THE DYNAMICS OF THE GENDER WAGE GAP IN LATE CAREER: A THEORETICAL MODEL

A Master's Thesis

by
GÜLCE DEMİREL

Department of
Economics
İhsan Doğramacı Bilkent University
Ankara
August 2024

To my sister and my cat

THE DYNAMICS OF THE GENDER WAGE GAP IN LATE CAREER: A THEORETICAL MODEL

The Graduate School of Economics and Social Sciences
of
İhsan Doğramacı Bilkent University

by

Gülce DEMİREL

In Partial Fulfillment of the Requirements for the Degree of
MASTER OF ARTS IN ECONOMICS

THE DEPARTMENT OF
ECONOMICS
İHSAN DOĞRAMACI BİLKENT UNIVERSITY
ANKARA

August 2024

THE DYNAMICS OF THE GENDER WAGE GAP IN LATE CAREER: A THEORETICAL MODEL
By Gülce Demirel

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Ahmet Arda Gitmez
Advisor

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Çağla Ökten
Examining Committee Member

I certify that I have read this thesis and have found that it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Arts in Economics.

Emre Ekinci
Examining Committee Member

Approval of the Graduate School of Economics and Social Sciences

Refet S. Gürkaynak
Director

ABSTRACT

THE DYNAMICS OF THE GENDER WAGE GAP IN LATE CAREER: A THEORETICAL MODEL

Demirel, Gülce
M.A., Department of Economics
Supervisor: Asst. Prof. Ahmet Arda Gitmez

August 2024

I examine the gender wage gap in the late career stage, utilizing data from the Turkish Household Labor Force Survey spanning 2005 to 2023. By integrating elements of contract theory and narratives, the study offers a novel perspective on discrimination that challenges the traditional statistical discrimination theory. The two-period career concerns model incorporates rational Bayesian updating and narrative-driven extremist updating mechanisms to explore how different belief systems affect wages. The findings reveal that narrative-driven belief updating exacerbates wage variance for women and reinforces gender wage gaps. Empirical analysis supports the model's predictions, showing significantly higher wage variance for women and a widening gender wage gap in the late career stage, particularly in the service sector. The research underscores the critical role of sector-specific characteristics in shaping wage outcomes and highlights the importance of addressing biases formed by social expectations in labor markets.

Keywords: Gender wage gap, Contract Theory, Narratives

ÖZET

GEÇ KARİYER DÖNEMİNDE CİNSİYET ÜCRET FARKLILIKLARININ DİNAMİKLERİ: TEORİK BİR MODEL

Demirel, Gülce
Yüksek Lisans, İktisat Bölümü
Tez Danışmanı: Dr. Öğr. Üyesi Ahmet Arda Gitmez

Ağustos 2024

Bu tez, 2005-2023 yıllarını kapsayan Türkiye Hanehalkı İşgücü İstatistikleri'ni kullanarak, kariyerin geç aşamasına odaklanarak cinsiyet ücret farkını incelemektedir. Sözleşme teorisi ve anlatılar literatürünü bir araya getirerek, geleneksel istatistiksel ayrımcılık teorisine alternatif sunan yenilikçi bir ayrımcılık perspektifi sunmaktadır. İki dönemli kariyer endişeleri modeli, farklı inanç sistemlerinin ücretleri nasıl etkilediğini keşfetmek için rasyonel Bayes güncellemesi ve anlatı temelli aşırı güncelleme mekanizmalarını içermektedir. Bulgular, anlatı temelli güncellemenin kadınlar için ücret varyansını artırdığını ve cinsiyet ücret farkını pekiştirdiğini göstermektedir. Ampirik analiz, modelin öngörülerini destekleyerek, özellikle hizmet sektöründe kariyerin geç aşamasında kadınlar için önemli ölçüde daha yüksek ücret varyansı ve genişleyen cinsiyet ücret farkını ortaya koymaktadır. Araştırma, sektörel özelliklerin ücretler üzerindeki etkisini ve işgücü piyasalarındaki sosyal önyargıların önemini vurgulamaktadır.

Anahtar Kelimeler: Cinsiyet ücret farkı, Sözleşme Teorisi, Anlatılar

ACKNOWLEDGMENTS

First and foremost, I want to extend my heartfelt thanks to my advisor, Arda Gitmez, for their invaluable guidance, insightful feedback, and continuous support throughout this research. Their expertise and encouragement have been instrumental in the completion of this thesis, and I'm truly grateful for their mentorship.

I would like to thank Çağla Ökten and Emre Ekinci for their valuable comments and suggestions.

I would like to express my deepest gratitude to Mahmut Sefa İpek, who has been my friend and a constant source of support. I am incredibly thankful for your belief in me.

Special thanks to my colleagues and fellow students, especially Serkan Karademir and Muhammet Ali Canşı, for their camaraderie, stimulating discussions, and helpful suggestions. Your companionship made this journey more enjoyable and enriching.

Lastly, I would like to extend my heartfelt thanks to my family and friends for their unwavering support and patience throughout this process. Your love and understanding have been my greatest source of strength.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Gender and its implications have long been critical aspects of societal structures, influencing the distribution of opportunities, resources, and power. As we navigate the modern world complexities, the discourse surrounding gender inequality becomes increasingly momentous. The recognition of gender disparities is essential not only for promoting fairness but also for unlocking the full potential of individuals and societies.

The discrepancies between men and women in various contexts have garnered considerable attention in academic, policy, and public discourse over the years. A gender gap is defined as “a gap in any area between women and men in terms of their levels of participation, access, rights, remuneration or benefits” by the European Commission (1998). An important aspect is the gender wage or pay gap. OECD defines the gender wage gap as the difference between men’s and women’s median earnings relative to men’s median earnings. In 2021, it was 11.9% for all OECD countries, 10.4% for the European Union, and 17% for the United States.

While the data offer a numerical portrayal, they alone do not communicate the subtleties of the dynamics at play. To understand the details, we must explore the underlying factors contributing to the observed patterns more profoundly. This calls for a more comprehensive investigation to understand the “why” and "how" behind these disparities.

Understanding late-career dynamics is crucial for a comprehensive analysis of gender wage disparities. According to Ortiz-Ospina, Hasell, and Roser (2018), the gender pay gap is not static throughout a worker's life. Their analysis shows that the gender pay gap is smaller at the beginning of careers but grows significantly in later stages. This trend underscores the persistent challenges women face in achieving wage parity with men, particularly in late-career stages.

The late-career widening of the gender pay gap is linked to several factors, including career interruptions due to family responsibilities and occupational segregation. This dynamic is particularly pronounced in the service sector, where the 'inference' is difficult. In other words, it's more challenging to link high productivity with high effort in the service sector than in manufacturing. That is because there is a less direct correlation between effort and output compared to the manufacturing sector.

The primary objective of this thesis is to construct a formal model to understand the late-career gender wage gap, its factors, and the conditions that magnify their effects. To this end, I integrate insights from two essential strands of economic literature: contract theory and narratives. The model is designed to encapsulate the social norms' influence on wage determination by conceptualizing them as "narratives" that employers inherently accept. These narratives capture societal expectations; specifically, when a woman fails, it must be because she did not work enough, and if she succeeds, she cannot succeed without working.

For instance, consider a scenario in a corporate environment where a female employee misses a project deadline. The prevailing narrative may lead supervisors to assume that her failure was due to a lack of effort rather than considering the stochastic nature of the relationship between success and effort by taking external stochastic factors that contribute to failure into account. For example, unexpected changes in market conditions can disrupt even the most well-planned projects, or a key team member's sudden illness can cause delays. These external, unpredictable elements play a significant role in the success or failure of a task yet are often overlooked in favor of simplistic attributions about individual effort. These factors might be readily accepted explanations for a male counterpart.

Conversely, when a woman reaches an accomplishment, such as securing a major client or successfully leading a project, the narratives automatically attribute her success to high effort rather than acknowledging her inherent ability to seize an opportunity, take advantage of it, and achieve success without exerting much effort by taking risks, strategic thinking, being in the right place at the right time or leadership abilities, unlike men.

These narratives create a paradox where both conflicting premises are considered valid, as follows. On the one hand, they can result in women feeling the constant need to prove their worth through performance as they strive to meet or exceed expectations to counteract negative stereotypes. Women may push themselves to work harder and achieve more to gain the same level of recognition and advancement that men might receive based on perceived potential or standard performance.

On the other hand, this bias can lead women to avoid aiming for high achievements or working hard initially because they anticipate their efforts will not be recognized if they fail. Knowing that failure might be attributed to a lack of effort rather than external factors or bad luck, women might be discouraged from taking risks or striving for significant accomplishments, leading to underperformance.

Some might argue that attributing women's success to luck and men's to effort is an alternative explanation for gender wage disparities; however, this perspective further undermines women's capabilities and contributions. Dismissing their achievements as mere chance fails to recognize the structured and strategic efforts women employ to overcome inherent biases and barriers.

This is the reason behind studying a setting where the market attributes a woman's success to pure effort rather than pure luck. Even this generates a persistent gender wage gap due to the contradictory narrative bias. These skewed perceptions and expectations, coupled with the stochastic nature of effort and success, contribute to the perpetuation of the gender wage gap, particularly in late-career stages where accumulated biases and stereotypes have a more pronounced effect.

In this thesis, I develop a two-period career concerns model to analyze the interactions between workers and the labor market in a competitive setting, with a particular focus on gender differences. The model incorporates elements of stochastic cost heterogeneity and diverse belief-updating mechanisms, comparing rational Bayesian updating for men with narrative-driven extremist updating for women. By examining a competitive labor market where workers can be either behavioral or strategic types, and effort and output are binary variables influenced by stochastic processes, the model explores how different belief-updating mechanisms affect wage determination and worker welfare.

The results of the model reveal that the market's belief-updating mechanism significantly impacts wage dynamics and gender disparities. The narrative-driven extremist updating mechanism, which assumes high output for women is absolutely due to high effort and low output to low effort, leads to higher wage variance for women and perpetuates gender wage gap, especially in the service sector compared to manufacturing which is in line with data. Empirical analysis using Turkish Household Labor Force Survey data supports the theoretical findings, showing that women experience higher wage variance compared to men and a wider wage gap as they progress in their careers.

The structure of this thesis is as follows: Chapter 2 reviews the relevant literature, emphasizing contract theory, narratives, and discrimination. Chapter 3 presents the motivating stylized facts derived from Turkish data. Chapter 4 establishes the theoretical framework. Chapter 5 demonstrates the model's results and empirical implications. Finally, Chapter 6 provides concluding remarks.

CHAPTER 2

RELATED LITERATURE

This thesis is positioned at the intersection of career concerns, narratives, and discrimination literature. Combining insights from these areas of research allows me to conduct a comprehensive exploration of how implicit incentives, shaped by future career prospects, interact with the persuasive power of narratives and the pervasive issue of discrimination. By examining these intersections, I study how narratives influence perceptions and behaviors within labor markets and how these perceptions can create and perpetuate discrimination, especially as gender wage gap in the late-career stage.

2.1 Career Concerns in Contract Theory

Contract theory is a branch of economics that examines the mechanisms and structures through which a principal can incentivize an agent to undertake specific actions that align with the principal's objectives. A critical distinction in this field is between explicit and implicit incentives. Explicit incentives are directly tied to measurable outcomes and include mechanisms such as piece-rate pay and bonuses. Seminal studies in this area, such as those by Holmström (1999), Mirrlees (1999), and Guesnerie and Laffont (1984), have significantly advanced our understanding of how these incentives can be structured to optimize agent performance.

However, implicit incentives, which are less tangible and often linked to future career prospects such as promotions, recognition, and reputation, also play a crucial role in shaping agent behavior. These incentives are not immediately tied to financial rewards but are crucial for understanding long-term motivation and performance within organizations. Implicit incentives are particularly relevant in the context of career concerns, where the agent's current actions are influenced by their potential impact on future career opportunities. This intersection of contract theory and career concerns forms a rich area of research that seeks to understand how implicit incentives can be harnessed to align the interests of agents and principals effectively.

Career concerns refer to individuals' implicit incentives and motivations regarding their future career prospects, which can significantly influence their current behavior and decision-making processes. This is particularly relevant in the context of employment contracts, where the actions of agents (employees) are affected not only by their current compensation but also by the potential impact of their performance on future career opportunities and earnings.

In contract theory, career concerns can alter the traditional principal-agent relationship by introducing additional incentives for agents to perform well beyond immediate financial rewards. Agents may take certain actions, such as exerting high effort to signal their ability and competence to future employers.

Exploring career concerns in contract theory has led to a rich body of literature. This section will focus on seminal works that lay the theoretical foundation, providing a comprehensive overview of how career concerns shape employment contract design.

In that respect, Harris and Holmström (1982) provided a career concerns model that demonstrates how agents' future career prospects influence their current performance. Their model considers an equilibrium framework of long-term labor contracts under incomplete but symmetric information. Incomplete information means that there are certain key variables or parameters that are not fully known to either party. In particular, the exact ability or productivity of the worker is not known initially. Both the firm (principal) and the worker (agent) only have an estimate or a belief about the worker's ability, which gets updated over time as more performance data becomes available. Symmetric information indicates that whatever information is available is equally accessible to both the principal and the agent. In other words, both parties have the same information about the relevant variables at any given point in time. There is no information asymmetry where one party knows more than the other.

While neither the employer nor the employee knows the true productivity of the employee at the outset, they both observe the same performance signals over time and update their beliefs based on these signals. This shared learning process influences the design of employment contracts and wage dynamics, as both parties adjust their expectations and decisions based on the evolving information about the worker's productivity.

One of the key insights from this model is that wages are shown to never decline with age and increase only when a worker's market value surpasses their current wage. This explains why earnings positively correlate to experience even after controlling for productivity. Moreover, the equilibrium contracts derived from their model provide insights into wage structures and the importance of experience in determining wages. This model underscores the role of dynamic learning about a worker's ability and its impact on wage progression.

Further, Dewatripont et al. (1999) extended the career concerns model by illustrating the importance of information structures on implicit incentives. They showed that the flow of information from current actions to market assessment is crucial in determining the nature of these incentives. Information structures refer to the way how information is gathered, processed, and communicated within an organization or market. They determine what information is available to the principal (such as a market) about the agent's (such as an employee) actions and performance. An information structure defines how performance metrics and other relevant data are observed, interpreted, and utilized to decide incentives and rewards.

Different information structures can significantly affect the nature of these incentives. For instance, a sufficient statistic is a summary measure that captures all the relevant information needed to evaluate the agent's performance for decision-making purposes. When the principal uses a sufficient statistic, the agent's performance is evaluated efficiently, and the incentives are aligned optimally. In contrast, Blackwell garbling refers to a situation where the observed information is a noisy or imprecise version of the actual performance data. Garbling can increase or decrease incentives depending on how the noise affects the perceived effort. Another example is inclusive information structures, which provide additional details beyond a primary performance measure, potentially including various performance metrics or qualitative assessments. Adding more detailed information can sometimes enhance incentives by providing a fuller picture of the agent's efforts. Lastly, marginal information structures consider only partial or incremental information about performance, often focusing on changes or differences over time. This can influence incentives by highlighting specific aspects of performance that are most relevant for improvement.

Dewatripont et al. compare different information structures and derive general results that parallel standard results on information structures in the principal-agent model. Their model emphasizes that improved information can either increase or reduce incentives depending on how it affects market beliefs about the agent's future productivity.

The primary distinction between these two models lies in their treatment of information and its impact on career concerns. While Harris and Holmström focus on the dynamic learning process of a worker's ability over time and its influence on wage progression, Dewatripont et al. highlight the role of conjectures and beliefs in shaping the incentives arising from career concerns. Specifically, Harris and Holmström's model predominantly aligns with the concept of sufficient statistics, as discussed in Dewatripont et al. (1999). In Harris and Holmström, the observed output serves as a sufficient statistic, capturing all relevant information needed to update the firm's beliefs about the worker's productivity and determine wage adjustments efficiently. This comprehensive summary measure ensures that the incentives are aligned optimally based on performance. On the other hand, Dewatripont et al.'s model considers how different information structures, including those other than sufficient statistics, about the agent's actions and outcomes influence the market's perception of the agent's ability, thereby affecting the agent's effort decisions. By examining various information structures such as sufficient statistics, Blackwell garbling, and inclusive information structures, Dewatripont et al. provide a comprehensive understanding of how the flow and quality of information shape implicit incentives and career concerns.

This thesis builds on the foundational two-period career concerns model provided by these seminal works. However, unlike previous models that incorporate the agent's talent as a parameter, this study captures the agent's productivity directly through expected output. In this setup, the market (principal) pays the expected output, reflecting a more direct relationship between the agent's current performance and future rewards without the intermediary of talent as a latent variable.

Inspired by the insights from Dewatripont et al. (1999) regarding the impact of different information structures on implicit incentives and career concerns, this study explores the effects of various belief-updating schemes on implicit incentives and wage determination. The model introduced in this thesis examines a competitive labor market where workers can be of either behavioral or strategic type, with effort and output being binary variables influenced by stochastic processes. The market updates its beliefs about worker types that hinge on the unobserved effort based on observed outputs, employing both Bayesian rational updating and narrative-driven extremist updating. By comparing the effects of rational and narrative belief-updating mechanisms, this study aims to shed light on how different approaches to belief updating can significantly affect wage dynamics and worker welfare, thereby extending the theoretical understanding of career concerns in labor economics.

2.2 Narratives in Information Theory

Information theory is a mathematical framework for quantifying information transmission, processing, and storage. It provides fundamental limits on how efficiently information can be communicated and has wide-ranging applications in various fields, including telecommunications, computer science, and economics. In economics, information theory is crucial for understanding how information asymmetry and signal transmission influence decision-making and market outcomes. Building on these foundations, researchers have explored how information can be strategically used to influence behavior and decisions, leading to the study of persuasion in economic contexts.

Persuasion is an endeavor by a sender to influence the beliefs or actions of a receiver to achieve a more favorable outcome from the sender's perspective (Little, 2023). For instance, in the seminal work of Crawford and Sobel (1982), a sender can influence the receiver's actions by sending *messages* without incurring any cost, a concept known as "cheap talk." Additionally, the literature on

Bayesian persuasion, initiated by Kamenica and Gentzkow (2011) and building on Aumann et al. (1995), has flourished over the last decade. Key contributions to this field include the works of Alonso and Câmara (2016), Kolotilin et al. (2017), Ely (2017), and Galperti (2019). In the Bayesian persuasion framework, the sender can influence the receiver through the strategic provision of information, committing to an *experiment* that maps each state realization to a signal distribution. This paper diverges from this tradition by exploring the use of *narratives* rather than information to persuade a receiver. By focusing on how stories and narratives shape beliefs and decisions, this study aims to provide a deeper understanding of the persuasive power of narratives in economic contexts.

Incorporating narratives as a means of persuasion has recently gained prominence, beginning with the work of Shiller (2017, 2020). However, as Barron and Fries (2023) note, there is still no commonly recognized definition of narratives in economics. Aina (2023) emphasizes that narratives function to relate our observations to conclusions about them, influencing how we interpret the world. Consequently, narratives can significantly shape people's behaviors and beliefs.

Different formalizations of narratives have been adopted in the literature, including moral reasoning (Bénabou et al., 2018), directed acyclic graphs (Eliaz and Spiegler, 2020), and likelihood functions (Schwartzstein and Sunderam, 2021). The latter is an ex-post analysis where the receiver has a default narrative, and the sender strategically manipulates this narrative after signal realization. Aina (2023) also models narratives as likelihood functions à la Schwartzstein and Sunderam. Still, Aina's analysis is ex-ante, with the sender committing to their communication strategy before a public signal is realized.

In this study, I do not formalize narratives as mathematical objects, unlike the works mentioned above. Rather, I am inspired by their functions, i.e., being a story that maps observations with results. In my model, the labor market for women and men adopts different updating schemes. For women, the market is boundedly rational in the sense that it does not apply Bayes' Rule when updating its beliefs. Instead, the market believes in a preexisting narrative that when they observe a woman's success, it must absolutely come from high effort. Conversely, when the market observes failure, the narrative absolutely associates it with low effort, and so does the market. Therefore, the resulting posteriors are either zero or one. On the other hand, for men, the market is Bayesian.

2.3 Discrimination

Discrimination in economics refers to the differential treatment of individuals based on characteristics such as race, gender, age, or other attributes rather than their productivity or qualifications. This subsection reviews key theoretical works on discrimination, highlighting various models and their implications that have shaped our understanding of this pervasive issue.

The foundational model of discrimination was introduced by Becker (1957), which was later categorized as "taste-based discrimination." In this model, employers incur disutility from hiring certain groups with marginalized identities, such as women or African-American workers in the original model, despite there being no productivity differences between these groups and historically privileged groups, such as men or White workers[1]. As a result, discriminatory employers either avoid hiring these groups or hire them at lower wages. Becker's model predicts that non-discriminatory employers should outcompete discriminatory ones in a competitive market, as they incur lower costs by not indulging in

[1]There is a convention in the literature starting with Becker that the discriminated worker is denoted as B for black, and the nondiscriminatory worker is denoted as W for white. For the purpose of my thesis, I will adopt a different notation, W for women (discriminated workers), M for men (privileged workers).

their discriminatory preferences. However, this model has been critiqued for not aligning perfectly with empirical observations, often showing persistent wage differentials and segregation in labor markets despite the theoretical expectation of competitive pressures eliminating discrimination.

Another prominent theoretical framework in the study of discrimination is statistical discrimination. Unlike taste-based discrimination, driven by personal prejudice, statistical discrimination arises from using group characteristics as proxies for individual productivity due to imperfect information.

The seminal work of Phelps (1972) introduces the foundational model, later referred to as Phelpsian statistical discrimination. This model explains how employers use group characteristics to make inferences about an individual's productivity when they lack complete information. Statistical discrimination occurs when employers rely on observable characteristics (e.g., race, gender) to make decisions about hiring, wages, or promotions because these characteristics are correlated with productivity-related traits that are difficult to observe directly. The model suggests that even without prejudice, discrimination can arise purely from economic rationality.

In this model, employers face uncertainty about individual productivity and use statistical averages associated with group characteristics to make decisions. Phelpsian discrimination is considered rational behavior in imperfect information because employers use available data to minimize risk. This behavior, while rational, leads to systematic discrimination against groups based on characteristics. Phelps proposes a simple problem of statistically inferring workers' productivities based on imperfectly informative productivity signals and workers' payoff-irrelevant identities. Each worker is identified by a category $j \in \{W, M\}$, where W represents women and M represents men. Their productivity type, θ, is drawn from a normal distribution $N(\mu_j, \sigma_j^2)$, with the parameters μ_j and

σ_j^2 varying according to the worker's category. Employers observe the worker's identity and an imperfect productivity signal given by $\theta = p + \epsilon$ where ϵ is distributed according to $N(0, \sigma_\epsilon^2)$. Given an identity and signal observation, the employer's inference of the worker's expected productivity is:

$$\mathbb{E}[p \mid \theta, j] = \frac{\sigma_j^2}{\sigma_j^2 + \sigma_{\epsilon j}^2}\theta + \frac{\sigma_{\epsilon j}^2}{\sigma_j^2 + \sigma_{\epsilon j}^2}\mu_j$$

To calculate the expected productivity, the observer (employer) uses a weighted average of the observed signal (θ) and the group mean (μ_j). The relative group variances of the productivity distribution and the noise determine the weights. Specifically, the weight assigned to the observed signal is $\frac{\sigma_j^2}{\sigma_j^2+\sigma_{\epsilon j}^2}$, and the weight assigned to the group mean is $\frac{\sigma_{\epsilon j}^2}{\sigma_j^2+\sigma_{\epsilon j}^2}$. That is because the total variation consists of two parts: one part comes from the signal, and the other part comes from the group's productivity distribution. This means that if the signal is very noisy (i.e., $\sigma_{\epsilon j}^2$ is large), more weight is given to the group average. Conversely, if the signal is reliable (i.e., $\sigma_{\epsilon j}^2$ is small), more weight is given to the observed signal.

Suppose each group's productivity signal is perfectly informative, i.e., $\sigma_{\epsilon W}^2 = \sigma_{\epsilon M}^2 = 0$. In that case, the employer fully observes the worker's type, and there is no scope for discrimination. However, if the signal is imperfect, the inferred productivity of a worker may depend on their group identity. Hence, statistical discrimination can create a feedback loop where affected groups have fewer opportunities to demonstrate their true productivity, reinforcing the initial biases. This perpetuates inequality and hinders the affected group's ability to advance economically.

As Onuchic (2023) illustrates, when the two groups possess identical productivity distributions, denoted as $\mu_W = \mu_M$ and $\sigma_W^2 = \sigma_M^2$, but exhibit different levels of noise in productivity signals, discrimination can arise. For instance, if the variance of the signal noise $\sigma_{\epsilon W}^2$ is greater than $\sigma_{\epsilon M}^2$, the signal is noisier for women than for men. Consequently, the employer's productivity inference is more sensitive to the signal for men than women. Specifically, a woman who receives a higher-than-average signal is perceived to have lower expected productivity than a man with the same signal. Conversely, a woman with a lower-than-average signal is seen as more productive than if she were a man.

One important critique of Phelpsian statistical discrimination is that the resulting discrimination is on an individual level rather than a group level (Aigner and Cain, 1977). To see this, when we take the expected value of the expected value above, i.e., $\mathbb{E}[\mathbb{E}[p \mid \theta, j]] = \mu_j$ by the Law of Iterated Expectations. Hence, when the average productivity of women and men is the same, i.e., $\mu_W = \mu_M$, on average, the market pays equal wages to women and men. Therefore, this model's empirical prediction is that there is no group-level discrimination when the average productivities are the same. Moreover, a good measure of discrimination should be formed such that there is aggregate discrimination at a group level on top of the individual level.

There is also another implication of Phelpsian discrimination regarding the variance of expected productivity (or, equivalently, wage) that can be criticized. To find Var($\mathbb{E}[p \mid \theta, j]$), we need to consider the variability of the conditional expectation with respect to the random variable θ. The term $\frac{\sigma_{\epsilon j}^2}{\sigma_j^2+\sigma_{\epsilon j}^2}\mu_j$ is a constant given j, so it does not contribute to the variance. Therefore, we focus on the term $\frac{\sigma_j^2}{\sigma_j^2+\sigma_{\epsilon j}^2}\theta$. Given that θ is a random variable with a distribution, we can use the property of the variance of a linear transformation of a random variable. Specifically, if $\theta \sim N(\mu_j, \sigma_j^2)$), then:

$$\text{Var}\left(\frac{\sigma_j^2}{\sigma_j^2+\sigma_{\epsilon j}^2}\theta\right) = \left(\frac{\sigma_j^2}{\sigma_j^2+\sigma_{\epsilon j}^2}\right)^2 \text{Var}(\theta)$$

Since $\theta \sim N(\mu_j, \sigma_j^2)$, we have Var$(\theta) = \sigma_j^2$. Therefore:

$$\text{Var}\left(\mathbb{E}[p \mid \theta, j]\right) = \left(\frac{\sigma_j^2}{\sigma_j^2+\sigma_{\epsilon j}^2}\right)^2 \sigma_j^2$$

So, the variance of $\mathbb{E}[p \mid \theta, j]$ is:

$$\text{Var}\left(\mathbb{E}[p \mid \theta, j]\right) = \frac{\sigma_j^6}{(\sigma_j^2 + \sigma_{\epsilon j}^2)^2}$$

Notice that the wage variance decreases in $\sigma_{\epsilon j}^2$. Therefore, if women's productivity distribution is noisier than men, i.e., $\sigma_{\epsilon W}^2 > \sigma_{\epsilon M}^2$, this model implies that the variance of the wage for women is smaller than men, i.e., $\text{Var}\left(\mathbb{E}[p \mid \theta, W]\right) < \text{Var}\left(\mathbb{E}[p \mid \theta, M]\right)$.

I introduced a different source of discrimination than the fact that productivity inference is more sensitive to the signal for men than women whenever the signal is nosier for women, given the same productivity distributions of both genders. In my model, although the signal's noise is the same for both genders in addition to the same productivity distributions, we still have sensitivity differentials in the opposite direction. The market is more sensitive to women's signals because it adopts a narrative-based updating scheme that absolutely associates high output with high effort and low output with low effort. For men, the market applies Bayes' Rule to give meaning to signals. As a result, over-sensitivity to women also perpetuates persistent discrimination.

CHAPTER 3

MOTIVATING STYLIZED FACTS

This study utilizes cross sectional data from the Turkish Household Labor Force Survey (HLFS) spanning from 2005 to 2023, encompassing a total of 1,768,326 individual observations. The dataset includes variables such as gender, age, the NACE 2 sector of employment, weekly hours worked, monthly earnings (including bonuses), and the survey year. This comprehensive dataset allows for an in-depth analysis of wage dynamics across different demographic groups and sectors over an extended period.

To analyze the hourly wage distribution, I performed several steps to standardize the earnings to an hourly rate, considering the variability in the number of hours worked each week:

$$\text{Hourly Wages} = \frac{\text{Monthly Earnings}}{\text{Weekly Hours Worked} \cdot 4.3}$$

Then, I adjusted the hourly wages for inflation to ensure consistency across the wide range of observation years, using 2003 as the base year (2003=100):

$$\text{CPI Adjusted Hourly Wages} = \frac{\text{Hourly Wages}}{\text{CPI}_{\text{survey year}}} \cdot 100$$

Below, Table 3.1 provides the mean, median, and variance of CPI-adjusted hourly wages by gender and the gender wage gap by age groups, specifically focusing on the early career (ages 35-45) and late career (ages 45-55) phases.

	Age 35-45		Age 45-55	
	Women	Men	Women	Men
Mean	4.42	4.15	4.35	4.45
Median	2.85	2.95	2.66	3.20
Variance	54.8	40.4	71.0	47.6
Gender Wage Gap	3.2%		17.1%	

Table 3.1: Summary Statistics of CPI-Adjusted Hourly Wages by Gender and Age Group and Gender Wage Gap by Age Group in Turkiye between 2005-2023

The mean row indicates that women in their early careers earn a higher mean hourly wage (4.42) than men (4.15) in Turkiye. Conversely, in the late career phase, women's mean hourly wage slightly decreases to 4.35, while men's hourly wage increases to 4.45. A comparison of the median wages for both career stages reveals that women earn less than men, and as women progress in their careers, their median wage declines, whereas for men, it rises. This results in a wider gap in median wages during the late career stage.

Additionally, the variance row demonstrates that women experience significantly higher variance in hourly wages in both early (54.8) and late career (71.0) stages compared to men, whose variances are 40.4 and 47.6, respectively. These findings suggest that while women may initially earn more than men on average in Turkiye, their wage variance increases substantially over time, indicating a more dispersed wage distribution for women compared to men as they advance in their careers.

Furthermore, the gender wage gap, calculated as the difference between the median earnings of men and women divided by the median earnings of men, reveals a gap of 3% in the early career stage, which escalates to 17% in the late career stage. These figures indicate a more pronounced gender wage gap in the late career stage, meriting further investigation. This thesis explores this significant increase in the gender wage gap as women advance in their careers.

To further explore the wage dynamics across different age groups and genders, I present the CPI-adjusted hourly wage distributions for two age groups: 35-45 and 45-55. The following histograms provide a visual representation of the wage distributions for men and women in these age categories.

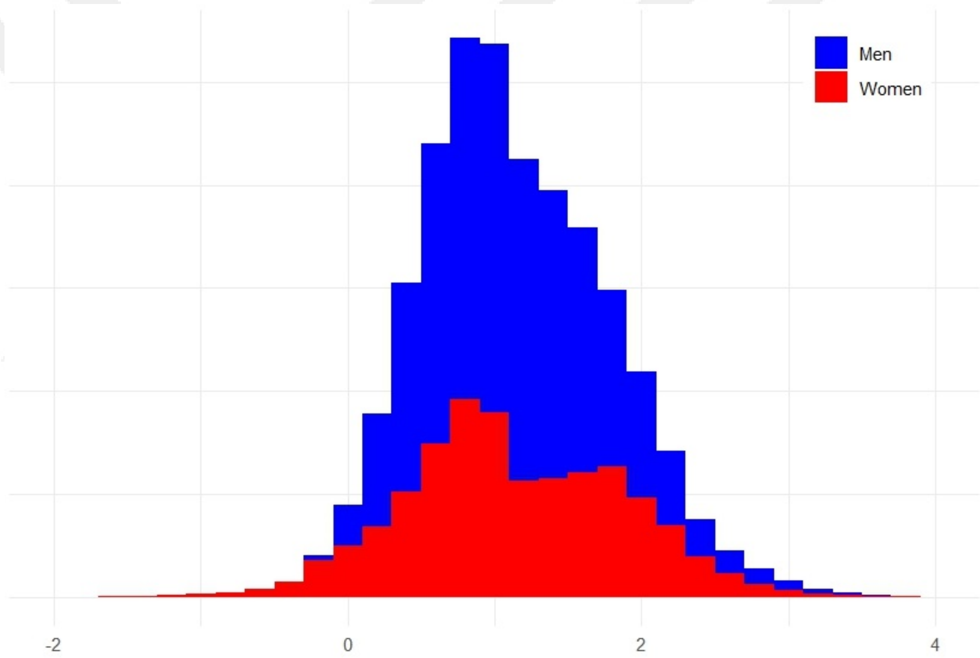

Figure 3.1: Log CPI-Adjusted Hourly Wages for Ages 35-45 in Turkiye between 2005-2023

In both Figure 3.1 and Figure 3.2, I use logarithms to make the visualization more insightful. Moreover, men's wages (blue) are more frequent, indicating more observations than women's (red) and low female employment rates. For each bin in the histograms, the observations of men are higher than those of women except for the extreme values.

Figure 3.1 depicts the log CPI-adjusted hourly wage distribution for the early

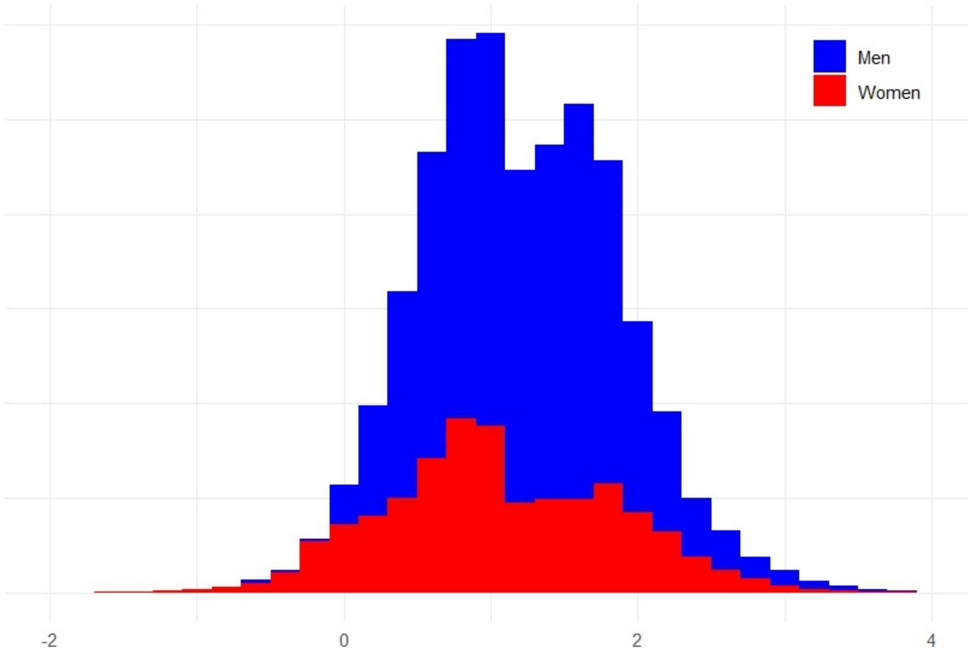

Figure 3.2: Log CPI-Adjusted Hourly Wages for Ages 45-55 in Turkiye between 2005-2023

career group (ages 35-45). While both distributions for men and women are relatively centered around the median value, the distribution for women appears more spread out given the lower frequency, suggesting greater wage variance for women at this career stage.

Figure 3.2 presents the log of CPI-adjusted hourly wages for individuals aged 45-55, differentiated by gender. The histogram reveals a bimodal distribution, particularly pronounced for men compared to women. This two-peak nature indicates that as individuals progress into the later stages of their careers, there is a greater divergence in wage levels, especially among men. The frequency of observations for men at the higher end of the wage distribution is notably greater, contributing to a higher variance in this stage of their careers, as reflected in Table 3.1. While the distribution is less pronounced, there is still observable spread for women, though the frequency remains lower compared to men. This pattern highlights the increasing wage variance as careers advance, with a more marked effect seen in male wages. Nonetheless, despite the higher frequency of male observations, the variance in women's wages is higher. While male observations are concentrated in a narrower range, female observations are spread over the same area with fewer observations, indicating higher variance for women in this career stage.

Both Table 3.1 and Figures 3.1 & 3.2 provide insights into the distribution of CPI-adjusted hourly wages across different career stages and between genders. By examining the data, we uncover critical patterns that highlight the dynamics of wage variability as careers progress. Specifically, we identify two primary stylized facts that emerge from the data:

Stylized Fact 1. *The variance of CPI-adjusted hourly wages increases with the career stage, for both women and men.*

This observation justifies employing a dynamic setting to model gender wage differentials. In this framework, the market sets a fixed wage in the initial period, resulting in zero variance for early-career wages. However, as individuals progress into the second period, wages start to vary based on performance and other factors, leading to a positive variance. This aligns with empirical observations, showing that wage variance not only exists but also increases in the later stages of one's career. To simplify what we observe in the data and to explain the gender wage differential dynamics in the late career—which is the main focus of this thesis—I utilized the career concerns model.

Stylized Fact 2. *Variance of women's CPI-adjusted hourly wages is higher than men's in both early and late careers. Additionally, the variance increases with the career stage, more for women than men.*

This finding is significant because the empirical implication of Phelps' (1972) theory of statistical discrimination does not account for the higher wage variance observed in the discriminated group (women) compared to the non-discriminated group (men), as discussed in the previous section. However, our findings indicate a higher wage variance for women than men, both in early and late career stages. This discrepancy highlights the necessity for a more comprehensive model that can better capture the observed higher variance in wages among the discriminated group.

Additionally, the data reveals sectoral differences in wage distributions. The sectors are divided into two main categories: manufacturing and service industry, following the Statistical Classification of Economic Activities in the European Community (NACE Rev.2). The manufacturing sector corresponds to NACE codes 10-33, which include industries such as manufacturing food products and textiles. The service sector corresponds to NACE codes 45-96, encompassing economic activities such as wholesale and retail trade, financial service activities, education, and human health activities. By analyzing the NACE 2 sector data, we can observe how wage variance differs across industries, providing further insight into gender wage disparities within specific economic contexts. This sectoral analysis allows us to identify particular industries where the wage gap and variance are more pronounced.

	Age 35-45		Age 45-55	
Sector	Women	Men	Women	Men
Manufacturing	21.6	24.8	18.9	39.6
Service	69.4	51.4	98.8	56.4

Table 3.2: Variance of CPI-Adjusted Hourly Wages in Turkiye between 2005-2023 by Age, Gender, and Sector

Table 3.2 presents the variance of hourly wages segmented by age group, gender, and sector (manufacturing and service). The data reveals notable differences in wage variability across these dimensions. For the age group 35-45, we observe that women in the manufacturing sector have a variance of 21.6, whereas men exhibit a slightly higher variance of 24.8. However, the variance in the service sector is significantly greater for both genders, with women showing a variance of 69.4 compared to 51.4 for men. This trend persists into the late career stage (ages 45-55), where the variance in the manufacturing sector for women decreases to 18.9, while it increases substantially for men to 39.6. In contrast, the service sector again exhibits higher variability, with women's wage variance escalating to 98.8, markedly higher than men's variance of 56.4. This data underscores the higher wage dispersion experienced by women, particularly in the service sector, as they progress in their careers. The substantial increase in variance for women in the service sector during the late career stage is particularly noteworthy, indicating a more pronounced spread in wage distribution for women compared to men in the same sector and age group.

Stylized Fact 3. *In the late career stage, the variance of women's CPI-adjusted hourly wages in the service sector is substantially higher than in the manufacturing sector. This disparity is also significantly greater when compared to men's variance in the same sectors.*

The significant difference in wage variance between the service and manufacturing sectors, particularly for late-career women, is an important finding. This difference underscores the role of sector-specific characteristics in shaping wage outcomes. The main difference between the service and manufacturing sectors is that the service sector suffers from an 'inference problem.' That is, in the service sector, it is more difficult to affiliate high productivity with high effort compared to the manufacturing sector. Then, Stylized Fact 3 implies that in sectors where the output is less directly tied to effort, there is a greater spread in women's wages than in men's. This disparity is critical as it highlights the potential for greater inequality and instability in wage outcomes for women in the service sector as they advance in their careers. The model introduced in the next section is designed to capture these dynamics by considering the unique characteristics of different sectors. By incorporating sector-specific sensitivities of output to the effort, the model aims to provide a more accurate representation of the mechanisms driving gender wage disparities, particularly in the service sector. The model introduced in the next section will be justified by considering how these sector-specific differences in sensitivity to effort affect late-career wage spread.

	Age 35-45	Age 45-55
Sector	Gender Wage Gap	Gender Wage Gap
Manufacturing	0.126	0.119
Service	-0.065	0.165

Table 3.3: Gender Wage Gap in Turkiye between 2005-2023 by Age and Sector

Table 3.3 presents the gender wage gap by age and sector, highlighting distinct disparities between manufacturing and service sectors across different career stages. For individuals aged 35-45, the gender wage gap is positive in the manufacturing sector (0.126), indicating that men earn more than women. Conversely, in the service sector, the gender wage gap is negative (-0.065), suggesting that women earn slightly more than men. A negative gender wage gap in favor of women is also a common phenomenon in middle-income countries with low female labor force participation (Ortiz-Ospina et al., 2018). For instance, in 2022, men's employment rate in Turkiye surpassed women's by 34.6 percentage points, compared to a 9.8 percentage point difference in the EU (Euronews, 2024). Conversely, in the late career phase, women's mean hourly wage slightly decreases to 4.35, whereas men's hourly wage increases to 4.45. As individuals progress into the late career stage (age 45-55), the gender wage gap in manufacturing slightly decreases to 0.119, while it significantly increases in the service sector to 0.165.

Stylized Fact 4. *The gender wage gap in the service sector increases significantly from early to late career stages, whereas the gap in the manufacturing sector remains relatively stable.*

These findings reveal that the gender wage gap not only persists but also widens in the service sector as careers advance. This indicates that women in the service sector experience greater wage disparities than their male counterparts in later career stages, which may reflect underlying sectoral differences in how wages are determined and the sensitivity of output to effort. The substantial increase in the service sector's gender wage gap in the late career stage underscores the need to consider sector-specific dynamics when addressing gender wage disparities. The subsequent model introduced will further explore these aspects, particularly focusing on the higher wage variance observed in the service sector.

CHAPTER 4

MODEL

In this chapter, I introduce a two-period career concerns model to analyze the interactions between workers and the observer (the market) in a competitive labor market. Through the model, I explore the implications of stochastic cost heterogeneity and various belief-updating mechanisms the market employs on wage determination and welfare.

The basic elements of the model are as follows. There are two sides of the market: employers and employees. Each employee makes the binary choice of exerting effort or not. An employee can be (i) a behavioral type, in which case they put effort, or (ii) a strategic type, in which case they decide to exert effort or not. For the strategic types, exerting effort is associated with a positive cost. The players live for two periods. In each period, the workers choose to exert effort e, and the market observes the output y and pays the workers. Effort and output are binary variables, taking on either high or low values. The relationship between output and effort is stochastic, governed by a commonly known production function, denoted as $\mathbb{P}r(y \mid e)$.

In the first period, the market pays a sunk wage to all workers prior to observing output. Then, strategic types choose an effort level based on the model's parameters, whereas behavioral types consistently exert high effort. Subsequently, the first period's output is observed. Some workers succeed by producing high output, while others fail by generating low output. I will interchangeably use "succeeding" and "producing high output" as well as "failing" and "producing low output" hereafter. Additionally, "working" will be used interchangeably with "exerting high effort," and "not working" will be used interchangeably with "putting in low effort."

The market then updates its beliefs regarding a worker's type, or equivalently, the probability that a worker will exert low effort in the second period, based on the output realization in the first period. These two probabilities converge because there is no third period. Consequently, it is never optimal for strategic types to work in the second period. They may have an incentive to mimic behavioral types in the first period to persuade the market that they will work in the second period, thereby increasing their second-period wage, which is determined by the probability of succeeding conditional on the first period's output realization.

In the second period, the market pays the probability of producing high output conditional on the first period's output. Then, the workers pick an effort level. In particular, behavioral types choose to work, and strategic types exert low effort since there is no tomorrow and high effort induces a cost. Lastly, the second period's output is drawn.

Certain elements in this setup give rise to information asymmetries. Firstly, the market has a prior belief about the proportion of behavioral and strategic types but does not particularly know who is who. On the contrary, workers are aware of their types. Secondly, due to the stochastic nature of production and effort, the market cannot essentially distinguish the employees who worked and succeeded from those who did not work but still succeeded. It can only observe the final output, not the effort.

The belief system of the market is of considerable importance. The market may adopt a rational approach, updating its beliefs according to Bayes' Rule, or it may exhibit bounded rationality by believing in “narratives" and updating its beliefs in an extremist fashion. Specifically, upon observing success, the market that believes in narratives unequivocally attributes it to the high effort, fully disregarding the possibility that high output may arise from not working. Conversely, when observing failure, the market updates its belief about the probability of workers exerting effort to zero, as if it were impossible for a worker to exert high effort and still fail.

I refer to this type of belief updating approach as believing in narratives because narratives are the stories that explain observations. However, they tend to be extremists. Narratives simplify the complex reality by creating a direct and absolute cause-effect relationship, which can be easily comprehended but often overlooks real-world phenomena' nuances and stochastic nature. In the labor market context, the narratives market might interpret a worker's success as absolute proof of high effort, thereby ignoring any random factors or luck that could also lead to high output. Similarly, if a worker fails, the market's narrative-driven belief system would assume the failure is due to a lack of effort, discounting any other possible reasons such as external conditions or sheer bad luck. This extremist approach to belief updating can significantly impact the market's wage-setting behavior and overall equilibrium, as it fails to account for the probabilistic and multifaceted nature of effort and output.

The following sections will systematically explore the implications of information asymmetries and diverse belief systems within the two-period career concerns model. First, the benchmark model will be presented, assuming no information asymmetries, and the equilibrium will be characterized. This foundational analysis will provide a baseline against which more complex scenarios can be compared. Next, the model will be extended to incorporate asymmetric information where the market updates its beliefs using Bayesian inference. The analysis will focus on the equilibrium behavior and the impact of Bayesian updating on the market's decisions and worker incentives. Finally, the model will consider asymmetric information in a setting where the market exhibits bounded rationality. Here, the belief updating process will deviate from the Bayesian framework, reflecting more extreme responses to observed outcomes. The implications of this boundedly rational behavior for equilibrium conditions and worker strategies will be thoroughly examined. Each section will build on the previous one, providing a comprehensive understanding of how varying belief frameworks influence the labor market dynamics within the career concerns framework.

4.1 Benchmark Model with No Information Asymmetries

A two-period interaction between the workers and the market is considered within the framework of a career concerns model, with a modification: the market possesses perfect knowledge regarding whether a worker is of behavioral or strategic type, in addition to the cost distribution among the population. Denote the cost of working by $c \in (0, \infty)$, with $\mathbb{P}\text{r}(\text{strategic type}) = \theta$, indicating that $\theta \in [0, 1]$ proportion of the population strategically decides to exert effort which incurs a positive cost c. Conversely, the proportion of behavioral types is given by $\mathbb{P}\text{r}(\text{behavioral type}) = 1 - \theta$ who always choose to exert high effort which is not associated with a cost.

The sequence of events unfolds as follows. At time $t = 1$, the market initially disburses a wage, denoted as w_1, which is considered sunk. Subsequently, workers select an effort level, denoted as $e_1 \in \{e_H, e_L\}$. The behavioral types will choose $e_1 = e_H$. Knowing that the market knows they are strategic and expects them not to work due to its cost, strategic types will not work and choose $e_1 = e_L$. Following this, the output for the first period, denoted as $y_1 \in \{y_H, y_L\}$, is realized according to a fixed and known distribution:

$$\mathbb{P}\mathrm{r}(y_H \mid e_H) = \alpha$$
$$\mathbb{P}\mathrm{r}(y_H \mid e_L) = \beta$$

where $\alpha, \beta \in [0,1]$. This implies that $\mathbb{P}\mathrm{r}(y_L \mid e_H) = 1-\alpha$ and $\mathbb{P}\mathrm{r}(y_L \mid e_L) = 1-\beta$. Thus, it is mutually understood by the players that the probability of success given high effort is α, while the probability of failure given high effort is $1-\alpha$. Similarly, for low effort, the probability of producing high output given low effort is β, and the probability of low output given low effort is $1-\beta$.

Assumption 1 *Effort increases the chance of success, i.e., $\alpha > \beta$.*

This assumption ensures a shared understanding among the players that increased worker effort correlates with a higher probability of success. This assumption will be maintained throughout this research.

At time $t = 2$, the market compensates the workers based on the expected output, denoted as $w_2 = \mathbb{E}[\ y_2 \mid e_1]$, which remains fixed and independent of the first period's output due to the market's perfect knowledge of which effort level each type will choose. By normalizing $y_H = 1$ and $y_L = 0$, we obtain $\mathbb{E}[\ y_2 \mid e_1] = \mathbb{P}\mathrm{r}(y_H \mid e_1)$, thus the wage in the second period corresponds to the probability of success in the second period.

In other words, due to the market's perfect knowledge, w_2 directly depends on the effort exerted (the known state of the world) rather than the output (the signal, which the principal would typically use to infer the unknown effort). Since the types are known, the effort levels are also known. Workers will choose their effort levels accordingly, knowing the market knows their types. Consequently, the market's initial expectations will be realized. Given this foreknowledge, the market will not consider the output realization in the first period, as it already possesses complete information about the state of the world. Then, the workers will select e_2: behavioral types will choose e_H by design. Strategic types will strictly prefer not to work, i.e., e_L, due to the cost associated with working and the fact that there is no next period to consider.

In a typical two-period career concerns model, the equilibrium is characterized by (w_2^*, e_1^*). This is since w_1 is sunk, and choosing $e_2 = e_H$ is never optimal unless working is costless. The market cannot incentivize high-cost types to exert effort in the second period when low-cost types are already doing so, as the second period's wage is determined prior to the workers' choice of effort. Consequently, the critical determinant of the equilibrium is e_1 rather than e_2[1].

Proposition 1 *In the Benchmark Model with No Information Asymmetries, the Subgame Perfect Nash Equilibrium is characterized by the following strategies:*

- *The market sets w_2 based on the observed effort level in the first period:*

$$w_2^* = \begin{cases} \alpha & \textit{if} \quad e_1 = e_H \\ \beta & \textit{if} \quad e_1 = e_L \end{cases}$$

- *The strategic workers with $c > 0$ choose high effort ($e_1^* = e_H$) if $\alpha - \beta > c$; otherwise, they choose low effort ($e_1^* = e_L$).*

Proof. To demonstrate that the proposed equilibrium strategies constitute a Subgame Perfect Nash equilibrium, it must be shown that the workers are playing best responses given that the market pays the expected output, i.e., $\mathbb{E}[\, y_2 \mid e_1] = \mathbb{P}\mathrm{r}(y_H \mid e_1)$.

Assume the market sets the second-period wage w_2^* according to the observed

[1]Although e_2 is not a part of the equilibrium, it provides critical information regarding cost distributions. The significance of this will be elucidated in the following sections.

effort level in the first period. For workers, the payoff is $\alpha - c$ if they exert high effort and β if they exert low effort. Hence, high-cost workers will exert high effort only if $\alpha - c > \beta$. This condition simplifies to $\alpha - \beta > c$.

Given Assumption 1, $\alpha > \beta$, the following holds:

- If $\alpha - \beta > c$, strategic workers will exert high effort ($e_1^* = e_H$).
- If $\alpha - \beta \leq c$, strategic workers will exert low effort ($e_1^* = e_L$).

Given the market's perfect knowledge, the second-period wage w_2 is determined by the observed effort level in the first period:

$$w_2^* = \begin{cases} \alpha & \text{if } e_1 = e_H \\ \beta & \text{if } e_1 = e_L \end{cases}$$

Thus, the equilibrium strategies are self-enforcing: strategic workers maximize their payoffs by choosing the effort level based on their cost c relative to $\alpha - \beta$. The market's wage-setting mechanism accurately reflects these effort levels, ensuring that no player has an incentive to deviate from their strategy. ■

4.2 Asymmetric Information with Bayesian Market

I now shift the analysis to the standard two-period career concerns framework to model the interaction between workers and the market under the condition that the market cannot observe the workers' cost of effort. The market holds a prior belief that $\mathbb{P}\text{r}(\text{strategic type}) = \theta$ and $\mathbb{P}\text{r}(\text{behavioral type}) = 1 - \theta$. The market updates its belief using the Bayes Rule after observing y_1. Let the posterior probabilities be denoted as follows:

$$\hat{\theta} = \mathbb{P}\text{r}(\text{strategic type} \mid y_1 = y_H)$$
$$\tilde{\theta} = \mathbb{P}\text{r}(\text{strategic type} \mid y_1 = y_L)$$

Thus, $\hat{\theta}$ represents the probability that the market believes a given worker is of a strategic type, conditional on the worker's success. Conversely, $1 - \hat{\theta} = \mathbb{P}\text{r}(\text{behavioral type} \mid y_1 = y_H)$ represents the probability that the market believes a given worker is of a behavioral type, given success. Similarly, $\tilde{\theta}$ represents the probability that the market believes a given worker is a strategic type, conditional on the worker's failure. Alternatively, $1 - \tilde{\theta} = \mathbb{P}\text{r}(\text{behavioral type} \mid y_1 = y_L)$ represents the probability that the market believes a given worker is of behavioral type, given failure. Since low-cost types always choose to work, and working increases the probability of success by Assumption 1, we have $\tilde{\theta} > \hat{\theta}$. Therefore, a worker who fails is more likely to be a strategic type.

On the other hand, the relationship between cost and effort in both periods can be expressed as:

$$\mathbb{P}\text{r}(\text{strategic type} \mid y_1) \neq \mathbb{P}\text{r}(e_1 = e_L \mid y_1)$$
$$\mathbb{P}\text{r}(\text{behavioral type} \mid y_1) \neq \mathbb{P}\text{r}(e_1 = e_H \mid y_1)$$
$$\mathbb{P}\text{r}(\text{strategic type} \mid y_1) = \mathbb{P}\text{r}(e_2 = e_L \mid y_1)$$
$$\mathbb{P}\text{r}(\text{behavioral type} \mid y_1) = \mathbb{P}\text{r}(e_2 = e_H \mid y_1)$$

This indicates that the probabilities associated with being a behavioral or strategic type are directly linked to the effort levels exerted by the workers only in the second period. In the first period, however, high-cost types may be incentivized to exert high effort despite the cost. Consequently, it cannot be conclusively stated that all workers who choose $e_1 = e_H$ are behavioral types. Additionally, not all strategic types are necessarily those who do not choose to work in the first period, as some high-cost types may pretend to be behavioral types by choosing

$e_1 = e_H$. Therefore, the proportion of strategic workers may not necessarily equal that of workers who choose $e_1 = e_L$, leading to inequalities in the first period [2].

Due to information asymmetry, *all* strategic types may imitate behavioral types by choosing $e_1 = e_H$ to persuade the market that their effort is costless. Consequently, the market might be led to believe that these workers will exert effort at time $t = 2$, which, by Assumption 1, increases $w_2 = \mathbb{E}[\, y_2 \mid y_1]$ and subsequently raises their expected payoff. This scenario results in a *pooling equilibrium* where all types exert high effort in the first period. Alternatively, *some* strategic types might exert high effort in the first period to achieve the same objective, leading to a *semi-separating equilibrium* where strategic types mix between the strategies e_H and e_L in the first period. It is important to note that the cost must not be prohibitively high for a strategic type to exert high effort. Otherwise, choosing $e_1 = e_H$ will never be optimal, resulting in a *separating equilibrium* where high-cost types do not exert effort at all.

The decision of strategic workers to exert effort depends on the equilibrium that the market conjectures, which subsequently influences their incentives. On the other hand, behavioral workers are not crucial to the equilibrium analysis as they are behavioral and consistently choose to work. Additionally, although posteriors are essentially defined as the same, they vary in different equilibria.

4.2.1 Rational Separating Equilibrium

In career concerns setup, the equilibrium is characterized by (e_1^*, w_2^*), as w_1 is sunk and $e_2 = e_L$ for all strategic workers. We consider an equilibrium where the market incentivizes strategic types to choose low effort in the first period,

[2] Note that equalities may arise in the first period if the market conjectures a separating equilibrium in which only behavioral types exert effort, while none of the strategic types do so.

i.e., $e_1^* = e_L$. Following are the three parts of the equilibrium.

Posteriors

Recall that the posteriors are

$$\hat{\theta} = \mathbb{P}\text{r}(\text{strategic type} \mid y_1 = y_H)$$

and

$$\tilde{\theta} = \mathbb{P}\text{r}(\text{strategic type} \mid y_1 = y_L).$$

By the Bayes' Rule, we have the following:

$$\hat{\theta} = \mathbb{P}\text{r}(\text{strategic type} \mid y_1 = y_H) = \frac{\mathbb{P}\text{r}(\text{strategic type}) \cdot \mathbb{P}\text{r}(y_1 = y_H \mid \text{strategic type})}{\mathbb{P}\text{r}(y_1 = y_H)}$$

$$\tilde{\theta} = \mathbb{P}\text{r}(\text{strategic type} \mid y_1 = y_L) = \frac{\mathbb{P}\text{r}(\text{strategic type}) \cdot \mathbb{P}\text{r}(y_1 = y_L \mid \text{strategic type})}{\mathbb{P}\text{r}(y_1 = y_L)}$$

Since the high or low output can come from either high effort or low effort, we have:

$$\begin{aligned}
\mathbb{P}\text{r}(y_1 = y_H) &= \mathbb{P}\text{r}(e_1 = e_L) \cdot \mathbb{P}\text{r}(y_1 = y_H \mid e_1 = e_L) \\
&\quad + \mathbb{P}\text{r}(e_1 = e_H) \cdot \mathbb{P}\text{r}(y_1 = y_H \mid e_1 = e_H) \\
&= \theta \cdot \beta + (1 - \theta) \cdot \alpha \\
\mathbb{P}\text{r}(y_1 = y_L) &= \mathbb{P}\text{r}(e_1 = e_L) \cdot \mathbb{P}\text{r}(y_1 = y_L \mid e_1 = e_L) \\
&\quad + \mathbb{P}\text{r}(e_1 = e_H) \cdot \mathbb{P}\text{r}(y_1 = y_L \mid e_1 = e_H) \\
&= \theta \cdot (1 - \beta) + (1 - \theta) \cdot (1 - \alpha)
\end{aligned}$$

Since we are now considering a *rational separating equilibrium*, the posteriors particular to this setup are going to be denoted as $\hat{\theta}_{RS}$ and $\tilde{\theta}_{RS}$. This distinction is important because $\mathbb{P}\text{r}(\text{strategic type}) = \mathbb{P}\text{r}(e_1 = e_L) = \theta$ for this particular equilibrium and for this particular posteriors: $\hat{\theta}_{RS}$ and $\tilde{\theta}_{RS}$. That is because the market conjectures an equilibrium where all strategic workers with $c > 0$ choose

$e_1 = e_L$. Consequently, we have

$$\mathbb{P}\text{r}(y_1 = y_H \mid \text{strategic type}) = \mathbb{P}\text{r}(y_1 = y_H \mid e_1 = e_L) = \beta$$
$$\mathbb{P}\text{r}(y_1 = y_L \mid \text{strategic type}) = \mathbb{P}\text{r}(y_1 = y_L \mid e_1 = e_L) = 1 - \beta$$

Therefore,

$$\hat{\theta}_{RS} = \frac{\theta \cdot \beta}{\theta \cdot \beta + (1-\theta) \cdot \alpha}$$

$$\tilde{\theta}_{RS} = \frac{\theta \cdot (1-\beta)}{\theta \cdot (1-\beta) + (1-\theta) \cdot (1-\alpha)}$$

Notice that $\hat{\theta}_{RS} < \tilde{\theta}_{RS}$ by Assumption 1.

Wages

Unlike the benchmark model with perfect information, the market cannot directly observe effort. Consequently, the market uses the first period's output as a signal to calculate the second period's expected output. After normalizing $y_H = 1$ and $y_L = 0$, the rational market pays $w_2 = \mathbb{P}\text{r}(y_2 = y_H \mid y_1)$. In other words, the second period's wage is a function of the first period's output rather than the first period's effort.

$$w_2 = \begin{cases} \mathbb{P}\text{r}(y_2 = y_H \mid y_1 = y_H) & \text{if } y_1 = y_H \\ \mathbb{P}\text{r}(y_2 = y_H \mid y_1 = y_L) & \text{if } y_1 = y_L \end{cases}$$

Like before, the high output in the second period can come from both effort levels. Hence,

$$\begin{aligned} \mathbb{P}\text{r}(y_2 = y_H \mid y_1 = y_H) &= \mathbb{P}\text{r}(e_2 = e_H \mid y_1 = y_H) \cdot \mathbb{P}\text{r}(y_2 = y_H \mid e_2 = e_H, y_1 = y_H) \\ &\quad + \mathbb{P}\text{r}(e_2 = e_L \mid y_1 = y_H) \cdot \mathbb{P}\text{r}(y_2 = y_H \mid e_2 = e_L, y_1 = y_H) \\ \mathbb{P}\text{r}(y_2 = y_H \mid y_1 = y_L) &= \mathbb{P}\text{r}(e_2 = e_H \mid y_1 = y_L) \cdot \mathbb{P}\text{r}(y_2 = y_H \mid e_2 = e_H, y_1 = y_L) \\ &\quad + \mathbb{P}\text{r}(e_2 = e_L \mid y_1 = y_L) \cdot \mathbb{P}\text{r}(y_2 = y_H \mid e_2 = e_L, y_1 = y_L) \end{aligned}$$

In addition to conjecturing a separating equilibrium where all workers with

$c > 0$ choose $e_1 = e_L$, the market also perfectly anticipates that these workers will choose $e_2 = e_L$ since there is no subsequent period. Consequently, we have:

$$\mathbb{P}\mathrm{r}(e_2 = e_H \mid y_1 = y_H) = \mathbb{P}\mathrm{r}(\text{behavioral type} \mid y_1 = y_H) = 1 - \hat{\theta}_{RS}$$
$$\mathbb{P}\mathrm{r}(e_2 = e_L \mid y_1 = y_H) = \mathbb{P}\mathrm{r}(\text{strategic type} \mid y_1 = y_H) = \hat{\theta}_{RS}$$

$$\mathbb{P}\mathrm{r}(e_2 = e_H \mid y_1 = y_L) = \mathbb{P}\mathrm{r}(\text{behavioral type} \mid y_1 = y_L) = 1 - \tilde{\theta}_{RS}$$
$$\mathbb{P}\mathrm{r}(e_2 = e_L \mid y_1 = y_L) = \mathbb{P}\mathrm{r}(\text{strategic type} \mid y_1 = y_L) = \tilde{\theta}_{RS}$$

Since the second period's output is solely a function of the second period's effort, it is independent of the first period's output, conditional on the second period's effort. Therefore, the final consideration before explicitly characterizing the wages pertains to the following:

$$\mathbb{P}\mathrm{r}(y_2 = y_H \mid e_2 = e_H, y_1 = y_H) = \mathbb{P}\mathrm{r}(y_2 = y_H \mid e_2 = e_H) = \alpha$$
$$\mathbb{P}\mathrm{r}(y_2 = y_H \mid e_2 = e_L, y_1 = y_H) = \mathbb{P}\mathrm{r}(y_2 = y_H \mid e_2 = e_L) = \beta$$

Therefore in the separating equilibrium, the market pays

$$w_2^{RS} = \begin{cases} \hat{\theta}_{RS} \cdot \beta + (1 - \hat{\theta}_{RS}) \cdot \alpha & \text{if } y_1 = y_H \\ \tilde{\theta}_{RS} \cdot \beta + (1 - \tilde{\theta}_{RS}) \cdot \alpha & \text{if } y_1 = y_L \end{cases}$$

Incentive Compatibility

Lastly, verifying the behavior in the first period under the conjectured equilibrium is necessary. Specifically, the workers must choose effort levels corresponding to the established beliefs and wages. The behavioral workers will choose $e_1 = e_2 = e_H$ by design, and the market is aware of this. Consequently, only the behavior of strategic workers is pertinent to the analysis.

We must ensure that the strategic types are better off when they choose not to work. Thus, the incentive compatibility condition for the high types is given by:

$$\mathbb{E}[w_2 \mid e_1 = e_H] - c \leq \mathbb{E}[w_2 \mid e_1 = e_L]$$

Since wages are a function of success or failure, we have the following:

$$\mathbb{P}\mathrm{r}(y_1 = y_H \mid e_1 = e_H) \cdot w_2(y_1 = y_H) + \mathbb{P}\mathrm{r}(y_1 = y_L \mid e_1 = e_H) \cdot w_2(y_1 = y_L) - c$$
$$\leq \mathbb{P}\mathrm{r}(y_1 = y_H \mid e_1 = e_L) \cdot w_2(y_1 = y_H) + \mathbb{P}\mathrm{r}(y_1 = y_L \mid e_1 = e_L) \cdot w_2(y_1 = y_L)$$

Substituting corresponding parameters yields:

$$\alpha \cdot [\hat{\theta}_{RS} \cdot \beta + (1 - \hat{\theta}_{RS}) \cdot \alpha] + (1 - \alpha) \cdot [\tilde{\theta}_{RS} \cdot \beta + (1 - \tilde{\theta}_{RS}) \cdot \alpha] - c$$
$$\leq \beta \cdot [\hat{\theta}_{RS} \cdot \beta + (1 - \hat{\theta}_{RS}) \cdot \alpha] + (1 - \beta) \cdot [\tilde{\theta}_{RS} \cdot \beta + (1 - \tilde{\theta}_{RS}) \cdot \alpha]$$

Further simplification results in:

$$(\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2 \leq c$$

This condition suggests that for strategic workers to find it optimal not to work (choosing $e_1 = e_L$), the cost c must be sufficiently high relative to the output's responsiveness to exerting high effort ($\alpha - \beta$) and the difference between posterior probabilities $\tilde{\theta}_{RS}$ and $\hat{\theta}_{RS}$. Essentially, it ensures that the disincentive to exert high effort (due to the cost c) outweighs any potential increase in expected wage from working.

An important final note is that the incentive compatibility condition provides a threshold level of c, denoted by c^R. In the rational market setting, if $c > c^R$, a separating equilibrium will occur where strategic types do not work and behavioral types do. Therefore, the threshold is given by $c^R = (\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2$. The following proposition elaborates on this threshold level of c.

Proposition 2 *In the rational setting, a separating equilibrium exists if and only if* $c \geq c^R = (\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2$.

Proof. Suppose that there exists a rational separating equilibrium. Then, for the strategic types, it is optimal not to work. Hence, $\mathbb{E}[w_2 \mid e_1 = e_H] - c \leq \mathbb{E}[w_2 \mid e_1 =$

$e_L]$. As shown above, it reduces to $(\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2 = c^R \leq c$. Conversely, suppose that we have $(\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2 = c^R \leq c$. Then,

$$
\begin{aligned}
\mathbb{E}[w_2 \mid e_1 = e_H] - c \leq \mathbb{E}[w_2 \mid e_1 = e_H] - c^R& \\
&= \alpha \cdot [\hat{\theta}_{RS} \cdot \beta + (1 - \hat{\theta}_{RS}) \cdot \alpha] + (1 - \alpha) \cdot [\tilde{\theta}_{RS} \cdot \beta + (1 - \tilde{\theta}_{RS}) \cdot \alpha] \\
&- (\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2 \\
&= \beta \cdot [\hat{\theta}_{RS} \cdot \beta + (1 - \hat{\theta}_{RS}) \cdot \alpha] + (1 - \beta) \cdot [\tilde{\theta}_{RS} \cdot \beta + (1 - \tilde{\theta}_{RS}) \cdot \alpha] \\
&= \mathbb{E}[w_2 \mid e_1 = e_L]
\end{aligned}
$$

Therefore, a rational separating equilibrium exists if and only if $c \geq c^R$. ■

4.2.2 Rational Pooling Equilibrium

In this section, the rational market conjectures an equilibrium where all strategic workers are successfully incentivized to work, in addition to the behavioral workers who always exert high effort. If all strategic types choose to work, it constitutes a pooling or pure strategy equilibrium. The following proposition addresses the impossibility of this scenario.

Proposition 3 *There does not exist a pooling or pure strategy equilibrium in which all strategic workers with $c > 0$ choose $e_1 = e_H$ when the market is rational.*

Proof. Assume, for the sake of contradiction, that a pooling equilibrium exists in which all strategic workers with $c > 0$ choose to exert high effort in the first period. The following are the three components of such an equilibrium:

- Since the market conjectures an equilibrium where everyone works, $\mathbb{P}\mathrm{r}(e_1 = e_H) = 1$ and $\mathbb{P}\mathrm{r}(e_1 = e_L) = 0$. Consequently, the posteriors for the rational pooling equilibrium are

$$\hat{\theta}_{RP} = \frac{\mathbb{P}\text{r(strategic type)} \cdot \mathbb{P}\text{r}(y_1 = y_H \mid \text{strategic type})}{\mathbb{P}\text{r}(y_1 = y_H)} = \frac{\theta\alpha}{\theta\alpha + (1-\theta)\alpha}$$
$$= \theta$$

$$\tilde{\theta}_{RP} = \frac{\mathbb{P}\text{r(strategic type)} \cdot \mathbb{P}\text{r}(y_1 = y_L \mid \text{strategic type})}{\mathbb{P}\text{r}(y_1 = y_L)} = \frac{\theta(1-\alpha)}{\theta(1-\alpha) + (1-\theta)(1-\alpha)}$$
$$= \theta$$

These posteriors indicate that regardless of the output observed, the market remains believing in the prior after observation. That is because all workers choose $e_1 = e_H$, and output does not provide additional information.

- The second-period wage w_2 at equilibrium is derived as:

$$w_2^{RP} = \begin{cases} \hat{\theta}_{RP} \cdot \beta + (1 - \hat{\theta}_{RP}) \cdot \alpha & \text{if } y_1 = y_H \\ \tilde{\theta}_{RP} \cdot \beta + (1 - \tilde{\theta}_{RP}) \cdot \alpha & \text{if } y_1 = y_L \end{cases}$$

Given the posteriors, this simplifies to $w_2^{RP} = \theta\beta + (1-\theta)\alpha$. Hence, the wage is fixed and not a function of the observed output or the unobservable effort.

- The incentive compatibility condition for strategic types is as follows:

$$\mathbb{E}[w_2 \mid e_1 = e_H] - c \geq \mathbb{E}[w_2 \mid e_1 = e_L]$$

Equivalently,

$$\mathbb{P}\text{r}(y_1 = y_H \mid e_1 = e_H) \cdot w_2^{RP} + \mathbb{P}\text{r}(y_1 = y_L \mid e_1 = e_H) \cdot w_2^{RP} - c \geq$$
$$\mathbb{P}\text{r}(y_1 = y_H \mid e_1 = e_L) \cdot w_2^{RP} + \mathbb{P}\text{r}(y_1 = y_L \mid e_1 = e_L) \cdot w_2^{RP}$$

After substituting, we get:

$$\alpha(\theta\beta + (1-\theta)\alpha) + (1-\alpha)(\theta\beta + (1-\theta)\alpha) - c \geq$$
$$\beta(\theta\beta + (1-\theta)\alpha) + (1-\beta)(\theta\beta + (1-\theta)\alpha)$$

$$\theta\beta + (1-\theta)\alpha - c \geq \theta\beta + (1-\theta)\alpha$$

$$c \leq 0$$

This is a contradiction because $c \in (0, \infty)$. Hence, a pooling equilibrium does not exist where all types choose $e_1 = e_H$. ■

The intuition behind Proposition 3 is that if all workers exert effort in the first period, the market gains no information about individual success related to cost and effort since everyone is already working. If the market anticipates that all individuals will work in the first period, there is no rationale for conditioning second-period wages on output. In equilibrium, all individuals work, with some achieving low output and others high output. Given that the market learns nothing in the first period, it sets a fixed wage for the second.

This fixed wage corresponds to $\theta\beta + (1-\theta)\alpha$, reflecting the market's expectation of average output without any new information from the first period. Since this wage is not influenced by individual performance, strategic workers have no incentive to exert high effort in the first period. Consequently, such an equilibrium, where all workers choose to work, does not exist.

4.2.3 Rational Semi-Separating Equilibrium

In this section, the rational market conjectures an equilibrium in which some strategic workers are effectively motivated to exert effort alongside behavioral workers. When the market incentivizes only a subset of strategic workers, it leads to a semi-separating or mixed strategy equilibrium.

Let the market conjecture an equilibrium where some strategic types work, i.e., the strategic workers mix strategies between e_H and e_L. With probability $\omega \in (0, 1)$, strategic types choose $e_1 = e_H$, and the $1 - \omega$ portion of the strategic types choose $e_1 = e_L$. In other words, the market conjectures an equilibrium where

$$\mathbb{P}\mathrm{r}(e_1 = e_H) = 1 - \theta + \theta \cdot \omega \text{ and } \mathbb{P}\mathrm{r}(e_1 = e_L) = \theta \cdot (1 - \omega)$$
$$\mathbb{P}\mathrm{r}(e_2 = e_H) = 1 - \theta \text{ and } \mathbb{P}\mathrm{r}(e_2 = e_L) = \theta$$

Initially, consider the first period. In equilibrium, the workers who exert effort at $t = 1$ comprise both types. Specifically, $1 - \theta$ of these workers are behavioral, while $\theta \cdot \omega$ are strategic types. The workers who abstain from exerting effort in the first period are the remaining strategic types, representing the $\theta \cdot (1 - \omega)$ fraction. In the second period, the situation remains consistent: the market cannot incentivize strategic workers, whereas behavioral workers persist in exerting effort. The following outlines the three components of the semi-separating equilibrium.

Posteriors

Recall that the posteriors are

$$\hat{\theta} = \mathbb{P}\mathrm{r}(\text{strategic type} \mid y_1 = y_H)$$

$$\tilde{\theta} = \mathbb{P}\mathrm{r}(\text{strategic type} \mid y_1 = y_L)$$

Notice that $\mathbb{P}\mathrm{r}(\text{strategic type} \mid y_1) \neq \mathbb{P}\mathrm{r}(e_1 = e_L \mid y_1)$ and $\mathbb{P}\mathrm{r}(\text{strategic type} \mid y_1) \neq \mathbb{P}\mathrm{r}(e_1 = e_H \mid y_1)$ because ω portion of the high types choose $e_1 = e_H$ in the semi-separating equilibrium. Therefore, the posteriors, denoted as $\hat{\theta}_{RSS}$ and $\tilde{\theta}_{RSS}$, in the rational semi-separating equilibrium are:

$$\begin{aligned} \hat{\theta}_{RSS} &= \mathbb{P}\mathrm{r}(\text{strategic type} \mid y_1 = y_H) \\ &= \frac{\mathbb{P}\mathrm{r}(\text{strategic type}) \cdot \mathbb{P}\mathrm{r}(y_1 = y_H \mid \text{strategic type})}{\mathbb{P}\mathrm{r}(y_1 = y_H)} \end{aligned}$$

$$\begin{aligned} \tilde{\theta}_{RSS} &= \mathbb{P}\mathrm{r}(\text{strategic type} \mid y_1 = y_L) \\ &= \frac{\mathbb{P}\mathrm{r}(\text{strategic type}) \cdot \mathbb{P}\mathrm{r}(y_1 = y_L \mid \text{strategic type})}{\mathbb{P}\mathrm{r}(y_1 = y_L)} \end{aligned}$$

Since success might result from high or low effort, we have

$$
\begin{aligned}
\hat{\theta}_{RSS} &= \mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_H, y_1 = y_H) \cdot \mathbb{P}\mathrm{r}(e_1 = e_H \mid y_1 = y_H) \\
&\quad + \mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_L, y_1 = y_H) \cdot \mathbb{P}\mathrm{r}(e_1 = e_L \mid y_1 = y_H) \\
\tilde{\theta}_{RSS} &= \mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_H, y_1 = y_L) \cdot \mathbb{P}\mathrm{r}(e_1 = e_H \mid y_1 = y_L) \\
&\quad + \mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_L, y_1 = y_L) \cdot \mathbb{P}\mathrm{r}(e_1 = e_L \mid y_1 = y_L)
\end{aligned}
$$

Proposition 4 *Effort (e) is a sufficient statistic for the worker's type.*

Proof. To demonstrate that e is a sufficient statistic for the worker's type, we need to show that $\mathbb{P}\mathrm{r}(\text{strategic type} \mid e, y) = \mathbb{P}\mathrm{r}(\text{strategic type} \mid e)$ for all $e \in \{e_H, e_L\}$ and $y \in \{y_H, y_L\}$. Since both effort and output are binary variables, we will consider all possible cases as follows:

$$
\begin{aligned}
\mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_H, y_1 = y_H) &= \frac{\mathbb{P}\mathrm{r}(\text{strategic type}, e_1 = e_H, y_1 = y_H)}{\mathbb{P}\mathrm{r}(e_1 = e_H) \cdot \mathbb{P}\mathrm{r}(y_1 = y_H \mid e_1 = e_H)} \\
&= \frac{\theta \cdot \omega \cdot \alpha}{(1 - \theta + \theta\omega) \cdot \alpha} \\
&= \frac{\theta \cdot \omega}{1 - \theta + \theta\omega} \\
&= \mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_H).
\end{aligned}
$$

The final equality holds because $\mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_H)$ represents the proportion of strategic types among the workers who choose to exert high effort in the first period.

$$
\begin{aligned}
\mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_L, y_1 = y_H) &= \frac{\mathbb{P}\mathrm{r}(\text{strategic type}, e_1 = e_L, y_1 = y_H)}{\mathbb{P}\mathrm{r}(e_1 = e_L) \cdot \mathbb{P}\mathrm{r}(y_1 = y_H \mid e_1 = e_L)} \\
&= \frac{\theta \cdot (1 - \omega) \cdot \beta}{\theta \cdot (1 - \omega) \cdot \beta} \\
&= 1 \\
&= \mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_L).
\end{aligned}
$$

This equality holds because $\mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_L)$ denotes the proportion of strategic types among the workers who opt not to exert high effort in the first period. By definition, behavioral workers never select $e_1 = e_L$.

$$\begin{aligned}
\mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_H, y_1 = y_L) &= \frac{\mathbb{P}\mathrm{r}(\text{strategic type}, e_1 = e_H, y_1 = y_L)}{\mathbb{P}\mathrm{r}(e_1 = e_H) \cdot \mathbb{P}\mathrm{r}(y_1 = y_L \mid e_1 = e_H)} \\
&= \frac{\theta \cdot \omega \cdot (1-\alpha)}{(1-\theta+\theta\omega) \cdot (1-\alpha)} \\
&= \frac{\theta \cdot \omega}{1-\theta+\theta\omega} \\
&= \mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_H).
\end{aligned}$$

The final equality holds because $\mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_H)$ represents the proportion of strategic types among the workers who choose to exert high effort in the first period. Conditioning on $e_1 = e_H$ inherently accounts for the strategic types, rendering the conditional probability $\mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_H, y_1 = y_L)$ equivalent to $\mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_H)$. The observed output $y_1 = y_L$ does not provide additional information about the type once conditioned on the effort level $e_1 = e_H$.

$$\begin{aligned}
\mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_L, y_1 = y_L) &= \frac{\mathbb{P}\mathrm{r}(\text{strategic type}, e_1 = e_L, y_1 = y_L)}{\mathbb{P}\mathrm{r}(e_1 = e_L) \cdot \mathbb{P}\mathrm{r}(y_1 = y_L \mid e_1 = e_L)} \\
&= \frac{\theta \cdot (1-\omega) \cdot (1-\beta)}{\theta \cdot (1-\omega) \cdot (1-\beta)} \\
&= 1 \\
&= \mathbb{P}\mathrm{r}(\text{strategic type} \mid e_1 = e_L).
\end{aligned}$$

The concluding equality is accurate because only the ω portion of the strategic types choose not to work. Therefore, we conclude that $\mathbb{P}\mathrm{r}(\text{strategic type} \mid e, y) = \mathbb{P}\mathrm{r}(\text{strategic type} \mid e)$ for all $e \in \{e_H, e_L\}$ and $y \in \{y_H, y_L\}$. ■

Consequently, the output becomes independent of being a strategic type or not when conditioned on effort, as the only linkage between that and output is mediated through effort. In statistical terms, effort (e) is a sufficient statistic for the worker's type; thus, knowing e renders y conditionally independent of the status of being a strategic type. In other words, once e is known, observing y provides no additional information about workers' designation as a strategic type, with y merely facilitating inference about the unobserved effort level. Following from Proposition 4, the resulting posterior probabilities are:

$$\begin{aligned}
\hat{\theta}_{RSS} &= \mathbb{P}\text{r}(\text{strategic type} \mid e_1 = e_H) \cdot \mathbb{P}\text{r}(e_1 = e_H \mid y_1 = y_H) \\
&\quad + \mathbb{P}\text{r}(\text{strategic type} \mid e_1 = e_L) \cdot \mathbb{P}\text{r}(e_1 = e_L \mid y_1 = y_H) \\
\tilde{\theta}_{RSS} &= \mathbb{P}\text{r}(\text{strategic type} \mid e_1 = e_H) \cdot \mathbb{P}\text{r}(e_1 = e_H \mid y_1 = y_L) \\
&\quad + \mathbb{P}\text{r}(\text{strategic type} \mid e_1 = e_L) \cdot \mathbb{P}\text{r}(e_1 = e_L \mid y_1 = y_L)
\end{aligned}$$

By the Bayes Rule, the market updates as follows

$$\begin{aligned}
\mathbb{P}\text{r}(e_1 = e_H \mid y_1 = y_H) &= \frac{\mathbb{P}\text{r}(e_1 = e_H) \cdot \mathbb{P}\text{r}(y_1 = y_H \mid e_1 = e_H)}{\mathbb{P}\text{r}(y_1 = y_H)} \\
\mathbb{P}\text{r}(e_1 = e_H \mid y_1 = y_L) &= \frac{\mathbb{P}\text{r}(e_1 = e_H) \cdot \mathbb{P}\text{r}(y_1 = y_L \mid e_1 = e_H)}{\mathbb{P}\text{r}(y_1 = y_L)}
\end{aligned}$$

Again, since high or low output can stem from both effort levels

$$\begin{aligned}
\mathbb{P}\text{r}(y_1 = y_H) &= \mathbb{P}\text{r}(e_1 = e_H) \cdot \mathbb{P}\text{r}(y_1 = y_H \mid e_1 = e_H) \\
&\quad + \mathbb{P}\text{r}(e_1 = e_L) \cdot \mathbb{P}\text{r}(y_1 = y_H \mid e_1 = e_L) \\
&= (1 - \theta + \theta \cdot \omega) \cdot \alpha + \theta \cdot (1 - \omega) \cdot \beta
\end{aligned}$$

$$\begin{aligned}
\mathbb{P}\text{r}(y_1 = y_L) &= \mathbb{P}\text{r}(e_1 = e_H) \cdot \mathbb{P}\text{r}(y_1 = y_L \mid e_1 = e_H) \\
&\quad + \mathbb{P}\text{r}(e_1 = e_L) \cdot \mathbb{P}\text{r}(y_1 = y_L \mid e_1 = e_L) \\
&= (1 - \theta + \theta \cdot \omega) \cdot (1 - \alpha) + \theta \cdot (1 - \omega) \cdot (1 - \beta)
\end{aligned}$$

Therefore,

$$\begin{aligned}
\mathbb{P}\text{r}(e_1 = e_H \mid y_1 = y_H) &= \frac{(1 - \theta + \theta \cdot \omega) \cdot \alpha}{(1 - \theta + \theta \cdot \omega) \cdot \alpha + \theta \cdot (1 - \omega) \cdot \beta} \\
\mathbb{P}\text{r}(e_1 = e_H \mid y_1 = y_L) &= \frac{(1 - \theta + \theta \cdot \omega) \cdot (1 - \alpha)}{(1 - \theta + \theta \cdot \omega) \cdot (1 - \alpha) + \theta \cdot (1 - \omega) \cdot (1 - \beta)}
\end{aligned}$$

Note that these are not the posteriors of primary interest; they are merely intermediate results. Additionally, observe that

$$\mathbb{P}\text{r}(e_1 = e_L \mid y_1 = y_H) = 1 - \mathbb{P}\text{r}(e_1 = e_H \mid y_1 = y_H)$$
$$\mathbb{P}\text{r}(e_1 = e_L \mid y_1 = y_L) = 1 - \mathbb{P}\text{r}(e_1 = e_H \mid y_1 = y_L)$$

Therefore, the posterior probabilities are given by:

$$\hat{\theta}_{RSS} = \frac{\theta \cdot (\omega \cdot \alpha + (1-\omega) \cdot \beta)}{\alpha \cdot (1 - \theta + \theta \cdot \omega) + \beta \cdot \theta \cdot (1-\omega)}$$
$$\tilde{\theta}_{RSS} = \frac{\theta \cdot (\omega \cdot (1-\alpha) + (1-\omega) \cdot (1-\beta))}{(1-\alpha) \cdot (1 - \theta + \theta \cdot \omega) + (1-\beta) \cdot \theta \cdot (1-\omega)}$$

Wages

To understand the wage-setting mechanism in this framework, recall that the rational market sets the second-period wage, w_2, based on the probability of achieving high output given the first-period output, i.e., $w_2 = \mathbb{P}\text{r}(y_2 = y_H \mid y_1)$. In equilibrium, this translates to:

$$w_2^{RSS} = \begin{cases} \hat{\theta}_{RSS} \cdot \beta + (1 - \hat{\theta}_{RSS}) \cdot \alpha & \text{if } y_1 = y_H \\ \tilde{\theta}_{RSS} \cdot \beta + (1 - \tilde{\theta}_{RSS}) \cdot \alpha & \text{if } y_1 = y_L \end{cases}$$

Incentive Compatibility

Finally, verifying the behavior in the first period under the proposed equilibrium is essential. Specifically, workers must select effort levels that align with the established beliefs and wages. Behavioral workers will inherently choose $e_1 = e_2 = e_H$, which the market fully anticipates. Therefore, the analysis focuses primarily on the behavior of strategic workers.

We must ensure that the strategic types are indifferent between $e_2 = e_L$ and $e_1 = e_H$. Thus, the incentive compatibility condition for the strategic types is given by:

$$\mathbb{E}[w_2 \mid e_1 = e_H] - c = \mathbb{E}[w_2 \mid e_1 = e_L]$$

$$\mathbb{P}\mathrm{r}(y_1 = y_H \mid e_1 = e_H) \cdot w_2(y_1 = y_H) + \mathbb{P}\mathrm{r}(y_1 = y_L \mid e_1 = e_H) \cdot w_2(y_1 = y_L) - c$$
$$= \mathbb{P}\mathrm{r}(y_1 = y_H \mid e_1 = e_L) \cdot w_2(y_1 = y_H) + \mathbb{P}\mathrm{r}(y_1 = y_L \mid e_1 = e_L) \cdot w_2(y_1 = y_L)$$

Substituting corresponding parameters yields:

$$\alpha \cdot [\hat{\theta}_{RSS} \cdot \beta + (1 - \hat{\theta}_{RSS}) \cdot \alpha] + (1 - \alpha) \cdot [\tilde{\theta}_{RSS} \cdot \beta + (1 - \tilde{\theta}_{RSS}) \cdot \alpha] - c$$
$$= \beta \cdot [\hat{\theta}_{RSS} \cdot \beta + (1 - \hat{\theta}_{RSS}) \cdot \alpha] + (1 - \beta) \cdot [\tilde{\theta}_{RSS} \cdot \beta + (1 - \tilde{\theta}_{RSS}) \cdot \alpha]$$

Further simplification results in:

$$(\tilde{\theta}_{RSS} - \hat{\theta}_{RSS})(\alpha - \beta)^2 = c$$

Given the other parameters, this condition determines a unique level of ω. This value of ω dictates the specific mixing behavior of strategic workers. When $\omega \in (0, 1)$, the equilibrium is characterized as semi-separating, indicating that strategic workers mix between high and low effort. If $\omega = 0$, the equilibrium transitions to a fully separating equilibrium where no strategic workers exert high effort. In this case, the posteriors $\hat{\theta}_{RSS}$ and $\tilde{\theta}_{RSS}$ reduce to $\hat{\theta}_{RS}$ and $\tilde{\theta}_{RS}$, respectively. This condition ensures that the expected benefit of exerting high effort, adjusted for the cost c, does not exceed the expected benefit of exerting low effort, thereby maintaining the equilibrium. Below is a concluding proposition that summarizes how equilibria are determined in a rational market setting, contingent on the cost of effort.

Proposition 5 *Let $c^R = (\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2$ be the critical threshold that delineates different equilibrium regimes based on the cost level c. The equilibrium outcomes are as follows:*

- *If $c \geq c^R$, the resulting equilibrium is a separating equilibrium, wherein* *strategic types do not exert high effort.*

- *If $c < c^R$, the equilibrium that emerges is semi-separating and characterized by a mix of high and low effort among strategic types.*

Proof. Let us proceed step by step:

- If $c > c^R$, by the incentive compatibility condition of the rational separating equilibrium, we know that the derived equilibrium is a separating equilibrium.

- Suppose for a contradiction that if $c = c^R$, the equilibrium emerges as semi-separating with a valid $\omega \in (0, 1)$ according to the definition of the semi-separating equilibrium. Then, using the incentive compatibility condition of the rational semi-separating equilibrium, we have the following:

$$c = (\tilde{\theta}_{RSS} - \hat{\theta}_{RSS})(\alpha - \beta)^2 = (\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2 = c^R$$

  This implies:

$$\begin{aligned}
&\frac{\theta \cdot (\omega \cdot (1-\alpha) + (1-\omega) \cdot (1-\beta))}{(1-\alpha) \cdot (1-\theta+\theta \cdot \omega) + (1-\beta) \cdot \theta \cdot (1-\omega)} \\
&- \frac{\theta \cdot (\omega \cdot \alpha + (1-\omega) \cdot \beta)}{\alpha \cdot (1-\theta+\theta \cdot \omega) + \beta \cdot \theta \cdot (1-\omega)} \\
&= \frac{\theta \cdot (1-\beta)}{\theta \cdot (1-\beta) + (1-\theta) \cdot (1-\alpha)} - \frac{\theta \cdot \beta}{\theta \cdot \beta + (1-\theta) \cdot \alpha}
\end{aligned}$$

  This equality holds only when $\omega = 0$. As $\omega \notin (0, 1)$, we reached a contradiction; the resulting equilibrium is not semi-separating. Put differently, since ω is the proportion of strategic types who choose $e_1 = e_H$, none of the strategic types work in the first period. Hence, the equilibrium appears to be separating.

- Lastly, we employ proof by contrapositive: we need to demonstrate that if the equilibrium is not semi-separating, then $c > c^R$. By Proposition 2, the absence of a semi-separating equilibrium implies a separating equilibrium, as a pooling equilibrium does not exist. Therefore, we aim to show that a separating equilibrium implies $c > c^R$. Suppose we are in a separating equilibrium. Thus, $(\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2 < c$. Recall that $c^R = (\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2$. Therefore, we have $c^R < c$. ■

This concludes our analysis of the model under the assumption of a rational market, where beliefs are updated according to Bayes' Rule. Consider the following graphical illustration of Proposition 5. Since it is impossible to visualize a four-dimensional plot, I fix $\beta = 0.2$ and $\theta = 0.5$, which will be maintained henceforth.

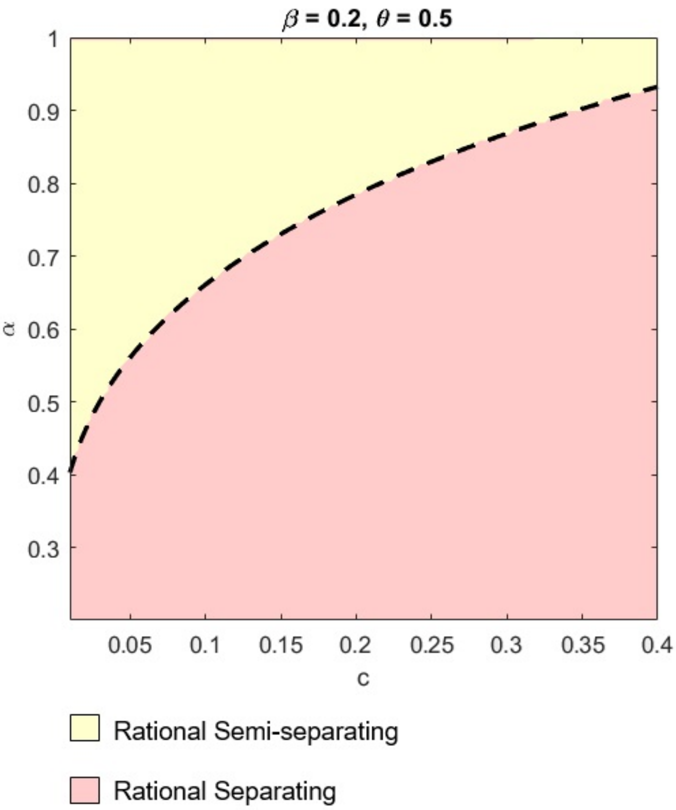

Figure 4.1: Rational Market Wage Determination: Semi-Separating and Separating Equilibria for $\beta = 0.2$ and $\theta = 0.5$

The incentive compatibility condition of the rational separating equilibrium provides c^R, which delineates the c-α plane into two regions. Thus, $c^R =$

$(\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2$ corresponds to the dashed concave curve. In other words, as the incentive compatibility condition of the rational semi-separating equilibrium creates different curves reflected by unique ω given each (c, α), ω corresponding to this curve is 0. To see this pluging $\omega = 0$ on $\tilde{\theta}_{RSS}$ and $\hat{\theta}_{RSS}$ gives $\tilde{\theta}_{RSS} = \tilde{\theta}_{RS}$ and $\hat{\theta}_{RSS} = \hat{\theta}_{RS}$.

Above this cutoff, the equilibrium that emerges is semi-separating, reflecting that $\alpha - \beta = \mathbb{P}\mathrm{r}(y_H \mid e_H) - \mathbb{P}\mathrm{r}(y_H \mid e_L)$, i.e., the responsiveness of output to the effort must be large enough compared to the cost of effort to incentivize strategic workers to mix between e_H and e_L. Therefore, in this region, the incentive compatibility condition of the semi-separating equilibrium is satisfied with an endogenous $\omega \in (0, 1)$.

Under the cutoff, the resulting equilibrium is separating as $\alpha - \beta$ is not large enough compared to c; subsequently, strategic types do not find the optimal to work. Put differently, considering the curves implied by the incentive compatibility condition of the rational semi-separating equilibrium are merely satisfied with $\omega \notin (0, 1)$. However, the ones corresponding to separating equilibrium fall into this region.

4.3 Asymmetric Information with Narratives Market

In this section, I will shift the focus from a rational market to examine the implications of strategic workers' behavior and second-period wages. The primary distinction between a rational market and a market that subscribes to 'narratives,' henceforth referred to as a 'narratives market,' lies in their method of updating prior beliefs. Unlike the rational market, which employs Bayesian inference, the narratives market adopts an extremist approach. In this framework, the narratives market interprets failure as exclusively resulting from not working and success as entirely attributable to working.

Consequently, while the posteriors are fundamentally defined in the same manner as before, with $\hat{\theta} = \mathbb{P}(\text{strategic type} \mid y_1 = y_H)$ and $\tilde{\theta} = \mathbb{P}(\text{strategic type} \mid y_1 = y_L)$, their values will differ under the narratives market approach. Specifically, due to this extremist perspective, the posteriors will be either 1 or 0, reflecting a binary interpretation of success and failure.

4.3.1 Narratives Separating Equilibrium

In this section, I consider the scenario where the narratives market conjectures an equilibrium in which none of the strategic workers choose $e_1 = e_H$. Consequently, we have $\mathbb{P}\text{r}(e_1 = e_L) = \mathbb{P}\text{r}(\text{strategic type}) = \theta$ and $\mathbb{P}\text{r}(e_1 = e_H) = \mathbb{P}\text{r}(\text{behavioral type}) = 1 - \theta$. The following are the three components of the separating equilibrium.

Posteriors

The market adopts the narrative based on the highest fit, as in Schwartzstein and Sunderam (2021) and Aina (2023). The fit of a narrative is defined as the likelihood of generating the observed outcome. That is, if the observed outcome is $y_1 = y_H$,

$$\text{Fit}(e_1 = e_H) = \mathbb{P}\text{r}(y_H \mid e_H) = \alpha$$
$$\text{Fit}(e_1 = e_L) = \mathbb{P}\text{r}(y_H \mid e_L) = \beta$$

Since $\alpha > \beta$ by Assumption 1, we conclude that $\text{Fit}(e_1 = e_H) > \text{Fit}(e_1 = e_L)$ upon observing success in the first period. Consequently, the market adopts the narrative that the worker is a behavioral type because they always work. Similarly, if the observed outcome is $y_1 = y_L$,

$$\text{Fit}(e_1 = e_H) = \mathbb{P}\text{r}(y_L \mid e_H) = 1 - \alpha$$
$$\text{Fit}(e_1 = e_L) = \mathbb{P}\text{r}(y_L \mid e_L) = 1 - \beta$$

Again, by Assumption 1, it follows that Fit$(e_1 = e_L) >$ Fit$(e_1 = e_H)$ upon observing failure. As a result, the market interprets the worker as a strategic type who does not always work.

Since the narratives market does not employ Bayes' Rule, it infers that observing $y_1 = y_H$ definitively indicates $e_1 = e_H$, and observing $y_1 = y_L$ definitively indicates $e_1 = e_L$. Given this extremist approach and the conjectured equilibrium where all strategic types choose $e_1 = e_L$, the posteriors are

$$\hat{\theta}_{NS} = \mathbb{P}\text{r}(\text{strategic type} \mid y_1 = y_H) = 0$$
$$\tilde{\theta}_{NS} = \mathbb{P}\text{r}(\text{strategic type} \mid y_1 = y_L) = 1$$

Wages

To comprehend the wage-setting mechanism under the narratives market, remember that the market, regardless of rationality, determines the second-period wage, w_2, based on the likelihood of attaining high output given the first-period output. Thus, $w_2 = \mathbb{P}\text{r}(y_2 = y_H \mid y_1)$. In equilibrium, this can be expressed as

$$\begin{aligned} w_2^N &= \begin{cases} \hat{\theta}_{NS} \cdot \beta + (1 - \hat{\theta}_{NS}) \cdot \alpha & \text{if } y_1 = y_H \\ \tilde{\theta}_{NS} \cdot \beta + (1 - \tilde{\theta}_{NS}) \cdot \alpha & \text{if } y_1 = y_L \end{cases} \\ &= \begin{cases} \alpha & \text{if } y_1 = y_H \\ \beta & \text{if } y_1 = y_L \end{cases} \end{aligned}$$

Incentive Compatibility

Lastly, it is crucial to validate the actions of strategic workers in the first period within the suggested equilibrium framework. We need to confirm that strategic types gain more by opting for $e_1 = e_L$ rather than $e_1 = e_H$. Hence, the incentive compatibility condition for strategic types is formulated as

$$\mathbb{E}[w_2 \mid e_1 = e_H] - c \leq \mathbb{E}[w_2 \mid e_1 = e_L]$$

$$\mathbb{P}\mathrm{r}(y_1 = y_H \mid e_1 = e_H) \cdot w_2(y_1 = y_H) + \mathbb{P}\mathrm{r}(y_1 = y_L \mid e_1 = e_H) \cdot w_2(y_1 = y_L) - c$$
$$\leq \mathbb{P}\mathrm{r}(y_1 = y_H \mid e_1 = e_L) \cdot w_2(y_1 = y_H) + \mathbb{P}\mathrm{r}(y_1 = y_L \mid e_1 = e_L) \cdot w_2(y_1 = y_L)$$

Replacing the relevant parameters results in:

$$\alpha^2 + (1 - \alpha) \cdot \beta - c \leq \beta \cdot \alpha + (1 - \beta) \cdot \beta$$

Further reduction leads to:

$$(\alpha - \beta)^2 \leq c$$

This condition ensures that when the market believes in narratives, for strategic workers find it optimal not to work, their cost of effort must be higher than the responsiveness of output to the effort, i.e., $(\alpha - \beta)^2$. It represents the squared difference between the probabilities of achieving high output under high and low effort, respectively, effectively measuring the impact of exerting high versus low effort. A larger value of $(\alpha - \beta)^2$ indicates a greater disparity in output due to effort differences.

The cost c, representing the burden of high effort, must be sufficiently large to dissuade strategic workers from choosing high effort. When this condition holds, it implies that the cost of exerting high effort is too high to be justified by the potential benefits in terms of expected wages, leading strategic workers to opt for low effort. Therefore, this condition ensures a separating equilibrium.

More importantly, similar to the rational market scenario, this condition provides a threshold level denoted by c^N such that for all $c > c^N$, the resulting equilibrium is separating and $c^N = (\alpha - \beta)^2$.

Proposition 6 *In the narratives setting, a separating equilibrium exists if and only if* $c \geq c^N = (\alpha - \beta)^2$.

Proof. Suppose that there exists a narratives separating equilibrium. Then, for the strategic types, it is optimal not to work. Hence, $\mathbb{E}[w_2 \mid e_1 = e_H] - c \leq \mathbb{E}[w_2 \mid e_1 = e_L]$. As shown above, it reduces to $(\alpha - \beta)^2 = c^N \leq c$. Conversely, suppose that we have $(\alpha - \beta)^2 = c^N \leq c$. Then,

$$\begin{aligned}
\mathbb{E}[w_2 \mid e_1 = e_H] - c \leq \mathbb{E}[w_2 \mid e_1 = e_H] - c^N \\
&= \alpha^2 + (1 - \alpha) \cdot \beta - (\alpha - \beta)^2 \\
&= \beta \cdot \alpha + (1 - \beta) \cdot \beta \\
&= \mathbb{E}[w_2 \mid e_1 = e_L]
\end{aligned}$$

Hence, a narratives separating equilibrium exists if and only if $c \geq c^N$ ∎

4.3.2 Narratives Pooling Equilibrium

Lastly, let us examine the implications of a narratives market conjecturing a pooling equilibrium where strategic types optimally choose to work. Consequently, we have $\mathbb{P}\mathrm{r}(e_1 = e_L) = 0$ and $\mathbb{P}\mathrm{r}(e_1 = e_H) = 1$. The prior probabilities remain unchanged, with $\mathbb{P}\mathrm{r}(\text{strategic type}) = \theta$ and $\mathbb{P}\mathrm{r}(\text{behavioral type}) = 1 - \theta$. The subsequent analysis will outline the three key components of the pooling equilibrium.

Posteriors

The narratives market disregards Bayes' Rule, interpreting $y_1 = y_H$ as a definitive indicator of $e_1 = e_H$ and $y_1 = y_L$ as or of $e_1 = e_L$. Therefore, the posteriors are the same as before:

$$\hat{\theta}_{NP} = \mathbb{P}\text{r}(\text{strategic type} \mid y_1 = y_H) = 0$$
$$\tilde{\theta}_{NP} = \mathbb{P}\text{r}(\text{strategic type} \mid y_1 = y_L) = 1$$

Wages

The same posteriors will lead to the same wages as before. In equilibrium, we have:

$$\begin{aligned} w_2^N &= \begin{cases} \hat{\theta}_{NP} \cdot \beta + (1 - \hat{\theta}_{NP}) \cdot \alpha & \text{if } y_1 = y_H \\ \tilde{\theta}_{NP} \cdot \beta + (1 - \tilde{\theta}_{NP}) \cdot \alpha & \text{if } y_1 = y_L \end{cases} \\ &= \begin{cases} \alpha & \text{if } y_1 = y_H \\ \beta & \text{if } y_1 = y_L \end{cases} \end{aligned}$$

Incentive Compatibility

Lastly, the incentive compatibility condition for strategic types is formulated as before since the posteriors and wages remain the same, but the direction of the inequality sign shifts this time:

$$\mathbb{E}[w_2 \mid e_1 = e_H] - c \geq \mathbb{E}[w_2 \mid e_1 = e_L]$$

Substituting the pertinent parameters and simplifying leads to the following:

$$(\alpha - \beta)^2 \geq c$$

This condition dictates that in a market governed by narratives, strategic workers will only find it optimal to exert high effort if their effort cost is less than the output's responsiveness to effort. Specifically, the effort cost c must be sufficiently low to make high effort worthwhile for strategic workers. When this condition is satisfied, it indicates that the cost of high effort is outweighed by the expected wage benefits, prompting strategic workers to choose high effort. Consequently, this condition supports a pooling equilibrium. Furthermore, remember the threshold level c^N. For $c < c^N$, the equilibrium remains pooling.

Finally, the following concluding proposition dictates the existence of a semi-separating equilibrium when the market believes in narratives instead of applying Bayes' Rule.

Proposition 7 *When the market believes narratives, a semi-separating or mixed strategy equilibrium exists if and only if $(\alpha - \beta)^2 = c$.*

Proof. Suppose a semi-separating equilibrium exists for which the strategic types mix between $e_1 = e_H$ and $e_1 = e_L$. The posteriors and second period's wages remain the same as the market believes in narratives. As before, the incentive compatibility condition that ensures the strategic types are indifferent between working and not working is as follows:

$$\mathbb{E}[w_2 \mid e_1 = e_H] - c = \mathbb{E}[w_2 \mid e_1 = e_L]$$
$$(\alpha - \beta)^2 = c$$

Conversely, suppose that $(\alpha - \beta)^2 = c$. Then,

$$\begin{aligned}
\mathbb{E}[w_2 \mid e_1 = e_H] - c &= \mathbb{E}[w_2 \mid e_1 = e_H] - (\alpha - \beta)^2 \\
&= \alpha^2 + (1 - \alpha) \cdot \beta - \alpha^2 - \beta^2 + 2 \cdot \alpha \cdot \beta \\
&= \alpha \cdot \beta + \beta + \beta^2 \\
&= \mathbb{E}[w_2 \mid e_1 = e_L]
\end{aligned}$$

Therefore, the strategic types are indifferent between $e_1 = e_H$ and $e_1 = e_L$, implying the resulting equilibrium is semi-separating. ■

This concludes our analysis of the model under the assumption of a market that believes in narratives, where beliefs are not updated according to Bayes' Rule. Examine the following graphical representation of the narratives framework alongside the threshold curve of the rational market. It is clear that the boundary of the rational market is above the narratives market's because $c^R = (\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2 < (\alpha - \beta)^2 = c^N$.

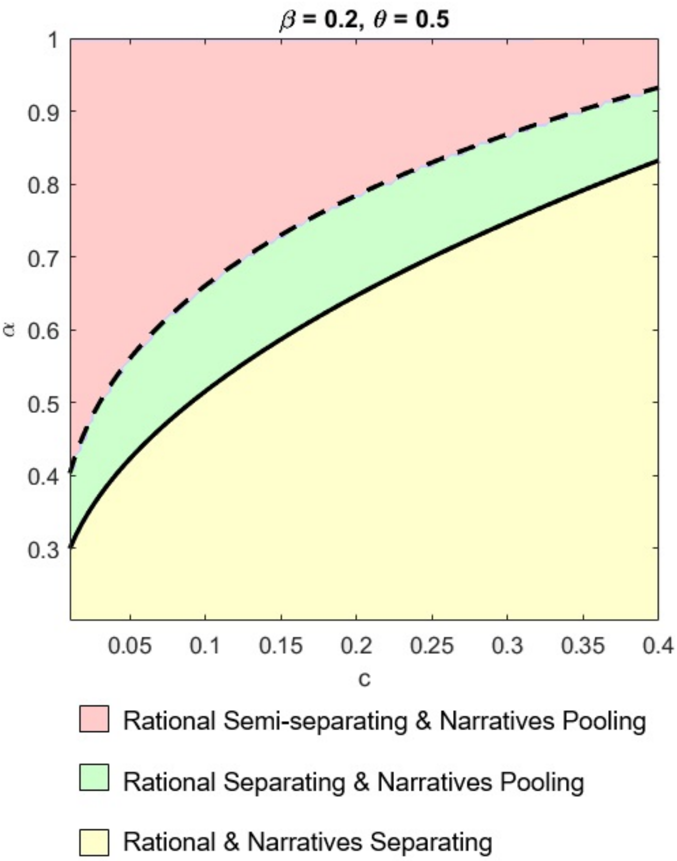

Figure 4.2: Comparative Analysis of Rational and Narratives Markets: Semi-Separating, Separating, and Pooling Equilibria for $\beta = 0.2$ and $\theta = 0.5$

The dashed curve corresponds to $c^R = (\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2$ and the black curve is $(\alpha - \beta)^2 = c^N$. Above the black curve, we observe pooling equilibrium in the narratives setting, and below, separating equilibrium occurs. On the curve, the resulting equilibrium is semi-separating.

CHAPTER 5

RESULTS

This section examines the implications of the model. First, I will illustrate the relationship between the variances of the second-period wage under rational and narrative frameworks as a function of α. Second, I will examine how the model's predicted gender wage gap varies with α in order to explore its implications across different sectors. Lastly, I will analyze the welfare implications of the model by determining whether exposure to the narratives market improves or diminishes workers' well-being, contingent on the parameter region.

5.1 Variance of Late Career Wage

It is feasible to compute the variance of second-period wages within the equilibrium framework for both rational and narrative markets. This calculation is possible because the second-period wage is a nonbinary binomial variable that takes on one value with a certain probability and another with a complementary probability. To illustrate this, recall that w_2 is a piecewise function of cost, which means that the equilibrium regime shifts depending on the cost. For the Bayesian market, as the cost of effort increases, we transition from a semi-separating equilibrium to a separating equilibrium. Consequently, the second-period wages for the rational market are given by:

$$w_2^R = \begin{cases} \hat{\theta}_{RSS} \cdot \beta + (1 - \hat{\theta}_{RSS}) \cdot \alpha & \text{if } y_1 = y_H \text{ and } c < c^R \\ \tilde{\theta}_{RSS} \cdot \beta + (1 - \tilde{\theta}_{RSS}) \cdot \alpha & \text{if } y_1 = y_L \text{ and } c < c^R \\ \hat{\theta}_{RS} \cdot \beta + (1 - \hat{\theta}_{RS}) \cdot \alpha & \text{if } y_1 = y_H \text{ and } c \geq c^R \\ \tilde{\theta}_{RS} \cdot \beta + (1 - \tilde{\theta}_{RS}) \cdot \alpha & \text{if } y_1 = y_L \text{ and } c \geq c^R \end{cases}$$

Instead of defining the rational market's second-period wages as a piecewise function of cost, we can also characterize in the following fashion, separately for rational semi-separating and separating equilibria:

$$\begin{aligned} w_2^{RSS} &= \begin{cases} w_2^{RSS}(y_1 = y_H) & \text{wp } \mathbb{P}\text{r}^{RSS}(y_1 = y_H) \\ w_2^{RSS}(y_1 = y_L) & \text{wp } \mathbb{P}\text{r}^{RSS}(y_1 = y_L) \end{cases} \\ &= \begin{cases} \hat{\theta}_{RSS} \cdot \beta + (1 - \hat{\theta}_{RSS}) \cdot \alpha & \text{wp } \mathbb{P}\text{r}^{RSS}(y_1 = y_H) \\ \tilde{\theta}_{RSS} \cdot \beta + (1 - \tilde{\theta}_{RSS}) \cdot \alpha & \text{wp } \mathbb{P}\text{r}^{RSS}(y_1 = y_L) \end{cases} \end{aligned}$$

$$\begin{aligned} w_2^{RS} &= \begin{cases} w_2^{RS}(y_1 = y_H) & \text{wp } \mathbb{P}\text{r}^{RS}(y_1 = y_H) \\ w_2^{RS}(y_1 = y_L) & \text{wp } \mathbb{P}\text{r}^{RS}(y_1 = y_L) \end{cases} \\ &= \begin{cases} \hat{\theta}_{RS} \cdot \beta + (1 - \hat{\theta}_{RS}) \cdot \alpha & \text{wp } \mathbb{P}\text{r}^{RS}(y_1 = y_H) \\ \tilde{\theta}_{RS} \cdot \beta + (1 - \tilde{\theta}_{RS}) \cdot \alpha & \text{wp } \mathbb{P}\text{r}^{RS}(y_1 = y_L) \end{cases} \end{aligned}$$

Next, we calculate the probabilities by considering that success can arise from either high or low effort, along with the corresponding equilibrium behavior as conjectured by the market. To start with a simpler case, we have $\mathbb{P}\text{r}^{RS}(y_1 = y_H) = \theta \cdot \beta + (1 - \theta) \cdot \alpha$. Here, θ represents the proportion of strategic workers who do not exert effort in the separating equilibrium, with the probability of success given low effort being β. The remaining $1 - \theta$ represents the behavioral workers who consistently exert high effort, with the probability of success given high effort being α. By employing a similar rationale, we obtain:

$$\mathbb{P}\mathrm{r}^{RS}(y_1 = y_L) = \theta \cdot (1-\beta) + (1-\theta) \cdot (1-\alpha)$$

$$\mathbb{P}\mathrm{r}^{RSS}(y_1 = y_H) = \theta \cdot (1-\omega) \cdot \beta + (1-\theta+\theta \cdot \omega) \cdot \alpha$$

$$\mathbb{P}\mathrm{r}^{RSS}(y_1 = y_L) = \theta \cdot (1-\omega) \cdot (1-\beta) + (1-\theta+\theta \cdot \omega) \cdot (1-\alpha)$$

Let us further consider the final probability, $\mathbb{P}\mathrm{r}^{RSS}(y_1 = y_L)$, to enhance our understanding. In this context, we are dealing with a semi-separating equilibrium. A proportion $\theta \cdot (1-\omega)$ of the workers are strategic and choose low effort, with the probability of failure given low effort being $1-\beta$. The remaining component accounts for the other workers, who are either behavioral types that consistently exert high effort or strategic types that decide to work. For these groups, the probability of failure given high effort is $1-\alpha$.

Thus, the expected wages for the rational market in the second period are:

$$\begin{aligned}
\mathbb{E}[w_2^{RSS}] &= [\hat{\theta}_{RSS} \cdot \beta + (1-\hat{\theta}_{RSS}) \cdot \alpha] \cdot \mathbb{P}\mathrm{r}^{RSS}(y_1 = y_H) \\
&\quad + [\tilde{\theta}_{RSS} \cdot \beta + (1-\tilde{\theta}_{RSS}) \cdot \alpha] \cdot \mathbb{P}\mathrm{r}^{RSS}(y_1 = y_L) \\
\mathbb{E}[w_2^{RS}] &= [\hat{\theta}_{RS} \cdot \beta + (1-\hat{\theta}_{RS}) \cdot \alpha] \cdot \mathbb{P}\mathrm{r}^{RS}(y_1 = y_H) \\
&\quad + [\tilde{\theta}_{RS} \cdot \beta + (1-\tilde{\theta}_{RS}) \cdot \alpha] \cdot \mathbb{P}\mathrm{r}^{RS}(y_1 = y_L)
\end{aligned}$$

Which is equivalent to

$$\mathbb{E}[w_2^{R}] = \begin{cases}
[\hat{\theta}_{RSS} \cdot \beta + (1-\hat{\theta}_{RSS}) \cdot \alpha] \cdot \mathbb{P}\mathrm{r}^{RSS}(y_1 = y_H) + & \\
[\tilde{\theta}_{RSS} \cdot \beta + (1-\tilde{\theta}_{RSS}) \cdot \alpha] \cdot \mathbb{P}\mathrm{r}^{RSS}(y_1 = y_L) & \text{if } c < c^R \\
[\hat{\theta}_{RS} \cdot \beta + (1-\hat{\theta}_{RS}) \cdot \alpha] \cdot \mathbb{P}\mathrm{r}^{RS}(y_1 = y_H) + & \\
[\tilde{\theta}_{RS} \cdot \beta + (1-\tilde{\theta}_{RS}) \cdot \alpha] \cdot \mathbb{P}\mathrm{r}^{RS}(y_1 = y_L) & \text{if } c > c^R
\end{cases}$$

Therefore, the variance of second-period wages for the rational market is:

$$\text{Var}(w_2^R) = \begin{cases} \mathbb{E}[{w_2}^{RSS}] \cdot (1 - \mathbb{E}[{w_2}^{RSS}]) \cdot (\tilde{\theta}_{RSS} - \hat{\theta}_{RSS})^2 \cdot (\alpha - \beta)^2 & \text{if } c < c^R \\ \mathbb{E}[{w_2}^{RS}] \cdot (1 - \mathbb{E}[{w_2}^{RS}]) \cdot (\tilde{\theta}_{RS} - \hat{\theta}_{RS})^2 \cdot (\alpha - \beta)^2 & \text{if } c > c^R \end{cases}$$

In the narratives market, although the equilibrium transitions from pooling to separating as the cost increases, the posteriors remain constant across these regimes. Consequently, the second-period wages are independent of the cost. Therefore, the second-period wages for the narratives market are given by:

$${w_2}^N = \begin{cases} \alpha & \text{if } y_1 = y_H \\ \beta & \text{if } y_1 = y_L \end{cases}$$

The second-period wages in the narratives market can be described as follows:

$${w_2}^{NP} = \begin{cases} \alpha & \text{wp } \mathbb{P}\text{r}^{NP}(y_1 = y_H) \\ \beta & \text{wp } \mathbb{P}\text{r}^{NP}(y_1 = y_L) \end{cases}$$

$${w_2}^{NS} = \begin{cases} \alpha & \text{wp } \mathbb{P}\text{r}^{NS}(y_1 = y_H) \\ \beta & \text{wp } \mathbb{P}\text{r}^{NS}(y_1 = y_L) \end{cases}$$

Specific probabilities for different output scenarios arise for each equilibrium setting in both the narratives market as well. In the pooling equilibrium, where all workers exert effort, and the probability of high output given high effort is α, the probabilities are given by:

$$\mathbb{P}\text{r}^{NP}(y_1 = y_H) = \alpha$$
$$\mathbb{P}\text{r}^{NP}(y_1 = y_L) = 1 - \alpha$$

Similarly, in the separating equilibrium, where strategic workers either exert low effort or do not work at all, the probabilities of achieving high and low output are derived as follows:

$$\mathbb{P}\text{r}^{NS}(y_1 = y_H) = \theta \cdot \beta + (1-\theta) \cdot \alpha$$

$$\mathbb{P}\text{r}^{NS}(y_1 = y_L) = \theta \cdot (1-\beta) + (1-\theta) \cdot (1-\alpha)$$

The second-period expected wages in the narratives market can be determined as follows:

$$\mathbb{E}[{w_2}^{NP}] = \alpha \cdot \mathbb{P}\text{r}^{NP}(y_1 = y_H) + \beta \cdot \mathbb{P}\text{r}^{NP}(y_1 = y_L)$$

$$\mathbb{E}[{w_2}^{NS}] = \alpha \cdot \mathbb{P}\text{r}^{NS}(y_1 = y_H) + \beta \cdot \mathbb{P}\text{r}^{NS}(y_1 = y_L)$$

Equivalently,

$$\mathbb{E}[{w_2}^{N}] = \begin{cases} \alpha \cdot \mathbb{P}\text{r}^{NP}(y_1 = y_H) + \beta \cdot \mathbb{P}\text{r}^{NP}(y_1 = y_L) & \text{if } c < c^N \\ \alpha \cdot \mathbb{P}\text{r}^{NS}(y_1 = y_H) + \beta \cdot \mathbb{P}\text{r}^{NS}(y_1 = y_L) & \text{if } c > c^N \end{cases}$$

Therefore, the wage spread for the narratives market is:

$$\text{Var}(w_2^N) = \begin{cases} \mathbb{E}[{w_2}^{NP}] \cdot (1 - \mathbb{E}[{w_2}^{NP}]) \cdot (\alpha-\beta)^2 & \text{if } c < c^N \\ \mathbb{E}[{w_2}^{NS}] \cdot (1 - \mathbb{E}[{w_2}^{NS}]) \cdot (\alpha-\beta)^2 & \text{if } c > c^N \end{cases}$$

Theorem 1 *Given fixed values of β and θ,*

- *There exists some $\alpha_1 \in (\beta, 1)$ such that, if $\alpha < \alpha_1$, then $Var(w^R) < Var(w^N)$ for all $c > 0$.*
- *There exists some $\alpha_2 \in (\alpha_1, 1)$ such that, if $\alpha > \alpha_2$, then $Var(w^R) \geq Var(w^N)$ for all $c > 0$.*

Proof. We will show the existence of such cutoffs step by step as follows.

Step 1. We note that $\text{Var}(w_2^R)$ is convex in and increases with c for all $c < c^R$

as demonstrated using Mathematica. However, we have omitted the first and second derivatives because of their complexity. Furthermore, for all $c \geq c^R$, $\text{Var}(w_2^R)$ is constant in c since the resulting equilibrium is separating, strategic types do not work regardless of the cost of effort.

Step 2. As $c \to 0$, the resulting equilibrium in the rational setting is semi-separating since $c^R > 0$. By the incentive compatibility condition, we have $c = (\tilde{\theta}^{RSS} - \hat{\theta}^{RSS})(\alpha - \beta)^2$. Thus, as $c \to 0$, we have $\tilde{\theta}^{RSS} \to \hat{\theta}^{RSS}$. Both posteriors are equal to each other when $\omega \to 1$. Therefore, as $c \to 0$, we have $\omega \to 1$ and $w_2^{RSS}(y_1 = y_H) = w_2^{RSS}(y_1 = y_H) = \theta\beta + (1-\theta)\alpha$. Intuitively, as the cost of effort decreases to zero, all strategic types tend to exert high effort, and the resulting equilibrium tends to a pooling equilibrium where $\omega \to 1$. Consequently, the output signal is not informative, and the posteriors are equal to the prior.

Step 3. The first two steps imply that $\max_c \text{Var}(w_2^R)$ is realized at $c = c^R$ and $\min_c \text{Var}(w_2^R)$ is as $c \to 0$. Moreover,

$$\max_c \text{Var}(w_2^R) = \frac{\theta^2(\alpha-\beta)^2(\alpha\theta - \alpha - \beta\theta + \beta)^2}{-\alpha^2\theta^2 + 2\alpha^2\theta - \alpha^2 + 2\alpha\beta\theta^2 - 2\alpha\beta\theta - \alpha\theta + \alpha - \beta^2\theta^2 + \beta\theta}$$

$$\min_c \text{Var}(w_2^R) = 0$$

Step 4. Now, consider the narratives market. Notice that $\text{Var}(w_2^N)$ is constant in c for all c. The threshold level c^N makes the function jump up or down depending on the other parameter values. To compare $\text{Var}(w_2^N(c < c^N))$ and $\text{Var}(w_2^N(c > c^N))$ as $\alpha \to \beta$, we start by examining their expressions:

$$\text{Var}(w_2^N(c < c^N)) = (\alpha^2 + \beta - \beta\alpha)\left(1 - (\alpha^2 + \beta - \beta\alpha)\right)(\alpha - \beta)^2$$

$$\begin{aligned}\text{Var}(w_2^N(c > c^N)) &= \left(-\alpha^2\theta + \alpha^2 + 2\alpha\beta\theta - \alpha\beta - \beta^2\theta + \beta\right)\cdot \\ &\quad \left(1 - \left(-\alpha^2\theta + \alpha^2 + 2\alpha\beta\theta - \alpha\beta - \beta^2\theta + \beta\right)\right)(\alpha - \beta)^2\end{aligned}$$

Now, substituting $\alpha = \beta + \epsilon$, where ϵ is a small perturbation approaching 0, and expanding each function yields:

$$\begin{aligned}\text{Var}(w_2^N(c<c^N)) &\approx (\beta(1-\beta)+\epsilon^2(1-2\beta)+\beta\epsilon(1-2\beta))\epsilon^2\\ \text{Var}(w_2^N(c>c^N)) &\approx \big(\beta(1+(1-\theta\beta))+\epsilon^2(1-\theta\beta)+2\beta\epsilon(\theta-1)\big)\cdot\\ &\qquad \big(1-\big(\beta(1+(1-\theta\beta))+\epsilon^2(1-\theta\beta)+2\beta\epsilon(\theta-1)\big)\big)\,\epsilon^2\end{aligned}$$

Both Var$(w_2^N(c<c^N))$ and Var$(w_2^N(c>c^N))$ approach 0 as $\epsilon\to 0$. The leading coefficients in the expansions are crucial in determining the relative growth rates of Var$(w_2^N(c<c^N))$ and Var$(w_2^N(c>c^N))$. For small ϵ, Var$(w_2^N(c<c^N))$ has the term $\beta(1-\beta)\epsilon^2$, while Var$(w_2^N(c>c^N))$ has terms involving $\epsilon^2(1-\theta\beta)$ and $\epsilon\beta(\theta-1)$. When we compare the coefficients, Var$(w_2^N(c<c^N))$ has positive terms like $\beta(1-\beta)$, which is straightforwardly positive. Var$(w_2^N(c>c^N))$ has a more complex expression involving both positive and potentially negative terms depending on θ and β. Given this analysis, we can conclude that as α decreases towards β, Var$(w_2^N(c<c^N))$ will indeed grow faster and become larger than Var$(w_2^N(c>c^N))$.

Step 5. To show that as $\alpha\to 1$, Var$(w_2^N(c<c^N))$ gets smaller than Var$(w_2^N(c>c^N))$, we analyze the behavior of both functions as $\alpha\to 1$. After substituting $\alpha=1-\epsilon$, where ϵ is a small positive number, and simplifying we obtain:

$$\begin{aligned}\text{Var}(w_2^N(c<c^N)) &\approx 2\beta\epsilon(1-\beta-\epsilon)^2\\ \text{Var}(w_2^N(c>c^N)) &\approx (\theta+2\beta\theta-\beta-\beta^2\theta)(\theta-2\beta\theta+\beta+\beta^2\theta)(1-\beta-\epsilon)^2\end{aligned}$$

As α approaches 1 ($\epsilon\to 0$), the leading coefficient in Var$(w_2^N(c<c^N))$ is $2\beta\epsilon$, which diminishes faster compared to the leading coefficients in Var$(w_2^N(c>c^N))$. Hence, as α approaches 1, Var$(w_2^N(c<c^N))$ gets smaller than Var$(w_2^N(c>c^N))$.

Step 6. As $\alpha\to\beta$, to determine if $\max\limits_c \text{Var}(w_2^R) < \min\limits_c \text{Var}(w_2^N)$, let's compare the expressions for each as α approaches β. Substituting $\alpha = \beta + \epsilon$:

$$\max_c \text{Var}(w_2^R) \approx \frac{\theta^2\epsilon^4(\theta-1)^2}{-\beta^2\theta^2 + 2\beta^2\theta - \beta^2 + \beta}$$
$$\min_c \text{Var}(w_2^N) \approx (\beta(1-\beta\theta))\,(1-\beta(1-\beta\theta))\,\epsilon^2$$

To compare these, let's consider the leading terms as $\epsilon \to 0$. For $\max_c \text{Var}(w_2^R)$, the numerator is $\theta^2\epsilon^4(\theta-1)^2$. For $\min_c \text{Var}(w_2^N)$, the numerator is $(\beta(1-\beta\theta))\,(1-\beta(1-\beta\theta))\,\epsilon^2$. For $\max_c \text{Var}(w_2^R)$, the denominator is approximately $-\beta^2\theta^2 + 2\beta^2\theta - \beta^2 + \beta$. As $\epsilon \to 0$, ϵ^4 term in $\max_c \text{Var}(w_2^R)$ will go to zero faster than the ϵ^2 term in $\min_c \text{Var}(w_2^N)$. Hence, for sufficiently small ϵ, we have:

$$\max_c \text{Var}(w_2^R) \approx 0$$
$$\min_c \text{Var}(w_2^N) \approx (\beta(1-\beta\theta))\,(1-\beta(1-\beta\theta))\,\epsilon^2$$

Given this, it is clear that $\max_c \text{Var}(w_2^R) < \min_c \text{Var}(w_2^N)$. As α approaches β, $\text{Var}(w_2^R)$ goes to zero faster than $\text{Var}(w_2^N)$, making the inequality true.

Step 7. By Step 6. $\exists\, \alpha_1$ such that for all $\alpha < \alpha_1$, we have $\text{Var}(w_2^N) > \text{Var}(w_2^R)$ for all c by continuity.

Step 8. We have shown that as $\alpha \to 1$, $\text{Var}(w_2^N)$ diminishes faster than $\text{Var}(w_2^R)$ in Step 5. Consequently, there exists an α_2 such that for every $\alpha > \alpha_2$, we have $\text{Var}(w_2^N) < \text{Var}(w_2^R)$ for all c due to continuity. This completes the proof. ■

This theorem posits that when $\alpha - \beta = \mathbb{P}\text{r}(y_H \mid e_H) - \mathbb{P}\text{r}(y_H \mid e_L)$, indicating the responsiveness of output to effort, decreases, the variance of second-period wages in the narratives market surpasses that in the rational market. As a result, in sectors where production is less sensitive to effort, the narratives market leads to a broader distribution of second-period wages compared to the rational market.

Below is the visualization of the variance for both markets, with $\beta = 0.2$ and

$\theta = 0.5$ held constant, and α varied. The plot shows four subplots, each corresponding to different values of α ($\alpha = 0.21$, $\alpha = 0.55$, $\alpha = 0.91$, and $\alpha = 0.99$). In each subplot, the wage spread is plotted against the cost c. The blue line represents the variance of wages in the narratives market, while the red line represents the variance in the rational market. The dashed vertical lines indicate the thresholds c^N and c^R, where the wage spreads change significantly because the resulting equilibrium changes.

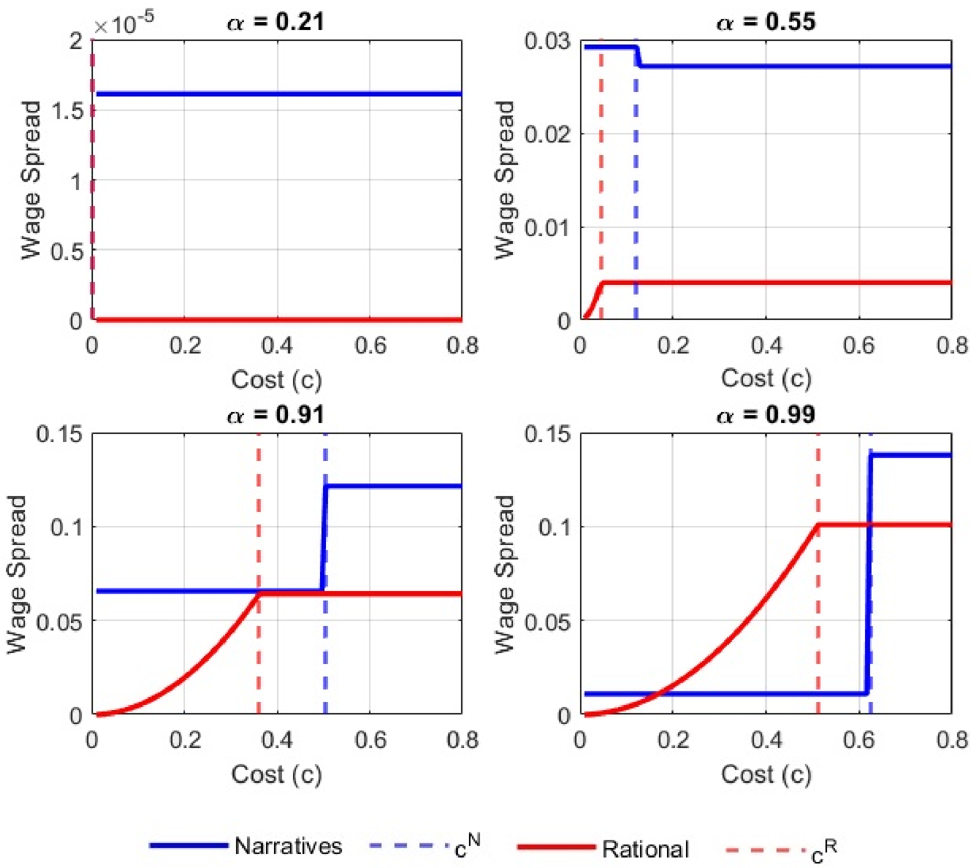

Figure 5.1: Wage Spread as a Function of Cost for Different Values of α: Narratives vs. Rational Market for $\beta = 0.2$ and $\theta = 0.5$

For lower values of α (e.g., $\alpha = 0.21$ and $\alpha = 0.55$), the variance in the rational market (Var(w_2^R)) is generally lower than in the narratives market (Var(w_2^N)) for all values of c. This is consistent with the first part of the theorem, which states that for $\alpha < \alpha_1$, $Var(w^R) < Var(w^N)$ for all $c > 0$. As α increases (e.g., $\alpha = 0.91$), we observe that Var(w_2^R) begins to increase more rapidly with c, and for some values of c, it can become comparable to or even exceed Var(w_2^N). This transition highlights the existence of the cutoff α_2, beyond which Var(w^R) can surpass Var(w^N) for higher values of α. In the case where α is very close to 1 (e.g., $\alpha = 0.99$), Var(w_2^R) is generally higher than Var(w_2^N) for most values of c. This supports the second part of the theorem, which asserts that for $\alpha > \alpha_2$, Var$(w^R) \geq$ Var(w^N) for all $c > 0$.

Moreover, according to the theorem, there exist parameter regions where the variance in wages created by the narratives market is higher than that produced by the rational market. Given Stylized Fact 2, which indicates that the variance of women's CPI-adjusted hourly wages is consistently higher than men's, especially in late career, we can infer that the labor market operates under different assumptions for women and men. The higher observed wage variance for women suggests that the labor market relies on narrative-driven beliefs when assessing female workers, leading to more extreme posteriors in their wage determination. In contrast, the relatively lower wage variance for men indicates that the market employs Bayesian rationality in evaluating male workers, resulting in fewer spread posteriors and, subsequently, fewer spread wages. This distinction highlights the differential treatment of gender in the labor market.

More importantly, recall that Phelpsian statistical discrimination theory empirically predicts that the wage variance for women should be smaller than that for men due to the noisier productivity signal associated with women. However, our empirical observations, as evidenced by Stylized Fact 2, reveal the opposite: women's wage variance is higher than men's. This discrepancy underscores the need for a model accurately reflecting this observed stylized fact. By accounting for a higher variance in women's wages within certain parameter regions, our model provides a more accurate representation of the labor market dynamics for the discriminated group.

Furthermore, the Stylized Fact 3 states that there exists a pronounced difference in wage variance between the service and manufacturing sectors, particularly for late-career women. This observation corresponds with Theorem 1, suggesting a relationship between output sensitivity to effort and the variance of second-period wages. Given that the service sector generally exhibits lower sensitivity to effort than the manufacturing sector, the narratives market results in a broader distribution of second-period wages within the service sector as $\alpha - \beta$ decreases. The higher variance observed in late-career women's wages in the service sector, compared to manufacturing, implies that the labor market employs narratives-based wage determination for women, leading to more extreme wage outcomes. In contrast, for men, the labor market seems to follow Bayesian updating, resulting in less spread in wage distribution.

5.2 Late Career Gender Wage Gap

The gender wage gap is the difference in men's and women's median earnings divided by men's median earnings. We have argued by Theorem 1 and Stylized Fact 2 that for women, the labor market believes in narratives, and for men, it is Bayesian. Therefore, the gender wage gap for late-career implied by this model is given by:

$$\frac{\mathbb{E}[{w_2}^R] - \mathbb{E}[{w_2}^N]}{\mathbb{E}[{w_2}^R]}$$

Below, Figure 5.2 presents the simulation results from the model, illustrating the relationship between the gender wage gap and α for different values of cost. The simulations reveal a hump-shaped pattern in the gender wage gap as a function of α. Starting from a gender wage gap of zero at $\alpha = \beta$, the gender wage gap increases, reaching its peak, and then declines, becoming negative as α approaches 1. This indicates that a positive gender wage gap emerges only when α is less than a certain threshold α^*. Consequently, industries characterized by low α

values—where inference is more challenging, such as service industries—are more likely to exhibit a larger wage gap. This finding aligns with Stylized Fact 4, which observed a higher wage gap in service sectors compared to manufacturing, particularly for women in their late careers.

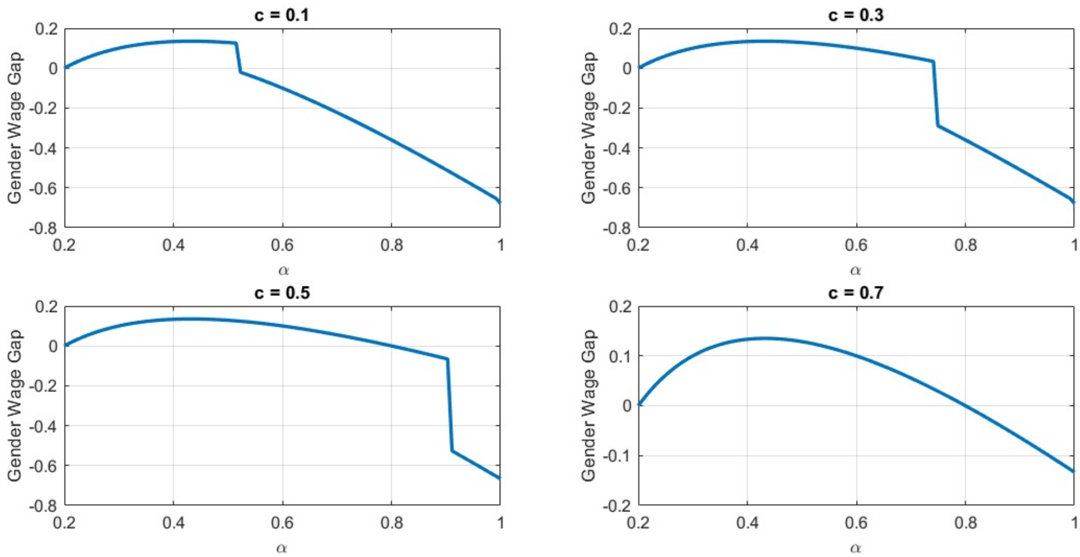

Figure 5.2: Gender Wage Gap as a Function of α for Different Values of Cost for $\beta = 0.2$ and $\theta = 0.5$

5.3 Welfare Analysis

This subsection extends the analysis to a normative perspective by comparing the total welfare for each market. The following theorems establish the conditions under which workers subjected to the narratives market are better off or worse off. We denote the second period's expected payoff when the market is rational with $\mathbb{E}U^R$, and when the market believes in narratives, it is $\mathbb{E}U^N$. The first-period wage is not included because it is just a shift parameter which is the same for all types, markets, and equilibria. Since the cost affects the type of equilibrium and, subsequently, the behavior and payoff, the total expected payoffs of both types for both markets are given as piecewise functions of c as follows:

$$\mathbb{E}U^R = \begin{cases} \theta \cdot \Big[\alpha \cdot \Big(\hat{\theta}_{RSS} \cdot \beta + (1-\hat{\theta}_{RSS}) \cdot \alpha\Big) \\ \quad +(1-\alpha) \cdot \Big(\tilde{\theta}_{RSS} \cdot \beta + (1-\tilde{\theta}_{RSS}) \cdot \alpha\Big) - c\Big] \\ \quad + (1-\theta) \cdot \Big[\alpha \cdot \Big(\hat{\theta}_{RSS} \cdot \beta + (1-\hat{\theta}_{RSS}) \cdot \alpha\Big) \\ \quad +(1-\alpha) \cdot \Big(\tilde{\theta}_{RSS} \cdot \beta + (1-\tilde{\theta}_{RSS}) \cdot \alpha\Big)\Big] & \text{if } c < c^R \\ \theta \cdot \Big[\beta \cdot \Big(\hat{\theta}_{RS} \cdot \beta + (1-\hat{\theta}_{RS}) \cdot \alpha\Big) \\ \quad +(1-\beta) \cdot \Big(\tilde{\theta}_{RS} \cdot \beta + (1-\tilde{\theta}_{RS}) \cdot \alpha\Big)\Big] \\ \quad + (1-\theta) \cdot \Big[\alpha \cdot \Big(\hat{\theta}_{RS} \cdot \beta + (1-\hat{\theta}_{RS}) \cdot \alpha\Big) \\ \quad +(1-\alpha) \cdot \Big(\tilde{\theta}_{RS} \cdot \beta + (1-\tilde{\theta}_{RS}) \cdot \alpha\Big)\Big] & \text{if } c \geq c^R \end{cases}$$

$$\mathbb{E}U^N = \begin{cases} \theta \cdot \big[\alpha^2 + (1-\alpha)\beta - c\big] \\ \quad + (1-\theta) \cdot \big[\alpha^2 + (1-\alpha)\beta\big] & \text{if } c < c^N \\ \theta \cdot [\beta\alpha + (1-\beta)\beta] \\ \quad + (1-\theta) \cdot \big[\alpha^2 + (1-\alpha)\beta\big] & \text{if } c \geq c^N \end{cases}$$

The expected payoffs are weighted averages of strategic and behavioral workers' payoffs according to the proportion of each type. To understand how the total payoffs are defined, consider the rational market. When $c < c^R$, $\theta\omega$ portion of workers are strategic and choose to work, with α probability they succeed and $1-\alpha$ they fail. The rest of the strategic workers, i.e., $\theta(1-\omega)$ choose not to work. However, by the incentive compatibility condition, strategic types are indifferent between working and not working, so the payoff associated with each action is equal. Hence, it is sufficient to include only one of them and weigh the payoffs with θ. Likewise, the $1-\theta$ portion is behavioral workers who choose to work and succeed with probability α and fail with probability $1-\alpha$. When $c > c^R$ strategic workers do not work and with β probability, they succeed, and $1-\beta$ they fail. The rest of the function is formed similarly.

The next theorem assesses the cases where the difference between α and β is large, implying a less responsive production function to the effort.

Theorem 2 *Fixing β and θ, there exist a cutoff level $\overline{\alpha}$ where $0 < \beta < \overline{\alpha} < 1$ such that if $\alpha > \overline{\alpha}$, then $\mathbb{E}U^N > \mathbb{E}U^R$ for all $c > 0$.*

Proof. We will demonstrate that there exists a threshold value for α beyond which the expected utility under the narratives setting exceeds that under the rational setting for any positive cost.

Step 1. We demonstrated in the proof of Theorem 1. that, as $c \to 0$, we have $\omega \to 1$ and $w_2^{RSS}(y_1 = y_H) = w_2^{RSS}(y_1 = y_L)$. Hence, $\lim_{c\to 0} \mathbb{E}U^R = \alpha - \theta\alpha + \theta\beta$. Furthermore, $\mathbb{E}U^R[c = c^R] = \alpha - \theta\alpha + \theta\beta$. Also, notice that for all $c > c^R$, the rational expected payoff is constant in c. Hence, $\lim_{c\to 0} \mathbb{E}U^R = \mathbb{E}U^R[c = c^R]$.

Step 2. We observe that $\mathbb{E}U^R$ is convex in c for all $c < c^R$ using Mathematica, but we do not include the first and second derivatives here due to their complexity.

Step 3. The first two steps imply that $\max_c \mathbb{E}U^R$ is attained when $c \geq c^R$ (or $c \to 0$) and $\min_c \mathbb{E}U^R$ is for $c^* \in (0, c^R)$.

Step 4. Consider $\mathbb{E}U^N$. For all $c < c^N$, $\frac{\partial \mathbb{E}U^N}{\partial c} = -\theta < 0$ and for all $c > c^N$, $\frac{\partial \mathbb{E}U^N}{\partial c} = 0$. Therefore, $\max_c \mathbb{E}U^N$ is attained as $c \to 0$ and $\min_c \mathbb{E}U^N$ is when $c \geq c^N$. Moreover, $\min_c \mathbb{E}U^N = \alpha^2 + \beta(1-\alpha) - \theta(\alpha - \beta)^2$.

Step 5. We have closed-form expressions in (α, β, θ) for $\min_c \mathbb{E}U^N$ and $\max_c \mathbb{E}U^R$. As $\alpha \to 1$, $\min_c \mathbb{E}U^N \to 1 - \theta(1-\beta)^2$ and $\max_c \mathbb{E}U^R \to 1 - \theta + \theta\beta$. Therefore, as $\alpha \to 1$, $\min_c \mathbb{E}U^N > \max_c \mathbb{E}U^R$. By Berge's Maximum Theorem, since these expected payoff functions are the value functions of expected payoff maximization, they are continuous. Therefore, there exists $\overline{\alpha}$ such that for all $\alpha > \overline{\alpha}$, $\mathbb{E}U^N > \mathbb{E}U^R$ for all $c > 0$. ■

This theorem suggests that when $\alpha - \beta = \Pr(y_H \mid e_H) - \Pr(y_H \mid e_L)$ is large, i.e., the responsiveness of output to the effort is high, the market believing in narratives does not harm workers. That is because the total expected payoff of all workers when the market is rational is less than when the market is boundedly rational.

The next theorem investigates the cases where the difference between α and β is small, implying a more responsive production function to the effort, but this time we have an additional condition regarding the difficulty of the task as well.

Theorem 3 *Fixing β and θ, there exist another cutoff level $\underline{\alpha}$ where $0 < \beta < \underline{\alpha} < \overline{\alpha} < 1$ such that*

- *If $\alpha < \underline{\alpha}$, then $\mathbb{E}U^N < \mathbb{E}U^R$ if and only if $\beta < 1 - \theta$ for all $c > 0$*
- *If $\alpha < \underline{\alpha}$, then $\mathbb{E}U^N > \mathbb{E}U^R$ if and only if $\beta > 1 - \theta$ for all $c > 0$.*

Proof. We will show that there exists a threshold value for α below which the expected utility under the narratives setting is less than that under the rational setting if and only if $\beta < 1 - \theta$ for any positive cost.

Step 1. Observe that as $\alpha \to \beta$, $c^R \to 0$ since $c^R = (\tilde{\theta}_{RS} - \hat{\theta}_{RS})(\alpha - \beta)^2$. Therefore, $\mathbb{E}U^R$ becomes a constant function of c. Then, as $\alpha \to \beta$ $\max_c \mathbb{E}U^R = \min_c \mathbb{E}U^R = \alpha - \theta\alpha + \theta\beta$.

Step 2. Moreover, as $\alpha \to \beta$ we also have $c^N \to 0$ since $c^N = (\alpha - \beta)^2$. Thus, $\max_c \mathbb{E}U^N = \min_c \mathbb{E}U^N = \alpha^2 + \beta(1 - \alpha) - \theta(\alpha - \beta)^2$.

Step 3. Now, substitute $\alpha = \beta + \epsilon$ into $\min_c \mathbb{E}U^R$ and $\max_c \mathbb{E}U^N$ where ϵ is a small perturbation approaching 0. We have $\min_c \mathbb{E}U^R = \beta + \epsilon(1 - \theta)$ and $\max_c \mathbb{E}U^N = \epsilon^2(1-\theta)+\beta(\epsilon+1)$. Consider $\min_c \mathbb{E}U^R > \max_c \mathbb{E}U^N$. After simplifications, this inequality boils down to $1 - \theta > \epsilon + \beta - \theta\epsilon$. As $\epsilon \to 0$ we have $1 - \theta > \beta$. By continuity, there exists a cutoff level $\underline{\alpha}$ such that for all $\alpha > \overline{\alpha}$, $\mathbb{E}U^N < \mathbb{E}U^R$ for all $c > 0$ if and only if $1 - \theta > \beta$.

Step 4. Otherwise, i.e. when $\beta > 1 - \theta$ we have $\min_c \mathbb{E}U^R < \max_c \mathbb{E}U^N$. Since both expected payoff functions are constant in c as $\alpha \to \beta$, this implies $\max_c \mathbb{E}U^R < \min_c \mathbb{E}U^N$. By continuity, there exists a cutoff level $\underline{\alpha}$ such that for all $\alpha > \overline{\alpha}$, $\mathbb{E}U^N > \mathbb{E}U^R$ for all $c > 0$ if and only if $1 - \theta < \beta$. ■

This theorem indicates that when $\alpha - \beta$ is small, i.e., the responsiveness of output to the effort is low and the additional condition $\beta < 1 - \theta = \mathbb{P}\text{r(behavioral type)}$ is satisfied, the market believing in narratives does harm workers. That is because the total expected payoff of all workers when the market is rational exceeds the case where the market is boundedly rational. This extra condition entails that the probability of success given not working is lower than the portion of the behavioral types that always work. Hence, these are more difficult tasks because succeeding in them without working is less probable. Consequently, the theorem states when the task is arduous and less responsive to effort, being subjected to the narrative market rather than the rational market harms the workers. On the other hand, if the task is easier, although the output is less responsive to effort, the market's belief in narratives does not harm workers.

To illustrate Theorems 2 and 3 regarding the expected payoffs under different setups, we can analyze Figure 5.3. The visualizations in the plot depict the expected payoffs (denoted as $\mathbb{E}U$) for both the narratives market (blue line) and the rational market (red line) as functions of cost (c) under different values of α. The different subplots represent different fixed values of α.

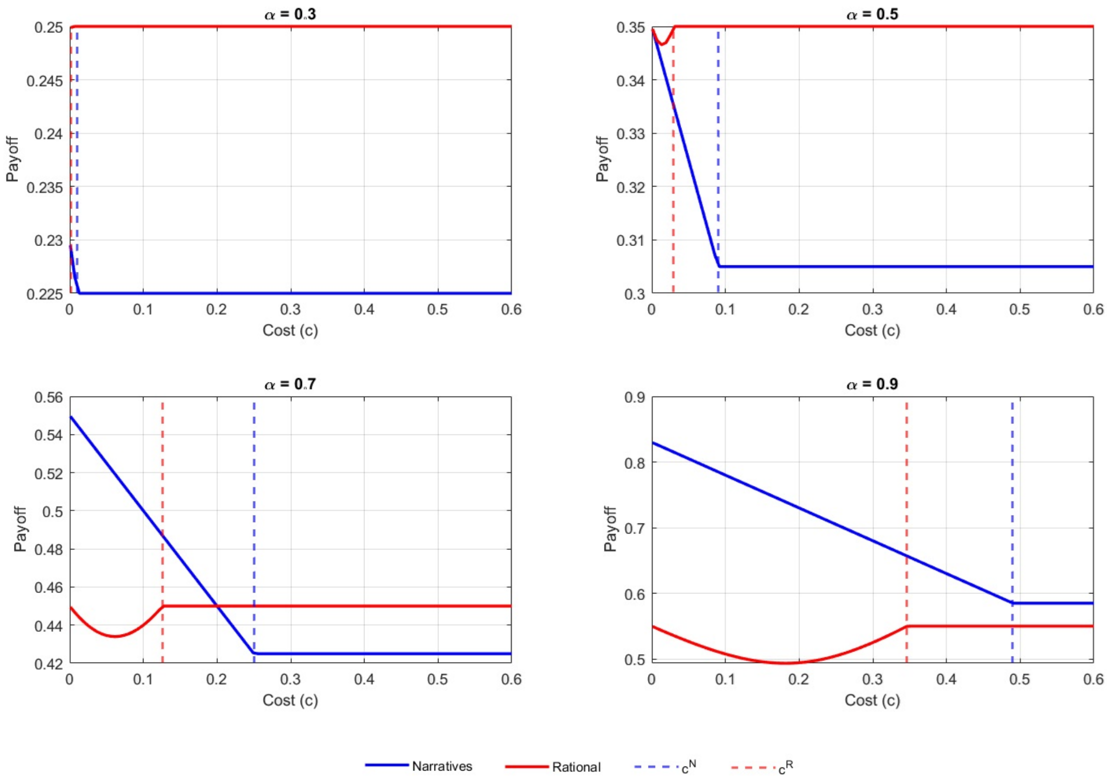

Figure 5.3: Expected Payoffs as Functions of Cost for Different Values of α: Narratives vs. Rational Market for $\beta = 0.2$ and $\theta = 0.5$

In the top-left subplot ($\alpha = 0.3$), the expected payoffs for the rational market are consistently higher than for the narratives market across all cost levels. This supports

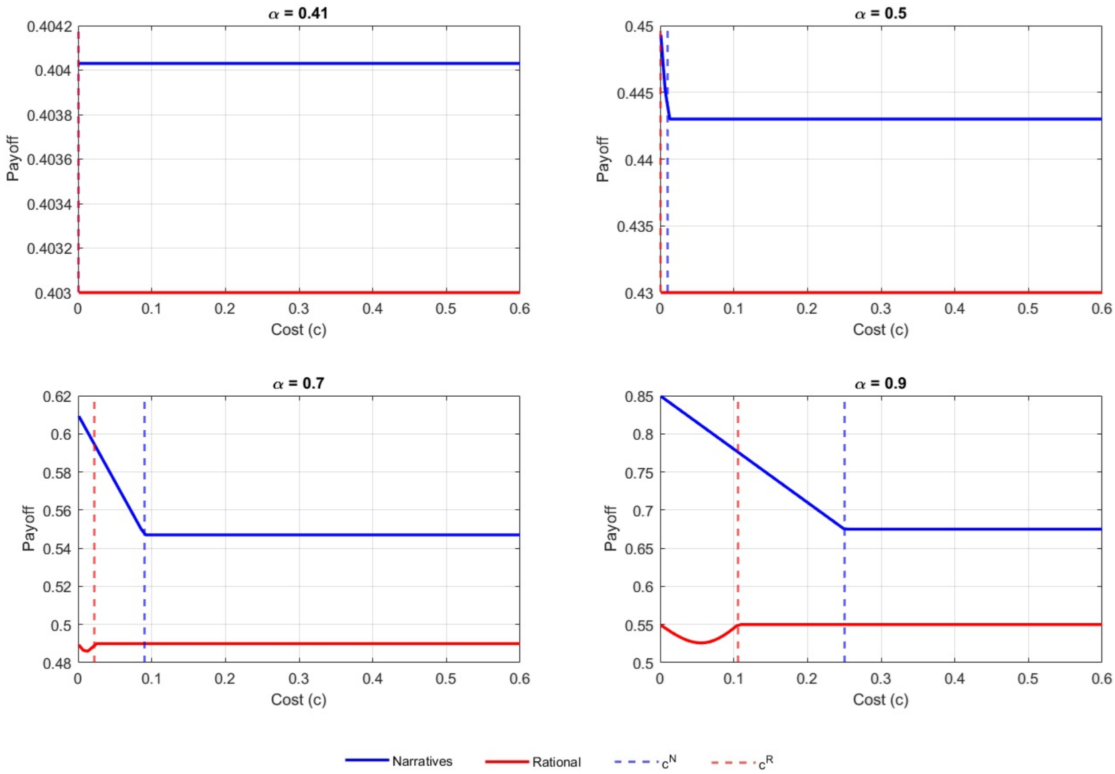

Figure 5.4: Expected Payoffs as Functions of Cost for Different Values of α: Narratives vs. Rational Market for $\beta = 0.4$ and $\theta = 0.7$

the part of Theorem 2 which states that when α is below a certain threshold, i.e., $\underline{\alpha}$, $\mathbb{E}U^R$ exceeds $\mathbb{E}U^N$ for all $c > 0$ if $\beta < 1 - \theta$, as $\beta = 0.2 < 0.5 = 1 - \theta$. In the top-right subplot ($\alpha = 0.5$), the expected payoffs for the rational market are still higher than the narratives market, but the difference narrows as the cost decreases. This illustrates the transition towards the region where the narratives market starts to provide higher payoffs for higher values of α and certain cost levels.

In the bottom-left subplot ($\alpha = 0.7$), as α increases further, we observe a region where the narratives market starts to have higher expected payoffs compared to the rational market for specific cost levels. This is in line with the transition described in Theorem 3 where α approaches $\overline{\alpha}$. In the bottom-right subplot ($\alpha = 0.9$), at high values of α, the narratives market consistently yields higher expected payoffs than the rational market, especially for lower costs. This supports Theorem 2, indicating that for α greater than $\overline{\alpha}$, $\mathbb{E}U^N$ is greater than $\mathbb{E}U^R$ for all $c > 0$.

Figure 5.4 below illustrates the expected payoffs in both the narratives and rational markets, with parameters β and θ fixed at 0.4 and 0.7, respectively, to capture the second part of Theorem 3, as $\beta = 0.4 > 1 - 0.7 = 1 - \theta$. We observe that for smaller values of $\alpha < \underline{\alpha}$, since the condition $\beta < 1 - \theta$ is not met, $\mathbb{E}U^N$ exceeds $\mathbb{E}U^R$ for all $c > 0$. Furthermore, for larger values of $\alpha > \overline{\alpha}$, we see that $\mathbb{E}U^N$ is greater than $\mathbb{E}U^R$ for all $c > 0$, which aligns with Theorem 2.

The findings from Theorem 2 suggest that in sectors where the output is more directly tied to effort, such as manufacturing, the narratives market does not render women worse off. In these industries, $\alpha - \beta$ is relatively high, indicating that the responsiveness of output to effort is significant. This implies that employers can more accurately infer the productivity of employees, leading to a situation where the application of narratives by the labor market does not disadvantage women. As a result, in manufacturing sectors, the variance in wages and the gender wage gap is lower, supporting the notion that women are not disproportionately affected by narratives. This finding aligns with the observed lower variance in women's wages in manufacturing, as shown in our data, highlighting the alignment between theoretical predictions and empirical evidence.

Conversely, Theorem 3 provides critical insights into the service sector, where $\alpha - \beta$ is low, indicating that output is less directly tied to the effort, making inference more challenging. In this context, particularly when the task is difficult ($\beta < 1 - \theta$), the application of narratives by the labor market results in significantly worse outcomes for women. The theorem posits that in such scenarios, $\mathbb{E}U^N < \mathbb{E}U^R$, implying that narratives lead to a lower expected utility for women compared to a rational market. This is evident in the higher wage variance observed in the service sector for women, suggesting that the narratives market generates more extreme wage outcomes. Therefore, the service sector's reliance on narratives exacerbates gender wage disparities, particularly in late career stages, where the variance is substantially higher than in manufacturing. This underscores the importance of addressing the differential impact of market narratives on women's welfare across different sectors.

CHAPTER 6

CONCLUSION

This thesis has undertaken an examination of the gender wage gap, with a specific focus on the late career stage, using a dataset from the Turkish Household Labor Force Survey spanning 2005 to 2023. By integrating elements of contract theory and narratives, this study provides an understanding of the factors contributing to wage disparities between men and women. The theoretical model developed in this research offers insights into the mechanisms through which implicit incentives, stochastic cost heterogeneity, and different belief-updating processes influence wage outcomes and perpetuate the gender wage gap.

The two-period career concerns model constructed in this thesis sheds light on the interplay between workers' efforts, market perceptions, and wage determination. The model incorporates both rational Bayesian updating and narrative-driven extremist updating mechanisms to explore how different belief systems affect labor market outcomes. The findings suggest that the narrative-driven updating mechanism, which associates women's success strictly with high effort and failure with low effort, exacerbates wage variance for women and reinforces gender wage gaps, especially in the service sector compared to manufacturing.

Empirical analysis using the Turkish Household Labor Force Survey data corroborates the theoretical predictions of the model. The data reveals that women experience significantly higher wage variance than men in both early and late career stages. This variance increases more rapidly for women as they advance in their careers, indicating a more dispersed wage distribution compared to men. Additionally, the gender wage gap is found to widen considerably in the late career stage, particularly in the service sector, where production is generally less sensitive to effort compared to the manufacturing sector due to the inference problem.

The sectoral analysis highlights the critical role of industry-specific characteristics in shaping wage dynamics. Women in the service sector face greater wage variability and a more pronounced wage gap than their counterparts in the manufacturing sector. This finding underscores the importance of considering sector-specific factors when addressing gender wage disparities. Policies aimed at reducing the gender wage gap should, therefore, be tailored to the unique characteristics of different industries to effectively target the underlying causes of inequality.

A key contribution of this thesis is its innovative perspective on discrimination by integrating contract theory and narrative literature. Traditional statistical discrimination theory suggests that the variance in wages for the discriminated group should be lower than for the non-discriminated group. However, the empirical data from this study contradicts this implication, showing higher wage variance for women. By incorporating narratives into the model, this thesis offers a new explanation for this discrepancy, highlighting how market perceptions and biases due to preexisting narratives, i.e., societal gender expectations, shape wage dynamics and contribute to gender inequality.

Future research could build on the findings of this thesis by exploring the impact of different policy interventions, such as wage transparency, on the gender wage gap. Experimental and longitudinal studies could provide further insights into the effectiveness of various measures aimed at reducing wage disparities. Moreover, expanding the analysis to include other demographic variables such as race, ethnicity, educational background, and the experiences of LGBTQ+ individuals, as well as all marginalized identities, could offer a more comprehensive understanding of the multifaceted nature of wage inequality.

In conclusion, this thesis highlights the persistent and complex nature of the gender wage gap, particularly in the late career stage. By integrating theoretical models with empirical data, it provides a robust framework for understanding the mechanisms driving wage disparities between men and women. The findings underscore the need for targeted policy interventions that address biases stemming from established social norms in the labor market. Through continued efforts to promote equality and inclusivity, significant progress can be made toward closing the gender wage gap and ensuring fair and equitable wage outcomes for all.

REFERENCES

Aigner, D. J. and Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review*, 30(2):175–187. 16

Aina, C. (2023). Tailored stories. [Unpublished manuscript]. 12, 52

Alonso, R. and Câmara, O. (2016). Persuading voters. *American Economic Review*, 106(11):3590–2605. 12

Aumann, R. J., Maschler, M., and Stearns, R. E. (1995). *Repeated games with incomplete information*. MIT Press. 12

Barron, K. and Fries, T. (2023). Narrative persuasion. (CESifo Working Paper No. 10206). http://dx.doi.org/10.2139/ssrn.4329465. 12

Becker, G. S. (1957). *The economics of discrimination*. University of Chicago Press. 13

Bénabou, R., Falk, A., and Tirole, J. (2018). Narratives, imperatives, and moral reasoning. (NBER Working Paper No. w24798). https://ssrn.com/abstract=3214341. 12

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, 50(6):1431–1451. 11

Dewatripont, M., Jewitt, I., and Tirole, J. (1999). The economics of career concerns, part ii: Application to missions and accountability of government agencies. *The Review of Economic Studies*, 66(1):199–217. 8

Eliaz, K. and Spiegler, R. (2020). A model of competing narratives. *American Economic Review*, 110(12):3786–3816. 12

Ely, J. C. (2017). Beeps. *American Economic Review*, 107(1):31–53. 12

Euronews (2024). Gender pay gap in europe: How do countries compare on narrowing the divide? Accessed: 2024-07-25. 25

European Commission (1998). *One hundred words for equality – A glossary of terms on equality between women and men*. Publications Office. 1

Galperti, S. (2019). Persuasion: The art of changing worldviews. *American Economic Review*, 109(3):996–1031. 12

Guesnerie, R. and Laffont, J.-J. (1984). Indirect public control of self-managed monopolies. *Journal of Comparative Economics*, 8(2):139–158. 6

Harris, M. and Holmström, B. (1982). A theory of wage dynamics. *The Review of Economic Studies*, 49(3):315–333. 7

Holmström, B. (1999). Managerial incentive problems: A dynamic perspective. *The Review of Economic Studies*, 66(1):169–182. 6

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615. 12

Kolotilin, A., Mylovanov, T., Zapechelnyuk, A., and Li, M. (2017). Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964. 12

Little, A. T. (2023). Bayesian explanations for persuasion. *Journal of Theoretical Politics*, 35(3):147–181. 11

Mirrlees, J. A. (1999). The theory of moral hazard and unobservable behaviour: Part i. *The Review of Economic Studies*, 66(1):3–21. 6

OECD (2023). *OECD Employment Outlook 2023*. Organisation of Economic Co-operation and Development. 1

Onuchic, P. (2023). Recent contributions to theories of discrimination. [Manuscript submitted for publication]. 15

Ortiz-Ospina, E., Hasell, J., and Roser, M. (2018). Economic inequality by gender. Our World in Data. Accessed: 2024-07-25. 25

Phelps, E. S. (1972). The statistical theory of racism and sexism. *American Economic Review*, 62:659–661. 14

Schwartzstein, J. and Sunderam, A. (2021). Using models to persuade. *American Economic Review*, 111(1):276–323. 12, 52

Shiller, R. J. (2017). Narrative economics. *American Economic Review*, 107(4):967–1004. 12

Shiller, R. J. (2020). *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press. 12