirectly through the optimal task assignment decision t.
2. Employee 1's payoff is maximised at $t = \frac{3}{2} - \frac{\sqrt{3}}{2} \approx 0{,}634$. Employee 1 prefers some level of discrimination in his/her favour. To reach this equilibrium, the manager would need to have a slight positive bias ($\rho_2 = 0{,}024$) in favour of employee 2.

Employee 2 payoff

Employee 2's payoff is equal to the sum of the utilities resulting from the two possible task assignment decisions.

For $t \in [0,1]$:

$$\underbrace{\int_0^1 \int_{ta_2}^1 \left(a_2 \frac{3-2t}{6-3t} - \frac{1}{2}\left(\frac{3-2t}{6-3t}\right)^2\right) da_1 da_2}_{m=1} + \underbrace{\int_0^1 \int_0^{ta_2} \left(a_2 \frac{2}{3} - \frac{1}{2}\left(\frac{2}{3}\right)^2\right) da_1 da_2}_{m=2} = \frac{9-4t}{36(2-t)}$$

For $t > 1$:

$$\underbrace{\int_0^1 \int_0^{\frac{a_1}{t}} \left(a_2 \frac{1}{3t} - \frac{1}{2}\left(\frac{1}{3t}\right)^2\right) da_2 da_1}_{m=1} + \underbrace{\int_0^1 \int_{\frac{a_1}{t}}^1 \left(a_2 \frac{1-3t^2}{3t-6t^2} - \frac{1}{2}\left(\frac{1-3t^2}{3t-6t^2}\right)^2\right) da_2 da_1}_{m=2} = \frac{9t^3-6t+2}{36t^2(2t-1)}$$

Combining the two functions above leads to two general insights:

1. The bias only affects the payoff of employee 2 indirectly through the optimal task assignment decision t.

2. Employee 2's payoff is maximised at $t = 1 + \frac{1}{\sqrt{3}} \approx 1{,}577$. Employee 2 prefers some level of discrimination in his/her favour. To reach this equilibrium, the manager would need to have a slight negative bias ($\rho_1 = -0{,}0377$) towards employee 2.

Combined employee payoff

For $t \in [0,1]$:

$$\underbrace{\frac{2t^3 - 6t^2 + 9}{36(2-t)}}_{Employee\ 1} + \underbrace{\frac{9-4t}{36(2-t)}}_{Employee\ 2} = \frac{t^3 - 3t^2 - 2t + 9}{36 - 18t}$$

For $t > 1$:

$$\underbrace{\frac{9t - 4}{36(2t-1)}}_{Employee\ 1} + \underbrace{\frac{9t^3 - 6t + 2}{36t^2(2t-1)}}_{Employee\ 2} = \frac{9t^3 - 2t^2 - 3t + 1}{36t^3 - 18t^2}$$

The combined utility for both employees is maximised only when $t = 1$. Even though the employees individually prefer some discrimination in their own favour, the combined payoff is maximised without discrimination.

Firm payoff:

The firm payoff is equal the manager payoff without bias. The only goal for the firm is to maximise the output generated by the two employees.

For $t \in [0,1]$:

$$\underbrace{\int_0^1 \int_{ta_2}^1 \left(a_1 \frac{3-t^2}{6-3t} + a_2 \frac{3-2t}{6-3t}\right) da_1 da_2}_{m=1} + \underbrace{\int_0^1 \int_0^{ta_2} \left(a_1 \frac{t}{3} + a_2 \frac{2}{3}\right) da_1 da_2}_{m=2} = \frac{t^3 - 3t^2 - 2t + 9}{9(2-t)}$$

For $t > 1$:

$$\underbrace{\int_0^1 \int_0^{\frac{a_1}{t}} \left(a_1 \frac{2}{3} + a_2 \frac{1}{3t}\right) da_2 da_1}_{m=1} + \underbrace{\int_0^1 \int_{\frac{a_1}{t}}^1 \left(a_1 \frac{2-3t}{3-6t} + a_2 \frac{1-3t^2}{3t-6t^2}\right) da_2 da_1}_{m=2} = \frac{9t^3 - 2t^2 - 3t + 1}{9t^2(2t-1)}$$

The payoff for the firm is maximised only when $t = 1$. The manager prefers some level of discrimination depending on whether he/she has a positive or negative bias towards employee 2. However, from a productivity point of view, it would be optimal for the manager to be unbiased and not to discriminate.

B.5) Equilibria of the extended game

To determine the optimal task assignment strategy, the expected employee ability levels given the task assignment decision have to be determined. The approach will be largely similar to the one in *Section B.2*. However, the task importance parameter η will now be equal to or greater than 1. Again, employees update their beliefs according to Bayes' rule. First, consider the case where $t \in [0,1]$. Expected employee ability levels can then be calculated the same as before:

	$\mathbf{m = 1}$	$\mathbf{m = 2}$
$\mathbf{a_1}$	$E(a_1\|1) = \frac{\int_0^1 \int_{ta_2}^1 a_1 da_1 da_2}{\int_0^1 \int_{ta_2}^1 da_1 da_2} = \frac{3 - t^2}{6 - 3t}$	$E(a_1\|2) = \frac{\int_0^1 \int_0^{ta_2} a_1 da_1 da_2}{\int_0^1 \int_0^{ta_2} da_1 da_2} = \frac{1}{3}t$
$\mathbf{a_2}$	$E(a_2\|1) = \frac{\int_0^1 \int_{ta_2}^1 a_2 da_1 da_2}{\int_0^1 \int_{ta_2}^1 da_1 da_2} = \frac{3 - 2t}{6 - 3t}$	$E(a_2\|2) = \frac{\int_0^1 \int_0^{ta_2} a_2 da_1 da_2}{\int_0^1 \int_0^{ta_2} da_1 da_2} = \frac{2}{3}$

Filling in these beliefs into the expression for t found in *Section B.1* yields:

$$\Rightarrow t = \frac{\eta^2 E(a_2|2) - E(a_2|1) + \rho_2}{\eta^2 E(a_1|1) - E(a_1|2)}$$

$$\Leftrightarrow t = \frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t} + \rho_2}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t}$$

$$\Leftrightarrow t = \frac{\frac{2(2-t)\eta^2}{3(2-t)} - \frac{3-2t}{3(2-t)} + \frac{3(2-t)\rho_2}{3(2-t)}}{\frac{(3-t^2)\eta^2}{3(2-t)} - \frac{(2-t)t}{3(2-t)}}$$

$$\Leftrightarrow t = \frac{4\eta^2 - 2\eta^2 t - 3 + 2t + 6\rho_2 - 3\rho_2 t}{3\eta^2 - \eta^2 t^2 - 2t + t^2}$$

$$\boxed{\Leftrightarrow (\boldsymbol{\eta}^2 - \mathbf{1})\boldsymbol{t}^3 + \mathbf{2}\boldsymbol{t}^2 + (\mathbf{2} - \mathbf{5}\boldsymbol{\eta}^2 - \mathbf{3}\boldsymbol{\rho}_2)\boldsymbol{t} + \mathbf{4}\boldsymbol{\eta}^2 + \mathbf{6}\boldsymbol{\rho}_2 - \mathbf{3} = \mathbf{0}}$$

On the other hand, in the case where $t > 1$, employees' beliefs about their level of ability are updated as follows:

	$\boldsymbol{m = 1}$	$\boldsymbol{m = 2}$
$\boldsymbol{a_1}$	$E(a_1\|1) = \frac{\int_0^1 \int_0^{\frac{a_1}{t}} a_1 da_2 da_1}{\int_0^1 \int_0^{\frac{a_1}{t}} da_2 da_1} = \frac{2}{3}$	$E(a_1\|2) = \frac{\int_0^1 \int_{\frac{a_1}{t}}^1 a_1 da_2 da_1}{\int_0^1 \int_{\frac{a_1}{t}}^1 da_2 da_1} = \frac{2 - 3t}{3 - 6t}$
$\boldsymbol{a_2}$	$E(a_2\|1) = \frac{\int_0^1 \int_0^{\frac{a_1}{t}} a_2 da_2 da_1}{\int_0^1 \int_0^{\frac{a_1}{t}} da_2 da_1} = \frac{1}{3t}$	$E(a_2\|2) = \frac{\int_0^1 \int_{\frac{a_1}{t}}^1 a_2 da_2 da_1}{\int_0^1 \int_{\frac{a_1}{t}}^1 da_2 da_1} = \frac{1 - 3t^2}{3t - 6t^2}$

Filling in these beliefs into the expression for t found in section A.1 yields:

$$\Rightarrow t = \frac{\eta^2 E(a_2|2) - E(a_2|1) + \rho_2}{\eta^2 E(a_1|1) - E(a_1|2)}$$

$$\Leftrightarrow t = \frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t} + \rho_2}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t}}$$

$$\Leftrightarrow t = \frac{\frac{(1-3t^2)\eta^2}{3t(1-2t)} - \frac{1-2t}{3t(1-2t)} + \frac{3t(1-2t)\rho_2}{3t(1-2t)}}{\frac{2(1-2t)\eta^2}{3(1-2t)} - \frac{2-3t}{3(1-2t)}}$$

$$\Leftrightarrow t = \frac{\eta^2 - 3\eta^2 t^2 - 1 + 2t + 3\rho_2 t - 6\rho_2 t^2}{2\eta^2 t - 4\eta^2 t^2 - 2t + 3t^2}$$

$$\boxed{\Leftrightarrow (3-4\eta^2)t^3 - (2-5\eta^2-6\rho_2)t^2 - (2+3\rho_2)t - \eta^2 + 1 = 0}$$

Case 1: No manager bias ($\rho_1 = 0$)

Without a manager bias towards employee 2, the equations found above can be solved for t as follows:

For $t \in [0,1]$:

$$t^* = \frac{\eta^2 - \sqrt{17\eta^4 - 26\eta^2 + 13} + 1}{2 - 2\eta^2} \qquad or \qquad t^* = 1$$

For $t > 1$:

$$t^* = 1 \qquad or \qquad t^* = \frac{\eta^2 + \sqrt{17\eta^4 - 26\eta^2 + 13} + 1}{8\eta^2 - 6}$$

In other words, there are two discriminatory equilibria and one non-discriminatory equilibrium. However, whenever $\eta \geq \sqrt{1{,}5}$, $t^* = 1$ becomes the unique equilibrium.

Case 2: Positive manager bias ($\rho_2 > 0$)

The following cases in which $\rho_1 \neq 0$ will not be solved algebraically as the third power in the equation makes it unnecessarily difficult to do so. However, the most important results can be smoothly derived the same way as in *Part A*.

For $t \in [0,1]$, a positive bias towards employee 2 moves the equilibrium task assignment decision t closer to 1 (no discrimination). However, for every level of bias, there is a maximum level of major task importance after which an equilibrium for $t \in [0,1]$ ceases to exist. This is because there are two forces at play in determining the equilibria: (1) the bias causes the manager to want to discriminate in favour of employee 2, while (2) the major task importance causes the manager to discriminate less. This is why an increasing positive bias combined with a high level of major task importance makes it impossible for an equilibrium to exist with discrimination in favour of employee 1.

On the other hand, for $t > 1$, an increasing bias requires a higher level of η to move closer to the non-discriminatory equilibrium. Additionally, reaching the non-discriminatory equilibrium becomes impossible with any level of bias.

Algebraically:
$(3 - 4\eta^2)t^3 - (2 - 5\eta^2 - 6\rho_2)t^2 - (2 + 3\rho_2)t - \eta^2 + 1 = 0$
$\Leftrightarrow 3 - 4\eta^2 - 2 + 5\eta^2 + 6\rho_2 - 2 - 3\rho_2 - \eta^2 + 1 = 0$ whenever $t = 1$
$\Leftrightarrow 3\rho_2 = 0$

The equality above can only hold if $\rho_1 = 0$. Any level of bias means that $t = 1$ cannot be an equilibrium regardless of the major task importance.

Case 3: Negative manager bias ($\rho_2 < 0$)

For $t \in [0,1]$, an increasing negative bias requires a higher level of η to move closer to the non-discriminatory equilibrium. Additionally, reaching the non-discriminatory equilibrium becomes impossible with any level of bias.

Algebraically:
$(\eta^2 - 1)t^3 + 2t^2 + (2 - 5\eta^2 - 3\rho_2)t + 4\eta^2 + 6\rho_2 - 3 = 0$
$\Leftrightarrow \eta^2 - 1 + 2 + 2 - 5\eta^2 - 3\rho_2 + 4\eta^2 + 6\rho_2 - 3 = 0$ whenever $t = 1$
$\Leftrightarrow 3\rho_2 = 0$

The equality above can only hold if $\rho_1 = 0$. Any level of bias means that $t = 1$ cannot be an equilibrium regardless of the major task importance.

On the other hand, for $t > 1$, a negative bias towards employee 2 also moves the equilibrium task assignment decision t closer to 1 (no discrimination). However, for every level of bias, there is a maximum level of major task importance after which an equilibrium for $t > 1$ ceases to exist. This is because here too, there are two forces at play in determining the equilibria: (1) the negative bias causes the manager to want to discriminate in favour of employee 1, while (2) the major task importance causes the manager to discriminate less. This is why an increasing positive bias combined with a high level of major task importance makes it impossible for an equilibrium to exist with discrimination in favour of employee 2.

B.6) Stability of the extended game

Stability for $t \in [0,1]$ ($\rho_1 \neq 0$)

Just like in the basic model, the equilibrium values for t are found as follows:

$$t = \frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t} + \rho_2}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t}$$

In order to check the stability of the equilibria, the following rule is used:

The equilibrium is stable if:

$$\frac{\partial}{\partial t}t > \frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t}}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} \Leftrightarrow \frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t} + \rho_2}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} < 1$$

The equilibrium is unstable if:

$$\frac{\partial}{\partial t}t < \frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t} + \rho_2}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} \Leftrightarrow \frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t} + \rho_2}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t} > 1$$

Solving the derivative gives:

$$\frac{\partial}{\partial t}\frac{\frac{2}{3}\eta^2 - \frac{3-2t}{6-3t} + \rho_2}{\frac{3-t^2}{6-3t}\eta^2 - \frac{1}{3}t}$$
$$= \frac{-2\eta^4(t^2-4t+3) + 2\eta^2(2t^2-7t+7) - 3\rho_2\eta^2(t^2-4t+3) + 3\rho_2(t-2)^2 - 2(t^2-3t+3)}{[(t-2)t - \eta^2(t^2-3)]^2}$$

Every possible combination of η and ρ_1 leading to a value of t will result in a derivative smaller than 1. All possible equilibria in the game with bias are discriminatory and stable.

Stability for $t > 1$ ($\rho_1 \neq 0$)

Just like in the basic model, the equilibrium values for t are found as follows:

$$t = \frac{\frac{1-3t^2}{3t-6t^2}\eta^2 - \frac{1}{3t} + \rho_2}{\frac{2}{3}\eta^2 - \frac{2-3t}{3-6t}}$$

In order to check the stability of the equilibria, the following rule is used:

The equilibrium is stable if:

$$\frac{\partial}{\partial t}t > \frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2-\frac{1}{3t}+\rho_2}{\frac{2}{3}\eta^2-\frac{2-3t}{3-6t}} \Leftrightarrow \frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2-\frac{1}{3t}+\rho_2}{\frac{2}{3}\eta^2-\frac{2-3t}{3-6t}} < 1$$

The equilibrium is unstable if:

$$\frac{\partial}{\partial t}t < \frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2-\frac{1}{3t}+\rho_2}{\frac{2}{3}\eta^2-\frac{2-3t}{3-6t}} \Leftrightarrow \frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2-\frac{1}{3t}+\rho_2}{\frac{2}{3}\eta^2-\frac{2-3t}{3-6t}} > 1$$

Solving the derivative gives:

$$\frac{\partial}{\partial t}\frac{\frac{1-3t^2}{3t-6t^2}\eta^2-\frac{1}{3t}+\rho_2}{\frac{2}{3}\eta^2-\frac{2-3t}{3-6t}} = \frac{[\eta^4(-6t^2+8t-2)+2\eta^2(7t^2-7t+2)+3\rho_1t^2-6t^2+6t-2]}{t^2[\eta^2(4t-2)-3t+2]^2}$$

Every possible combination of η and ρ_1 leading to a value of t will result in a derivative smaller than 1. All possible equilibria in the game with bias are discriminatory and stable.

B.7) Positive and negative bias effect comparison

This section will provide evidence on the following:

- A negative bias has the strongest effect in aggravating discrimination
- A positive bias has the strongest effect in decreasing discrimination

First, consider the following equations:

Model (a): $t^*_{A1} = \frac{-\sqrt{36\rho_1^2+24\rho_1+1}+6\rho_1+3}{6\rho_1+4}$ $(t^*_{A1} = 0{,}5 \text{ if } \rho_1 = 0)$

$t^*_{A2} = \frac{\sqrt{9\rho_1^2-30\rho_1+1}+3\rho_1+3}{12\rho_1+2}$ $(t^*_{A2} = 2 \text{ if } \rho_1 = 0)$

Model (b): $t^*_{B1} = \frac{-\sqrt{9\rho_2^2-30\rho_2+1}+3\rho_2+3}{4}$ $(t^*_{B1} = 0{,}5 \text{ if } \rho_2 = 0)$

$t^*_{B2} = \frac{\sqrt{36\rho_2^2+24\rho_2+1}+6\rho_2+3}{2}$ $(t^*_{B2} = 2 \text{ if } \rho_2 = 0)$

To show that a negative bias has the strongest effect on worsening discrimination compared to an equal positive bias, the following inequalities should hold:

(I) $t_{A1}^*(\rho) > t_{B1}^*(-\rho)$ for $0 < \rho < \frac{1}{6}$

(II) $t_{A2}^*(-\rho) > t_{B2}^*(\rho)$ for $0 < \rho < \frac{1}{6}$

(I) With a positive bias towards employee 1 (ρ), the equilibrium $t_{A1}^* = 0{,}5$ will become more discriminatory (move closer to zero). However, an equal negative bias ($-\rho$) against employee 2 should have a stronger effect on worsening discrimination and should thus be smaller. ρ cannot be greater than $\frac{1}{6}$ because this is the maximum possible negative bias in the model.

$$\Leftrightarrow \frac{-\sqrt{36\rho^2+24\rho+1}+6\rho+3}{6\rho+4} > \frac{-\sqrt{9(-\rho)^2-30(-\rho)+1}+3(-\rho)+3}{4} \quad \text{for } 0 < \rho < \frac{1}{6}$$

$$\Leftrightarrow \frac{-\sqrt{36\rho^2+24\rho+1}+6\rho+3}{6\rho+4} > \frac{-\sqrt{9\rho^2+30\rho+1}-3\rho+3}{4} \quad \text{for } 0 < \rho < \frac{1}{6}$$

As shown on Figure (I), the inequality above holds for the entire interval.

(II) With a negative bias against employee 1 ($-\rho$), the equilibrium $t_{A2}^* = 2$ will become more discriminatory in favour of employee 2 (move away from $t_{A2}^* = 2$). On the other hand, an equal positive bias (ρ) towards employee 2 also worsens discrimination but should have a weaker effect and should thus be smaller. ρ cannot be greater than $\frac{1}{6}$ because this is the maximum possible negative bias in the model.

$$\Leftrightarrow \frac{\sqrt{9(-\rho)^2-30(-\rho)+1}+3(-\rho)+3}{12(-\rho)+2} > \frac{\sqrt{36\rho^2+24\rho+1}+6\rho+3}{2} \quad \text{for } 0 < \rho < \frac{1}{6}$$

$$\Leftrightarrow \frac{\sqrt{9\rho^2+30\rho+1}-3\rho+3}{2-12\rho} > \frac{\sqrt{36\rho^2+24\rho+1}+6\rho+3}{2} \quad \text{for } 0 < \rho < \frac{1}{6}$$

As shown on Figure (II), the inequality above holds for the entire interval.

To show that a positive bias has the strongest effect on ameliorating discrimination compared to an equal negative bias, the following inequalities should hold:

(III) $t_{A1}^*(-\rho) < t_{B1}^*(\rho)$ for $0 < \rho < \frac{5}{3} - \frac{2\sqrt{2}}{\sqrt{3}}$

(IV) $t_{A2}^*(\rho) < t_{B2}^*(-\rho)$ for $0 < \rho < \frac{5}{3} - \frac{2\sqrt{2}}{\sqrt{3}}$

(III) With a negative bias against employee 1 ($-\rho$), the equilibrium $t_{A1}^* = 0{,}5$ will become less discriminatory (move closer to 1). On the other hand, an equal positive bias (ρ) towards employee 2 should have a stronger effect on ameliorating discrimination and should thus be greater (or closer to 1). ρ cannot be greater than $\frac{5}{3} - \frac{2\sqrt{2}}{\sqrt{3}} = (0{,}0337)$ because this is the maximum possible positive bias for which t_{B1}^* has a solution.

$$\Leftrightarrow \frac{-\sqrt{36(-\rho)^2+24(-\rho)+1}+6(-\rho)+3}{6(-\rho)+4} < \frac{-\sqrt{9\rho^2-30\rho+1}+3\rho+3}{4} \quad \text{for } 0<\rho<\frac{5}{3}-\frac{2\sqrt{2}}{\sqrt{3}}$$

$$\Leftrightarrow \frac{-\sqrt{36\rho^2-24\rho+1}-6\rho+3}{4-6\rho} < \frac{-\sqrt{9\rho^2-30\rho+1}+3\rho+3}{4} \quad \text{for } 0<\rho<\frac{5}{3}-\frac{2\sqrt{2}}{\sqrt{3}}$$

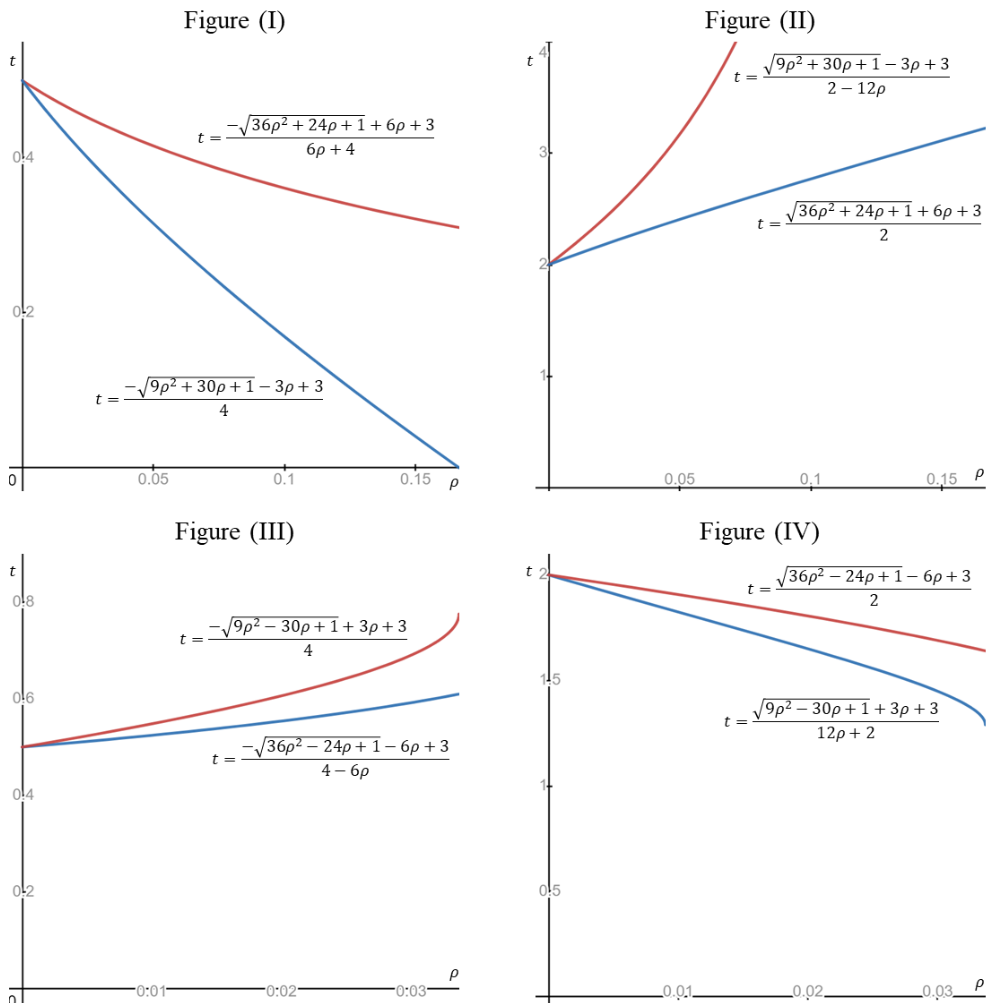
As shown on Figure (III), the inequality above holds for the entire interval.

(IV) With a positive bias towards employee 1 (ρ), the equilibrium $t_{A2}^* = 2$ will become less discriminatory (move closer to 1). An equal negative bias ($-\rho$) against employee 2 should have a weaker effect on ameliorating discrimination and should thus be greater (or closer to 2). ρ cannot be greater than $\frac{5}{3}-\frac{2\sqrt{2}}{\sqrt{3}} = (0{,}0337)$ because this is the maximum possible positive bias for which t_{A2}^* has a solution.

$$\Leftrightarrow \frac{\sqrt{9\rho^2-30\rho+1}+3\rho+3}{12\rho+2} < \frac{\sqrt{36(-\rho)^2+24(-\rho)+1}+6(-\rho)+3}{2} \quad \text{for } 0<\rho<\frac{5}{3}-\frac{2\sqrt{2}}{\sqrt{3}}$$

$$\Leftrightarrow \frac{\sqrt{9\rho^2-30\rho+1}+3\rho+3}{12\rho+2} < \frac{\sqrt{36\rho^2-24\rho+1}-6\rho+3}{2} \quad \text{for } 0<\rho<\frac{5}{3}-\frac{2\sqrt{2}}{\sqrt{3}}$$

As shown on Figure (IV), the inequality above holds for the entire interval.



C) Analysis of the task assignment game using *Model (c)*

C.1) Optimal task assignment for the basic game

The manager is indifferent between assigning the major task to employee 1 and employee 2 if:

$$U_M(m=1) = U_M(m=2)$$

Filling in using *Equations (1)* and *(3'')* yields:

$$\eta a_1 e_1 + a_2 e_2 + \rho_1 a_1 = a_1 e_1 + \eta a_2 e_2 + \rho_2 a_2$$

Given the optimal effort levels $e_i = \eta E(a_i|m)$ and $e_i = E(a_i|m)$, this becomes:

$$\eta a_1 \eta E(a_1|m=1) + a_2 E(a_2|m=1) + \rho_1 a_1 = a_1 E(a_1|m=2) + \eta a_2 \eta E(a_2|m=2) + \rho_2 a_2$$

Rearranging gives:

$$[\eta^2 E(a_1|1) - E(a_1|2) + \rho_1]a_1 = [\eta^2 E(a_2|2) - E(a_2|1) + \rho_2]a_2$$

Resulting in:

$$a_1 = ta_2$$
$$where \quad t = \frac{\eta^2 E(a_2|2) - E(a_2|1) + \rho_2}{\eta^2 E(a_1|1) - E(a_1|2) + \rho_1}$$

C.2) Equilibria of the basic game

To determine the optimal task assignment strategy, the expected employee ability levels given the task assignment decision have to be determined. Employees update their beliefs according to Bayes' rule. First, consider the case where $t \in [0,1]$. Expected employee ability levels can then be calculated as follows:

	$\mathbf{m = 1}$	$\mathbf{m = 2}$
$\mathbf{a_1}$	$E(a_1\|1) = \frac{\int_0^1 \int_{ta_2}^1 a_1 da_1 da_2}{\int_0^1 \int_{ta_2}^1 da_1 da_2} = \frac{3-t^2}{6-3t}$	$E(a_1\|2) = \frac{\int_0^1 \int_0^{ta_2} a_1 da_1 da_2}{\int_0^1 \int_0^{ta_2} da_1 da_2} = \frac{1}{3}t$
$\mathbf{a_2}$	$E(a_2\|1) = \frac{\int_0^1 \int_{ta_2}^1 a_2 da_1 da_2}{\int_0^1 \int_{ta_2}^1 da_1 da_2} = \frac{3-2t}{6-3t}$	$E(a_2\|2) = \frac{\int_0^1 \int_0^{ta_2} a_2 da_1 da_2}{\int_0^1 \int_0^{ta_2} da_1 da_2} = \frac{2}{3}$

Filling in these beliefs into the expression for t found in section A.1 yields:

$$\Rightarrow t = \frac{E(a_2|2) - E(a_2|1) + \rho_2}{E(a_1|1) - E(a_1|2) + \rho_1}$$

$$\Leftrightarrow t = \frac{\frac{2}{3} - \frac{3-2t}{6-3t} + \rho_2}{\frac{3-t^2}{6-3t} - \frac{1}{3}t + \rho_1}$$

$$\Leftrightarrow t = \frac{\frac{2(2-t)}{3(2-t)} - \frac{3-2t}{3(2-t)} + \frac{3(2-t)\rho_2}{3(2-t)}}{\frac{3-t^2}{3(2-t)} - \frac{(2-t)t}{3(2-t)} + \frac{3(2-t)\rho_1}{3(2-t)}}$$

$$\Leftrightarrow t = \frac{1 + 6\rho_2 - 3\rho_2 t}{3 - 2t + 6\rho_1 - 3\rho_1 t}$$

$$\Leftrightarrow -3t + 2t^2 - 6\rho_1 t + 3\rho_1 t^2 + 1 + 6\rho_2 - 3\rho_2 t = 0$$

$$\Leftrightarrow (2 + 3\rho_1)t^2 + (-3 - 6\rho_1 - 3\rho_2)t + 1 + 6\rho_2 = 0$$

$$\boldsymbol{\Leftrightarrow t^* = \frac{\sqrt{36\rho_1^2 - 36\rho_1\rho_2 + 24\rho_1 + 9\rho_2^2 - 30\rho_2 + 1} + 6\rho_1 + 3\rho_2 + 3}{6\rho_1 + 4}}$$

$$\boldsymbol{or\ \ t^* = \frac{-\sqrt{36\rho_1^2 - 36\rho_1\rho_2 + 24\rho_1 + 9\rho_2^2 - 30\rho_2 + 1} + 6\rho_1 + 3\rho_2 + 3}{6\rho_1 + 4}}$$

On the other hand, in the case where $t \geq 1$, employees' beliefs about their level of ability are updated as follows:

	$\boldsymbol{m = 1}$	$\boldsymbol{m = 2}$
$\boldsymbol{a_1}$	$E(a_1\|1) = \frac{\int_0^1 \int_0^{\frac{a_1}{t}} a_1 da_2 da_1}{\int_0^1 \int_0^{\frac{a_1}{t}} da_2 da_1} = \frac{2}{3}$	$E(a_1\|2) = \frac{\int_0^1 \int_{\frac{a_1}{t}}^1 a_1 da_2 da_1}{\int_0^1 \int_{\frac{a_1}{t}}^1 da_2 da_1} = \frac{2-3t}{3-6t}$
$\boldsymbol{a_2}$	$E(a_2\|1) = \frac{\int_0^1 \int_0^{\frac{a_1}{t}} a_2 da_2 da_1}{\int_0^1 \int_0^{\frac{a_1}{t}} da_2 da_1} = \frac{1}{3t}$	$E(a_2\|2) = \frac{\int_0^1 \int_{\frac{a_1}{t}}^1 a_2 da_2 da_1}{\int_0^1 \int_{\frac{a_1}{t}}^1 da_2 da_1} = \frac{1-3t^2}{3t-6t^2}$

Filling in these beliefs into the expression for *t* found in section A.1 yields:

$$\Rightarrow t = \frac{E(a_2|2) - E(a_2|1) + \rho_2}{E(a_1|1) - E(a_1|2) + \rho_1}$$

$$\Leftrightarrow t = \frac{\frac{1-3t^2}{3t-6t^2} - \frac{1}{3t} + \rho_2}{\frac{2}{3} - \frac{2-3t}{3-6t} + \rho_1}$$

$$\Leftrightarrow t = \frac{\frac{1-3t^2}{3t(1-2t)} - \frac{1-2t}{3t(1-2t)} + \frac{3t(1-2t)\rho_2}{3t(1-2t)}}{\frac{2(1-2t)}{3(1-2t)} - \frac{2-3t}{3(1-2t)} + \frac{3(1-2t)\rho_1}{3(1-2t)}}$$

$$\Leftrightarrow t = \frac{2t - 3t^2 + 3\rho_2 t - 6\rho_2 t^2}{-t^2 + 3\rho_1 t - 6\rho_1 t^2}$$

$$\Leftrightarrow t^3 - 3\rho_1 t^2 + 6\rho_1 t^3 + 2t - 3t^2 + 3\rho_2 t - 6\rho_2 t^2 = 0$$

$$\Leftrightarrow (1 + 6\rho_1)t^2 + (-3 - 3\rho_1 - 6\rho_2)t + 2 + 3\rho_2 = 0$$

$$\boxed{\Leftrightarrow \boldsymbol{t^*} = \frac{\sqrt{9\rho_1^2 - 36\rho_1\rho_2 - 30\rho_1 + 36\rho_2^2 + 24\rho_2 + 1} + 3\rho_1 + 6\rho_2 + 3}{12\rho_1 + 2}}$$

$$\boxed{\boldsymbol{or}\ \ \boldsymbol{t^*} = \frac{-\sqrt{9\rho_1^2 - 36\rho_1\rho_2 - 30\rho_1 + 36\rho_2^2 + 24\rho_2 + 1} + 3\rho_1 + 6\rho_2 + 3}{12\rho_1 + 2}}$$

C.3) Stability of the basic game

<u>$t \in [0,1]$</u>

The beliefs about the task assignment probabilities held by the employees ($\hat{t}$) differ from the equilibrium beliefs (t). For instance, the equilibrium $t = 1$ is stable if the manager's optimal response for $\hat{t}$ close but smaller than 1 is even closer to 1. Put differently: $\hat{t} < t < 1$. Otherwise, if $t < \hat{t}$, the manager's optimal response is further away from 1 than anticipated by the employees. In other words, the manager discriminates in favour of employee 1 even more than anticipated.

First, calculate the best response of the manager given the employee expectations updated using Bayes' rule:

$$t = \frac{E(a_2|2) - E(a_2|1) + \rho_2}{E(a_1|1) - E(a_1|2) + \rho_1} = \frac{\frac{2}{3} - \frac{3-2\hat{t}}{6-3\hat{t}} + \rho_2}{\frac{3-\hat{t}^2}{6-3\hat{t}} - \frac{1}{3}\hat{t} + \rho_1} = \frac{6\rho_2 - 3\rho_2\hat{t} + 1}{6\rho_1 - 3\rho_1\hat{t} - 2\hat{t} + 3}$$

Then, check when $t > \hat{t}$ (in other words, when is the equilibrium stable?):

$$\frac{6\rho_2 - 3\rho_2\hat{t} + 1}{6\rho_1 - 3\rho_1\hat{t} - 2\hat{t} + 3} > \hat{t}$$

$$6\rho_2 - 3\rho_2\hat{t} + 3\rho_1\hat{t}^2 + 2\hat{t}^2 - 6\rho_1\hat{t} - 3\hat{t} + 1 > 0$$

$$(3\rho_1 + 2)\hat{t}^2 - (6\rho_1 + 3\rho_2 + 3)\hat{t} + 1 + 6\rho_2 > 0$$

$\underline{t > 1}$

The beliefs about the task assignment probabilities held by the employees ($\hat{t}$) differ from the equilibrium beliefs (t). For instance, the equilibrium $t = 1$ is stable if the manager's optimal response for $\hat{t}$ close but greater than 1 is even closer to 1. Put differently: $1 < t < \hat{t}$. Otherwise, if $t > \hat{t}$, the manager's optimal response is further away from 1 than anticipated by the employees. In other words, the manager discriminates in favour of employee 2 even more than anticipated.

First, calculate the best response of the manager given the employee expectations updated using Bayes' rule:

$$t = \frac{E(a_2|2) - E(a_2|1) + \rho_2}{E(a_1|1) - E(a_1|2) + \rho_1} = \frac{\frac{1 - 3\hat{t}^2}{3\hat{t} - 6\hat{t}^2} - \frac{1}{3\hat{t}} + \rho_2}{\frac{2}{3} - \frac{2 - 3\hat{t}}{3 - 6\hat{t}} + \rho_1} = \frac{6\rho_2\hat{t} - 3\rho_2 + 3\hat{t} - 2}{6\rho_1\hat{t} - 3\rho_1 + \hat{t}}$$

Then, check when $t < \hat{t}$ (in other words, when is the equilibrium stable?):

$$\frac{6\rho_2\hat{t} - 3\rho_2 + 3\hat{t} - 2}{6\rho_1\hat{t} - 3\rho_1 + \hat{t}} < \hat{t}$$

$$6\rho_1\hat{t}^2 + \hat{t}^2 - 3\rho_1\hat{t} - 6\rho_2\hat{t} + 3\rho_2 - 3\hat{t} + 2 > 0$$

$$(6\rho_1 + 1)\hat{t}^2 - (3\rho_1 + 6\rho_2 + 3)\hat{t} + 2 + 3\rho_2 > 0$$

Taking together the inequalities above, the following conclusions can be drawn regarding the stability of the equilibria:

1. Whenever there are three equilibria, only the most discriminatory ones are stable.
2. An <u>equal</u> positive bias in the interval $0 \leq \rho_1 = \rho_2 < \frac{1}{3}$ results in two stable discriminatory equilibria and one unstable non-discriminatory equilibrium.
3. An <u>equal</u> positive bias greater than $\rho_1 = \rho_2 \geq \frac{1}{3}$ results in a unique and stable non-discriminatory equilibrium
4. An equal negative bias in the interval $-\frac{1}{6} \leq \rho_1 = \rho_2 \leq 0$ results in two stable discriminatory equilibria and one unstable non-discriminatory equilibrium.