What Prevents Productivity Gains from Translating into Wages: The Role of Deflation and Labor Market Frictions *

Nobuhiro Abe† Ryuichiro Hirano‡ Sohei Kaihatsu§
Bank of Japan

Under construction. Please do not circulate without the authors' permission.

January 30, 2026

Abstract

What prevents productivity gains from translating into wages? This paper disentangles the role of deflation and labor market frictions, particularly the prioritization of job security, as the causes of the widening wage-productivity gap. The empirical analysis reveals that wage-setting in Japan had been rigid in both upward and downward directions during the deflation period. Notably, upward wage rigidity during the deflationary period outweighed downward rigidity. This wage freeze led to subdued wage growth, resulting in the widening wage-productivity gap. With a heterogenous firm general equilibrium model, we quantitatively show that rigid wage-setting practices, coupled with deflation and the prioritization of job security, have induced firms to impose larger wage markdowns, suppressed wage growth, and hence widened the wage-productivity gap. Mitigating these factors contributes to increases in not only wages but also aggregate productivity through the reallocation effect.

Keywords: Wage-Productivity Gap, Downward Nominal Wage Rigidity, Job Security, Trend Inflation, Reallocation Effect
JEL classification: E24, E52, J30

*We are grateful to Kosuke Aoki, Ippei Fujiwara, Masao Fukui, Shinnosuke Kikuchi and Toshitaka Maruyama for their valuable comments and discussions. We thank Hiroshi Fujiki, Shigeru Fujita, Daisuke Ikeda, Mike Irwin, Yuichiro Ito, Takuji Kondo, Takushi Kurozumi, Koji Nakamura, Akio Okuno, the staff of the Bank of Japan and participants at the BOC-Philadelphia Fed-BOJ Joint Workshop for their comments. Wakako Kaku provided excellent research assistance. We acknowledge that the Ministry of Health, Labour and Welfare provides data from the *Basic Survey on Wage Structure*, and the Ministry of Internal Affairs and Communications provides data from the *Establishment Census, Establishment Directory Maintenance Survey, Establishment and Enterprise Census, and Economic Census for Business Frame.* The views expressed in this paper are those of the authors and do not necessarily reflect those of the Bank of Japan. Any remaining errors are those of the authors.

†Email: nobuhiro.abe@boj.or.jp
‡Email: ryuuichirou.hirano@boj.or.jp
§Email: souhei.kaihatsu@boj.or.jp

1 Introduction

When comparing the trends in productivity and real wages in the United States, Germany, France, and Japan, the gap between the two has widened over the past 30 years in the United States and Japan (Figure 1). Notably, this gap is more pronounced in Japan. The main aim of this paper is to study the factors that contribute to the differences between Western countries and Japan in the wage-productivity gap, which refers to the disparity between real wages and productivity.

This paper focuses on two key factors underlying the gap between real wages and productivity: (1) the role of inflation and deflation in rigid wage setting (the "Grease the Wheels Effect" of inflation in Tobin (1972)), and (2) the decline in workers' bargaining power in the labor market (e.g., Grossman and Oberfield (2022), Berger et al. (2022)). We study these two factors in two steps. First, as a fact finding, we conduct an empirical analysis using microdata of Japanese workers to determine whether wage setting is characterized by a wage freeze, which includes both upward and downward nominal wage rigidity (e.g., Daly and Hobijn (2014), Schaefer and Singleton (2023)). To discuss the role of the two factors, we conduct a subsample analysis, which covers both deflationary and non-deflationary periods, and examine the differences between these sub-sample periods.

Second, by employing a heterogeneous firm model that incorporates downward nominal wage rigidity (henceforth, DNWR), we quantitatively evaluate the extent to which (1) the decline in trend inflation and (2) the prioritization of job security during labor-management negotiations have contributed to the wage-productivity gap through nominal wage rigidity. Additionally, we study the mechanism by which the two factors, coupled with nominal wage rigidity, affect allocative efficiency and macro-level productivity through the worker allocation channel.

Our main findings and contributions to the literature can be summarized as follows. First, an estimation framework that simultaneously evaluates both upward and downward nominal wage rigidities is applied to extensive longitudinal microdata on individual workers in Japan, covering the

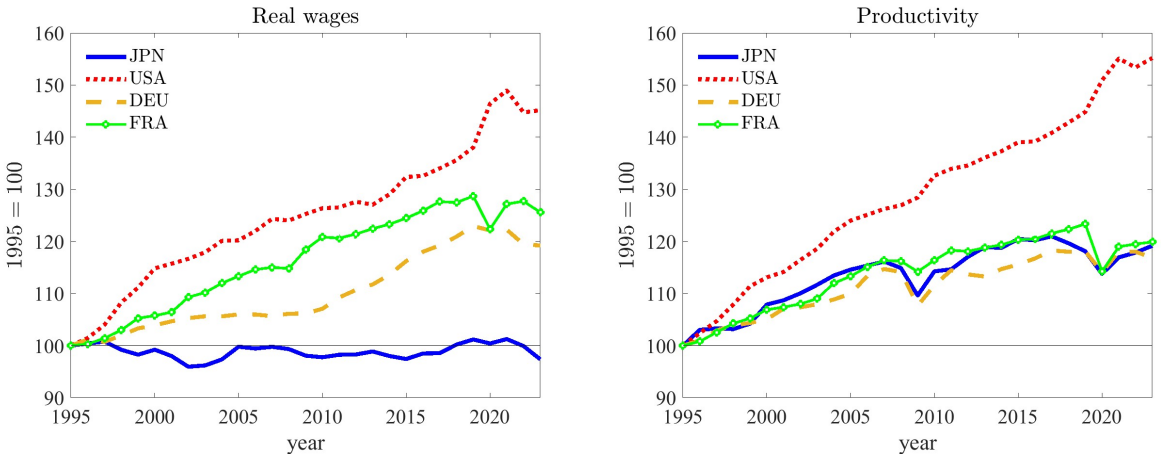


Figure 1: Real wages and productivity in the advanced economies

long period from 1990 to 2024. We find that both upward and downward nominal wage rigidity have been present in Japan over the last three decades. More importantly, we also find that the impact of upward rigidity strengthened after the late 1990s, when Japan entered a period of deflation, exerting downward pressure on actual wages. In this regard, although previous analyses have conducted quantitative evaluations only on downward rigidity, to the best of our knowledge, no studies have simultaneously evaluated upward rigidity within the context of long-term microdata.

As a theoretical contribution, this paper provides a dynamic general equilibrium framework to analyze two factors that amplify the impact of wage freezes: (1) a decline in trend inflation and (2) the prioritization of job security in labor-management negotiations. Specifically, we extend the partial equilibrium model of heterogeneous firms developed by Ehrlich and Montes (2024) into a general equilibrium framework.[1] Since inflation is important in its own right, we introduce both Rotemberg-type price adjustment costs and monetary policy, though our primary focus remains on the effects of exogenous productivity shocks. The construction of a general equilibrium model with heterogeneous firms enables us to investigate the impact of exogenous increases in aggregate productivity on macro-level wages and endogenous productivity, as well as the allocation effects of workers across firms.[2]

We show that, under DNWR, higher trend inflation and a greater preference for relative wages over job security lead to higher macro-level wages and productivity. The key mechanism is that higher inflation reduces the costs incurred by firms due to DNWR. We demonstrate that relaxing nominal wage rigidity can improve the efficiency of labor allocation by enhancing worker mobility based on wage signals, thereby increasing macro-level productivity. Furthermore, in an environment where relative wages are prioritized, workers are more likely to move toward high-productivity firms offering higher wages, which elevates the overall levels of wages and productivity at the macroeconomic level. We conduct both theoretical and quantitative analyses of these mechanisms under both static and dynamic settings. Since we analyze heterogeneous firms, our study contributes to the literature on heterogeneous households and firms (e.g., Ottonello and Winberry (2020), Auclert et al. (2021)). Moreover, our examination of the implications of worker allocation across firms connects our paper to the literature on allocative efficiency (e.g., Baqaee and Farhi (2020)).

Why we focus on deflation and preference on job security? One of the primary factors contributing to the wage-productivity gap is the rigidity of nominal wages. Under conditions of wage freezes, where nominal wages are inflexible both upwards and downwards, adjustments to wages

[1] Ehrlich and Montes (2024) propose a quasi-monopsony structure in which firms set wages without bargaining power on the side of workers, but workers can quit firms if their relative wages are low. Using this partial equilibrium model, they analyze the relationship between downward nominal wage rigidity and employment in the German labor market. For more details about the model, see Section 4.

[2] With a similar motivation, de la Barrera and Fukui (2025) propose a framework that integrates firm dynamics á la Hopenhayn and Rogerson (1993), wage posting with on-the-job search á la Burdett and Mortensen (1998), and DNWR in wage settings. They use this framework to quantify the efficiency losses caused by DNWR by comparing with a counterfactual economy with flexible wages.

in line with workers' productivity are hindered, widening the wage-productivity gap. In such an environment, various elements that suppress the nominal wage growth can increase the number of workers facing nominal wage rigidity. This issue is particularly pronounced in Japan, where the wage-productivity gap is notably wider compared to other advanced economies. This study, therefore, focuses on two key factors unique to Japan that contribute to the suppression of nominal wage growth, amplifying the impact of wage freezes: (1) deflation and (2) the labor market practice in Japan of prioritizing job security over wage increases.

First, with regard to the relationship between inflation (deflation) and nominal wage rigidity, it has traditionally been pointed out that when the inflation rate is positive, the adjustment of wages can be achieved through a decline in real wages without reducing nominal wages, thus mitigating the effects of nominal rigidity. This phenomenon is known as the "Grease the Wheels Effect" (Tobin (1972)). Regarding Japan, unlike other advanced economies, Japan has experienced a protracted period of deflation since the late 1990s. This suggests that the "Grease the Wheels Effect" has not been functioning effectively in Japan, and prolonged deflation may have contributed to the widening of the wage-productivity gap.

Second, the prioritization of job security in labor negotiations - or more generally, the decline in workers' bargaining power - has been identified in previous studies as one of the structural factors that have suppressed labor's share of income (e.g., Grossman and Oberfield (2022), Bergholt et al. (2022), Karabarbounis (2024)). Several prior studies focusing on Japan have similarly pointed out that the decline in workers' bargaining power has contributed to the reduction in labor's share of income (Dobbelaere and Kiyota (2018), Goodhart and Pradhan (2020)). In the context of declining bargaining power, it has been observed that job security started to be prioritized over wage increases following the deterioration of corporate earnings in the late 1990s (Ogura (2017); Watanabe (2024)). These Japanese employment practices may have further restrained the growth of nominal wages, thereby exacerbating the effects of nominal wage rigidity.

Structure This paper is structured as follows. Section 2 reviews relevant previous studies and explains the positioning of this paper in relation to them. Section 3 describes the estimation framework for the empirical analysis and presents the estimation results. In Section 4, we describe our theoretical framework and calibration strategy. Section 5 discusses analytical results on how inflation and preference for job security affect macro wage and productivity. In Section 6, we discuss the dynamic implications of the model under a positive productivity shock and show the results of counterfactual simulations under higher inflation and lower job security preference. Section 7 concludes the paper.

2 Extensive Literature Review

The analysis in this paper is related to (1) the labor share of income and wage markdown, (2) nominal wage rigidity, and (3) productivity and resource allocation. Below, we survey previous studies in

these fields and discuss this paper's contributions to each area.

2.1 Labor Share and Wage Markdown

Firstly, recent analyses have reported a global declining trend in the labor share, and a substantial body of research, primarily focusing on the United States, has examined its factors and underlying causes (Grossman and Oberfield (2022); Bergholt et al. (2022); Karabarbounis (2024)). As these studies point out, several factors such as technological innovation, globalization, the decline in workers' bargaining power, and demographic dynamics have collectively exerted downward pressure on the labor share. A representative recent study, Autor et al. (2020), points out that wages have been suppressed relative to productivity in highly productive large firms ("superstar firms") in the United States. In Europe, it has been observed that the increase in labor supply due to globalization has led to a decline in workers' bargaining power (Lodge et al. (2021)).

In Japan as well, the labor share of income has shown a declining trend, and several studies have reported that the decline in workers' bargaining power has contributed to this downward pressure. For instance, Fukao and Perugini (2021) highlight that the increasing proportion of non-regular workers has weakened the bargaining power of labor unions, exerting downward pressure on the labor income share. Also, the increase in FDI has strengthened the wage bargaining power of multinational corporations (Dobbelaere and Kiyota (2018)). Additionally, some studies highlight the weakening of workers' wage bargaining power in the context of labor hoarding during economic recessions (Goodhart and Pradhan (2020)). Furthermore, particularly since the late 1990s, worsening corporate earnings have led to a shift in labor-management negotiations, prioritizing employment stability over wage increases, which has also been identified as a factor contributing to the weakening of workers' wage bargaining power (Ogura (2017); Watanabe (2024)).

The decline in workers' wage bargaining power, or the strengthening of firms' bargaining power, has been incorporated into theoretical models. When assuming that firms have stronger bargaining power compared to workers (Monopsony market), it becomes possible to model wage markdowns, where wages fall below workers' productivity (Berger et al. (2022)). In this regard, recent empirical analyses using microdata have also reported the existence of wage markdowns (Yeh et al. (2022), Aoki et al. (2023)). Aoki et al. (2023) estimate wage markdowns and price markups in Japan and point out that while price markups have been shrinking, wage markdowns have been expanding. These estimation results suggest the possibility that wage markdowns have expanded in response to intensified international competition, which has made it difficult to raise product prices, as firms strive to secure their profits.

In this paper, based on these previous studies, we focus on the decline in workers' bargaining power as a factor contributing to the downward pressure on the labor share of income. Specifically, using a theoretical model, we quantitatively analyzed the extent to which the practice of prioritizing job security over wage increases in labor-management negotiations has contributed to the gap

between real wages and labor productivity.

2.2 Downward Nominal Wage Rigidity

Secondly, downward nominal wage rigidity (DNWR), driven by concerns that reducing nominal wages could diminish workers' motivation, has been highlighted in previous studies conducted across various countries and regions (Elsby (2009)). In recent years, analyses using large-scale microdata at the worker level have been conducted. Grigsby et al. (2021) analyze individual worker wage data in the United States and point out that fluctuations in wage aggregates are largely due to changes in wage levels from job transitions and bonuses, while the basic wages of individual workers remain rigid. Furthermore, Branten et al. (2018) analyze post-GFC individual wage surveys in EU countries and highlight that nominal wages were not reduced during economic recessions. Regarding Japan, studies by Yamamoto (2007) and Kambayashi (2011), among others, identify the existence of nominal wage rigidity based on the distribution patterns of individual workers' wage growth rates. Yamamoto (2007) compares the degree of rigidity (the proportion of workers whose wages have remained unchanged) across different types of wages internationally and reported that while the rigidity of total annual earnings, including bonuses, is low in Japan, the basic salary of regular employees is more rigid compared to other countries.

When DNWR exists, wage increases may also be suppressed, leading to wage rigidity in both upward and downward directions - a phenomenon known as wage freeze. Using theoretical models that incorporate DNWR, Elsby (2009) and Daly and Hobijn (2014) highlight a mechanism in which the fear of being unable to reduce wages during future periods when wage cuts might be necessary results in firms restraining current wage increases. While empirical studies on upward nominal wage rigidity (UNWR) are limited, there are some studies that quantitatively examine UNWR for specific time periods. For instance, in Japan, Yamamoto and Kuroda (2016) and Hirata et al. (2020) define workers who did not experience wage reductions during past economic downturns as having faced DNWR, and they point out that, during subsequent periods of economic recovery, the wage growth rate for workers facing DNWR is comparatively smaller than that for workers who do not experience such rigidity.

Furthermore, regarding nominal wage rigidity, as pointed out as early as Tobin (1972), it has been noted that when inflation rates are high, real wages decline without the need to lower nominal wages, thereby weakening the impact of nominal wage rigidity (the "Grease the Wheels Effect" of inflation). In this regard, recent empirical studies have also highlighted that in periods of low inflation, wage increases are more likely to be suppressed due to increased prospects of facing future downward rigidity (Schaefer and Singleton (2023)). In terms of the relationship between DNWR and inflation, recent analyses employing theoretical models that incorporate DNWR highlight that moderate levels of inflation can lead to improvements in social welfare (Aktug (2025), Mineyama (2022)).

This paper examines nominal wage rigidity in Japan from both empirical and theoretical perspectives. In the empirical section, it extends the friction model used by studies such as Fehr and Goette (2005) and Kuroda et al. (2003), which quantitatively analyze the degree of downward rigidity, into a framework that evaluates both upward and downward rigidities. With this framework, we investigate whether wage freeze has occurred in Japan, and examines the impact of the recent rise in inflation on rigidity. In the theoretical section, based on the partial equilibrium model of the labor market with DNWR proposed by Ehrlich and Montes (2024), a general equilibrium model with heterogeneous firms of varying productivity is constructed. Under conditions of nominal wage rigidity, we quantitatively analyze the impact of inflation rates on the gap between real wages and productivity.

2.3 Allocative Efficiency

Thirdly, this paper is also related to the analysis of the impact of resource allocation efficiency on productivity. Generally, the rise in macro-level productivity can be decomposed into the contributions of technological progress and improvements in the resource allocation of micro-agents. As for methods to measure the contribution of resource allocation to macro-level productivity growth, both statistical approaches and structural model approaches have been proposed (Goldin et al. (2024)).

In the statistical approach, macro-level productivity is decomposed into the within-firm effect, which reflects changes in productivity within individual firms, and the allocation effect, which reflects changes in the shares of firms with different productivity levels (e.g., Foster et al. (2001)). Analyses based on the statistical approach focusing on Japan have point out that the reallocation effect has significantly influenced productivity trends (Ikeuchi et al. (2018); Fukao et al. (2021)).

The structural model approach is a method to measure resource allocation inefficiencies arising from imperfect competition. A seminal study by Baqaee and Farhi (2020) quantifies inefficiencies due to imperfect competition through price markups, decomposing TFP growth into contributions from technological progress and resource allocation effects. Aoki et al. (2024), in applying a similar analysis to Japan, report that compared to the United States, Japan exhibits a smaller contribution from technological progress and a larger contribution from resource allocation effects.

Based on previous studies that suggest resource allocation effects contribute to macro-level productivity, this paper also examines the mechanism through which the reduction in nominal wage rigidity improves worker allocation and impacts macro-level productivity. We confirmed that if the impact of nominal wage rigidity is alleviated by factors such as a rise in trend inflation, macro-level productivity can improve through the mobility of workers driven by wage signals.

3 Empirical Analysis

In this section, we provide empirical evidence that Japanese workers have faced both downward and upward nominal wage rigidity. In particular, we show that upward nominal wage rigidity is observed to be greater during the deflationary period in Japan, which is new to the literature. We start the empirical section with a review of the empirical literature on wage rigidity and show our contribution. Then we describe the details of the datasets and the empirical strategy. Finally, we present the results and discuss the implications of our findings.

3.1 Review on Empirical Analysis on Nominal Wage Rigidity

Since the 1990s, empirical studies leveraging worker-level microdata have proliferated across various countries, shedding light on the rigidity of nominal wages, particularly in the context of downward rigidity. The seminal works by Altonji and Devereux (1999) and Fehr and Goette (2005) utilize a friction model. This framework posits that potential wage adjustments are subject to threshold constraints. These studies document downward nominal wage rigidity (DNWR) in the United States and Switzerland. Likewise, Dickens et al. (2007) examines worker-level microdata from 16 Western European countries to analyze the distribution of wage changes, finding significant size of clustering around zero percent and concluding that DNWR is prevalent to varying extents across these nations. More recently, Grigsby et al. (2021) develops a comprehensive dataset based on U.S. administrative payroll data, revealing that nominal base wage declines among job-stayers are much rarer than previously thought.

On the other hand, empirical analyses that leverage microdata to investigate the upward nominal wage rigidity (UNWR) have been relatively less developed. Schaefer and Singleton (2023) provides evidence for upward rigidity by showing that, during periods of low price inflation in the United Kingdom, the upper tail of the wage growth distribution became compressed. Similarly, Christofides and Li (2005) analyzes worker-level microdata from Canada, utilizing an estimation framework that incorporated an upper threshold into the friction model. They reveals the presence of upward thresholds in addition to downward ones.

Turning to studies focused on Japan, Kuroda et al. (2003) analyzes data from the 1990s using a friction model and identified evidence of downward rigidity in base salaries for regular employees. Additionally, Kambayashi (2011), employing an extended dataset covering the period from 1993 to 2006, finds concentration around zero percent in the distribution of base salary changes. They find that this concentration around zero percent becoming more pronounced in the late 1990s. With respect to studies focusing UNWR in Japan, Yamamoto and Kuroda (2016) and Hirata et al. (2020) find that the wage growth rates of workers experiencing downward wage rigidity lagged behind those of workers unaffected by such constraints.[3]

[3]They define workers who do not experience wage reductions during economic downturns as facing downward

We contribute to the literature as we conduct a quantitative analysis using worker-level micro-data spanning a long period, from 1990 to the most recent data available in 2024, which include both deflationary and non-deflationary periods. With our data sets, we are able to examine whether the extent of upward and downward wage rigidity has changed over time. Furthermore, the analysis adopts an estimation framework that accounts for not only downward but also upward rigidity in nominal wages.

3.2 Data Used for the Analysis

This study utilizes survey data from the Basic Survey on Wage Structure, conducted by the Ministry of Health, Labour and Welfare, covering the period from 1989 to 2024. The dataset consists of approximately 45.2 million worker-level samples in total. This survey provides comprehensive, long-term data on individual workers' wages and working hours, alongside demographic information such as gender, age, and educational background, as well as characteristics of the establishments, including industry sector and firm size. Our focus of the data is the scheduled wages of full-time workers.

The survey employed in this study is a repeated cross-sectional survey, not a panel survey. It draws new samples each year and lacks recorded identifiers to track individual workers across multiple survey periods. Therefore, calculating worker-level wage growth rates using this survey requires the identification of the same workers across survey years. Following the methodology outlined by Kambayashi (2011) and Date et al. (2024), this study matches samples that are assumed to correspond to the same individuals based on workplace, gender, educational background, age, and tenure. This matching process enables us to construct a pseudo-panel dataset.

In detail, the same workplace was identified by matching establishments with identical prefectures, municipalities, basic survey areas, and establishment numbers between the previous year and the current year. It is important to note that in the years when the sampling frame of the *Basic Survey on Wage Structure* was updated, the continuity of these identifiers was not always preserved. However, the linking of identifiers was facilitated by using corresponding tables between the old and new surveys, obtained from sources such as the *Economic Census* provided by the Ministry of Internal Affairs and Communications. Nonetheless, for the periods 1992-1993, 1995-1996, and 1997-1998, corresponding tables for the old and new surveys were unavailable. Therefore, in this study, connections for these years could not be established, and the wage growth rates for the years 1993, 1996, and 1998 are excluded from the dataset.

Subsequently, among the worker samples extracted from identified matching establishments, those with matching gender and educational background between the previous year and the current year, and whose age and tenure had both increased by one year from the previous year, were identified and linked as the same worker. To minimize the likelihood of erroneously linking samples of different individuals, observations were excluded if multiple workers within the same establish-

rigidity.

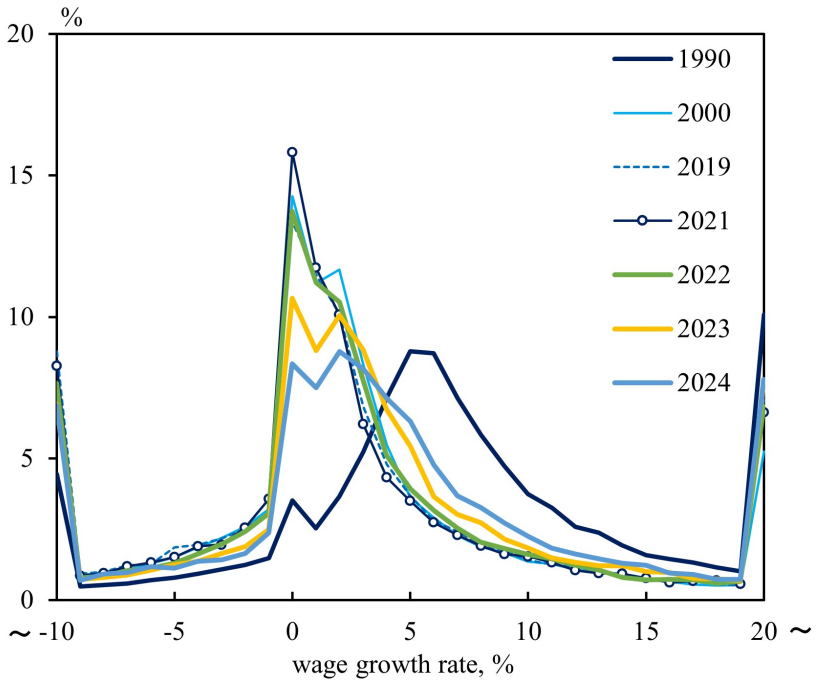


Figure 2: Distributions of nominal wage growth over last three decades

ment in either the previous or current year shared identical gender, educational background, age, and tenure. Furthermore, workers were excluded from the analysis if their wage growth rates or changes in working hours fell below -35% or exceeded +60%. Consequently, wage growth rates were calculated for a total of approximately 3.6 million samples, excluding the survey years 1993, 1996, and 1998. These calculated wage growth rates are illustrated in Figure 2.

Examining the distribution of wage growth rates in Figure 2, it is observed that since the 2000s, when price inflation stagnated, the distribution of wage growth rates shifted leftward and became concentrated around zero percent. This suggests the presence of workers facing nominal wage rigidity. Subsequently, throughout the 2010s, the concentration of wage growth rates around zero percent persisted without significant changes. However, from 2020 to the most recent data of 2024, while the overall wage growth rate distribution shifted to the right, the proportion around zero percent decreased, indicating a potential change in the extend of nominal wage rigidity. In the following sections, the degree of nominal wage rigidity and its potential changes will be analyzed using quantitative methods.

3.3 Empirical Analysis Framework

In this section, a quantitative analysis will be conducted using the constructed dataset of wage growth rates to examine the rigidity of wage growth rates, as well as potential changes in these rigidities over time.

Specifically, drawing on studies such as Christofides and Li (2005), this paper conducts an es-

timation using a friction model with upper and lower thresholds. The friction model assumes the existence of a potential wage growth rate, which is determined by the attributes of workers and firms. However, the observed wage growth rate is constrained by both upper and lower thresholds, as defined by eq. (1) and eq. (2).

$$\Delta w_i^* = x_i'\beta + \varepsilon_i, \quad \varepsilon_i \sim N\left(0, \tau^2\right) \tag{1}$$

$$\Delta w_i = \begin{cases} \Delta w_i^* - \theta_t^+ & \text{if} \quad \theta_t^+ \leq \Delta w_i^* \\ 0 & \text{if} \quad \theta_t^- < \Delta w_i^* < \theta_t^+ \\ \Delta w_i^* - \theta_t^- & \text{if} \quad \Delta w_i^* \leq \theta_t^- \end{cases} \tag{2}$$

In eq. (1), the dependent variable Δw_i^* represents the potential wage growth rate, and i denotes the worker. Regarding the explanatory variables x_i, they consist of data on corporate profitability (year-over-year changes in ordinary profits for the industry and firm size to which worker i belongs), inflation rate (prefecture-level consumer price index excluding fresh food), labor market conditions (Employment Condition DI by industry and company size), age, a dummy variable for workers aged 60 and over, working hours, a gender dummy, an education dummy, an external labor market dummy, and dummy variables indicating industry and firm size.[4]

Eq. (2) illustrates the relationship between the potential wage growth rate Δw_i^* and the observed wage growth rate Δw_i. Specifically, it is assumed that when the potential wage growth rate falls within the threshold range (θ_t^-, θ_t^+), the observed wage growth rate is zero, meaning wages remain unchanged. On the other hand, if the potential wage growth rate exceeds the thresholds, only the portion exceeding the specified thresholds is observed as the actual wage growth rate. Figure 3 illustrates the relationship between the two. Moreover, the upper threshold θ_t^+ and lower threshold θ_t^- were estimated while allowing them to vary across different periods: (i) 1990-1994, (ii) 1995-2003, (iii) 2004-2012, (iv) 2013-2019, and (v) 2020-2024. The estimation method is based on the Bayesian estimation approach for Tobit models proposed by Chib (1992), and the estimation technique proposed by Omori and Miyawaki (2010) was adopted.

3.4 Results of the Empirical Analysis

First, the determinants of the potential wage growth rate, as estimated from eq. (1), were examined in Table 1. Among the macroeconomic factors, it was observed that improvements in corporate profitability, higher inflation rates, and tighter labor market conditions significantly contributed to increases in the potential wage growth rate. With regard to worker attributes, the findings indicate that younger age, longer scheduled working hours, and being female were associated with signif-

[4]Using a finite mixture model, as in Date et al. (2024), full-time workers are categorized into two classes with different wage structures. Date et al. (2024) confirm that those two classes correspond to external labor market class, characterized by a relatively flat wage curve as indicated in previous studies, and internal labor market class.

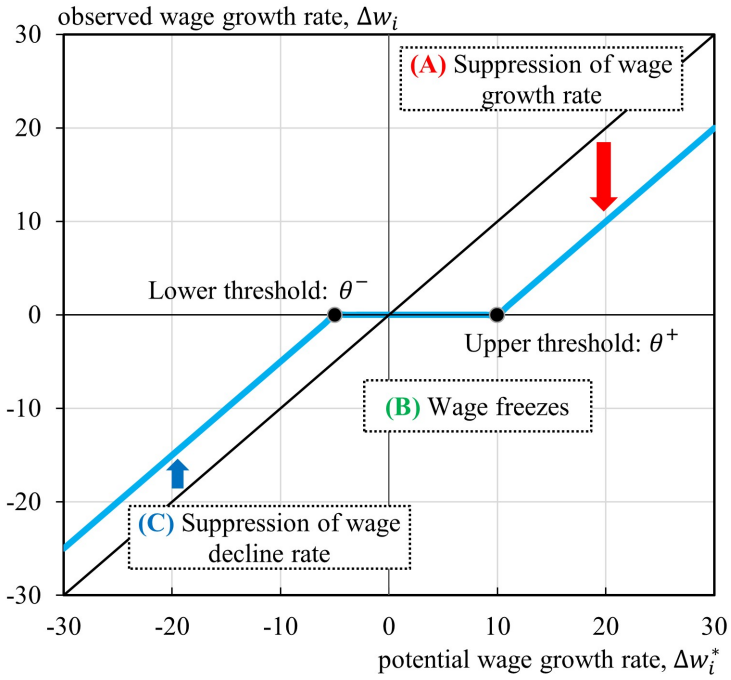


Figure 3: The conceptual diagram illustrating observed and potential wage growth

Table 1: Estimated factors affecting the potential wage

Variables	Medians	95% Confidence Intervals
Ordinary profit ratio (y/y chg., by firm industry and size)	0.09	(0.08, 0.09)
CPI inflation rate (by prefecture, excluding fresh foods)	0.64	(0.62, 0.65)
Employment conditions DI (by firm industry and size)	-0.03	(-0.03, -0.03)
Age	-0.11	(-0.12, -0.11)
Aged 60 or above	-0.58	(-0.64, -0.53)
Scheduled working hours (y/y growth rate)	0.13	(0.13, 0.13)
Female	0.17	(0.14, 0.20)
Senior High school graduates	0.00	(-0.04, 0.04)
Junior college etc. graduates	0.00	(-0.06, 0.05)
University etc. graduates	0.21	(0.16, 0.26)
External labor market	-1.24	(-1.28, -1.21)
Constant	4.88	(4.78, 4.97)

icant increases in wage growth rates. Regarding educational background, the results suggest that holding a university or graduate degree positively influenced wage growth rates. Furthermore, it was confirmed that workers affiliated with the external labor market experienced significant downward pressure on their wage growth rates.

In this study, it is assumed that the mechanism governing potential wage determination remains stable over the estimation period. However, the observed wage growth rate is subject to

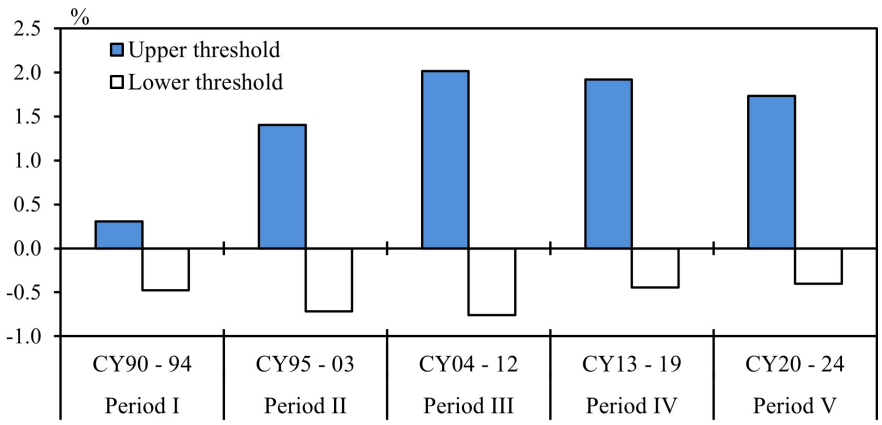


Figure 4: Estimation results of upper and lower thresholds

NOTE: The figures are medians of the posterior distributions.

upward and downward thresholds, which vary across different periods. As shown in Figure 4, the estimated thresholds for each period depict different trends for the upper and lower bounds. The lower threshold remained relatively steady, experiencing only minor fluctuations across all periods. In contrast, the upper threshold exhibited a noticeable upward trend from the early 1990s (Period I) to the early 2000s (Period III), followed by a modest decline as the economy transitioned into the early 2010s (Period IV) and subsequently into the early 2020s (Period V).

The proportions of workers facing upward and downward rigidity were depicted in Figure 5. Here, based on the estimated thresholds, workers experiencing upward rigidity were defined as those whose potential wage growth rates were positive but whose wages did not increase, while workers experiencing downward rigidity were defined as those whose potential wage growth rates were negative but whose wages remained unchanged. In the early 1990s (Period I), the proportion of workers encountering either form of wage rigidity was less than 1%, indicating that potential wage growth rates were largely reflected in actual wage adjustments during this period. However, the proportion of workers experiencing rigidity increased steadily, peaking in the early 2000s (Period III) at 7% for upward rigidity and 4% for downward rigidity.

3.5 Comparison Between Observed Data and Estimated Potential Wages

In this section, the potential wage growth rate for each worker is calculated based on the estimation results derived from the aforementioned framework. Subsequently, the distribution of potential wage growth rates is constructed and compared to the distribution of observed wage growth rates, as illustrated in Figure 6.

First, in 1990, the influence of both the upper and lower estimated thresholds was minimal, resulting in the distributions of potential and observed wage growth rates being approximately similar. Next, focusing on 2005 and 2015 – years within the deflationary period – the following key

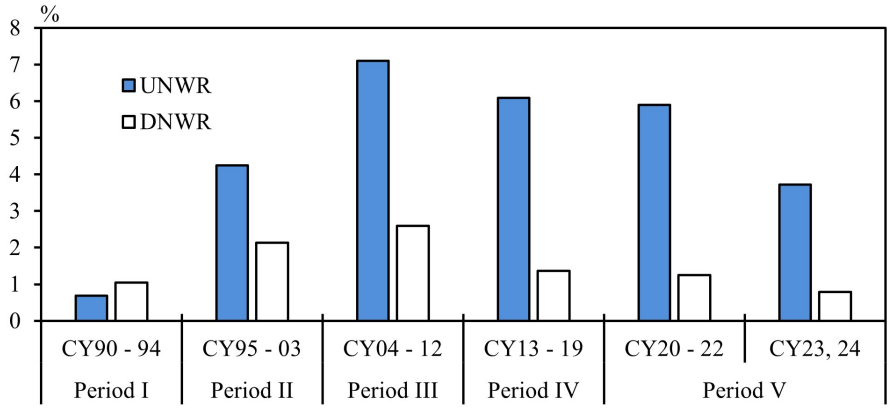


Figure 5: Proportion of workers facing upward and downward rigidity

NOTE: Workers facing UNWR (DNWR) are defined as those whose potential wage growth rate is positive (negative), but whose observed wages remain unchanged.

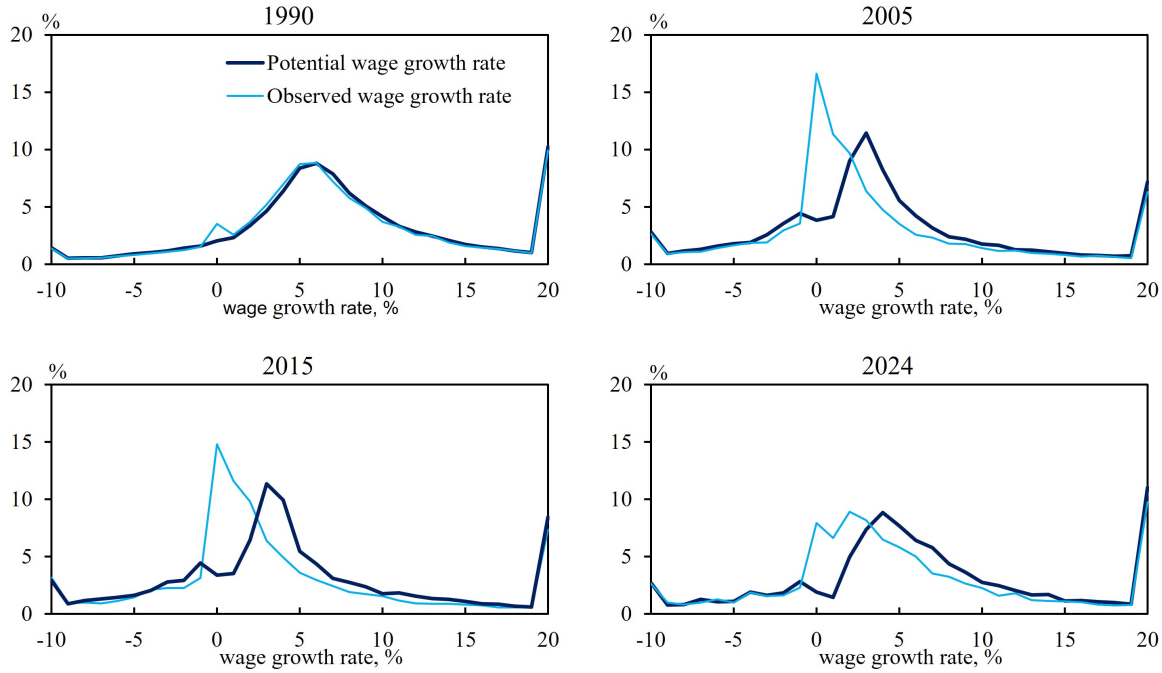


Figure 6: Comparison of potential and observed wage growth distributions

trends were identified: (i) an expansion of the upper threshold, and (ii) a decline in potential wage growth rates, driven by factors such as decreasing inflation rates. Consequently, the distribution of observed wage growth rates shifted leftward relative to the distribution of potential wage growth rates, forming a pronounced peak near zero percent.

For 2024, while the thresholds themselves have remained relatively unchanged, rising inflation rates have led to an increase in potential wage growth rates. As a result, the proportion of workers experiencing upward wage rigidity has declined, narrowing the gap between potential wage growth rates and observed wage growth rates. This trend indicates a mitigated impact of nominal wage

rigidity.

Based on the results of the empirical analyses, the following key findings can be indicated. First, the analysis quantitatively confirms that a significant proportion of workers in Japan have experienced nominal wage rigidity, both upward and downward. Second, the influence of upward wage rigidity intensified notably after the late 1990s, coinciding with the onset of deflation in the Japanese economy. During this period, observed wage growth rates consistently fell below potential wage growth rates. Third, in the most recent period since 2020, the impact of wage rigidity appears to have eased. This mitigation is likely attributable to rising inflation rates and other contributing factors, which have collectively driven an increase in potential wage growth rates.

4 Theoretical Framework

4.1 Model

In theoretical analysis, to investigate the factors behind the gap between productivity and wages, we construct a framework that incorporates not only the empirically observed downward nominal wage rigidity (DNWR) but also the trade-off between wage increases and job security in labor-management negotiations. When examining macro-level productivity, the effects of resource allocation among firms also become crucial. Thus, we extend the traditional New Keynesian model by introducing firms with heterogeneous productivity levels following Ehrlich and Montes (2024).

The model features two distinct margins of employment adjustment: retention of incumbent workers through wage-setting, and hiring of new workers through vacancy posting. Downward nominal wage rigidity makes wage-setting history-dependent and shifts the relative cost of using these margins. These mechanisms jointly determine the effective marginal cost of labor and the endogenous wage markdown, which are the focus of the subsequent analysis.

The economy consists of heterogeneous firms that produce intermediate goods and face idiosyncratic productivity shocks, wholesale firms that supply these intermediate goods monopolistically, and firms that combine the intermediate goods to produce final goods. Workers supply labor to firms producing intermediate goods and consume final goods. The remaining agents are the government and the central bank. The government conducts lump-sum taxation or transfers, while the central bank adjusts the nominal interest rate in response to the inflation rate. Time is discrete and indexed by $t = 0, 1, 2, \ldots$, and the unit of the time horizon for the model is annual.

4.1.1 Heterogeneous intermediate-goods producing firms

Intermediate goods-producing firms hire workers to conduct production. The only input is labor, and it is assumed that each worker supplies a fixed unit of labor. Thus, neither working hours nor capital is considered in the analysis.

There is a unit measure of intermediate-goods producing firms indexed by $j \in [0, 1]$. Produc-

tivity differs across firms, and the index j represents an individual firm. Given aggregate prices and quantities $\{P_t^w, W_t, \pi_t, u_t\}$, firm j chooses its current wage $w_{j,t}$ and vacancies $v_{j,t}$ to maximize the present discounted value of profits. Firms sell their output at a relative price P_t^w, which corresponds to the real price of intermediate goods faced by wholesale producers. The Bellman equation for firm j is as follows:

$$\mathbb{V}_{j,t}\left(a_{j,t}, w_{j,t-1}, n_{j,t-1}\right) = \max_{w_{j,t}, v_{j,t}} \Pi_{j,t} + \beta \mathbb{E}_t\left[\mathbb{V}_{j,t+1}\left(a_{j,t+1}, w_{j,t}, n_{j,t}\right)\right]. \tag{3}$$

The firm's state variables are current productivity $a_{j,t}$, last period's wage $w_{j,t-1}$, and last period's employment $n_{j,t-1}$.

The firm's period-t profit is given by:

$$\Pi_{j,t} = P_t^w y_{j,t} - w_{j,t} n_{j,t} - \underbrace{\mathcal{D}\left(w_{j,t}, w_{j,t-1}\right) n_{j,t}}_{\text{DNWR}} - \underbrace{\mathcal{C}\left(v_{j,t}, n_{j,t-1}\right)}_{\text{Vacancy-posting cost}}. \tag{4}$$

Revenue equals output multiplied by the relative price P_t^w. Costs consist of the wage bill, downward nominal wage adjustment costs that are proportional to current employment, and vacancy posting costs that depend on current vacancies and past employment.

Production function Firm j produces output using labor as the sole input according to

$$y_{j,t} = a_{j,t} n_{j,t}^{\alpha}, \tag{5}$$

where $y_{j,t}$ denotes firm-level output, $n_{j,t}$ denotes employment, and $\alpha \in (0,1)$ implies decreasing returns to scale. This assumption ensures an interior firm size distribution and prevents unbounded expansion of highly productive firms.

Firm-level productivity $a_{j,t}$ consists of an aggregate component and an idiosyncratic component. Specifically, the logarithm of productivity satisfies

$$\begin{aligned} \ln a_{j,t} &= \ln z_t + \ln u_{j,t}, \\ \ln z_t &= \gamma t + \rho^z \ln z_{t-1} + \varepsilon_t^z, \quad \varepsilon_t^z \sim N(0, \sigma_z^2), \\ \ln u_{j,t} &= (1-\rho^u)\ln \bar{u} + \rho^u \ln u_{j,t-1} + \varepsilon_{j,t}^u, \quad \varepsilon_{j,t}^u \sim N(0, \sigma_u^2). \end{aligned} \tag{6}$$

Here, z_t captures aggregate productivity common to all firms, while $u_{j,t}$ captures firm-specific productivity. The parameters ρ^z and ρ^u govern persistence, γ is the drift term of aggregate productivity, and ε_t^z and $\varepsilon_{j,t}^u$ are mean-zero innovations.

Downward nominal wage rigidity Following Ehrlich and Montes (2024), DNWR is introduced through an asymmetric adjustment cost:

$$\begin{aligned}\mathcal{D}\left(w_{j,t}, w_{j,t-1}\right) &= \left[\lambda'\left(\frac{w^n_{j,t-1}}{w^n_{j,t}} - 1\right)\right] \mathbb{1}_{\frac{w^n_{j,t-1}}{w^n_{j,t}} > 1} \\ &= \left[\lambda\left(\frac{w_{j,t-1}}{w_{j,t}} - (1+\pi_t)\right)\right] \mathbb{1}_{(1+\pi_t)w_{j,t} < w_{j,t-1}}.\end{aligned} \tag{7}$$

where π_t denotes the inflation rate and $\lambda \geq 0$ governs the severity of DNWR. The indicator function equals one whenever the firm attempts to cut the nominal wage. In such cases, the firm incurs an additional cost proportional to the size of the wage cut. Higher inflation allows real wages to decline mechanically even if nominal wages are unchanged, thereby relaxing the effective rigidity.

Vacancy-posting cost Following Bilal et al. (2022) and Bloesch et al. (2024), we assume:

$$\mathcal{C}(v_{j,t}, n_{j,t-1}) = c^v \left(\frac{v_{j,t}}{n_{j,t-1}}\right)^{\chi} v_{j,t}. \tag{8}$$

where $c^v > 0$ scales the level of hiring costs. $\chi \geq 0$ determines the degree of convexity of vacancy-posting costs with respect to the number of vacancies. Hence, per-vacancy costs increase with the ratio of vacancies to incumbent workers, $v_{j,t}/n_{j,t-1}$. This formulation implies that hiring is more costly, in marginal terms, for firms that expand rapidly relative to their existing workforce.

Firms' employment The law of motion for firm j's employment is given by:

$$n_{j,t} = \left[1 - \delta\left(\frac{w_{j,t}}{W_t}\right)\right] n_{j,t-1} + m_{j,t}. \tag{9}$$

The first term represents workers retained from the previous period, while the second term represents newly hired workers. The separation rate $\delta(\cdot)$ depends on the firm's wage relative to the aggregate wage index W_t, so that workers are more likely to leave firms that pay relatively low wages.

The separation rate is defined as:

$$\delta\left(\frac{w_{j,t}}{W_t}\right) = \bar{\delta}\left(\frac{w_{j,t}}{W_t}\right)^{-\epsilon^\delta}, \tag{10}$$

where $\bar{\delta}$ denotes the average separation rate, and ϵ^δ measures workers' preference for job security—higher values make separations more wage-sensitive. A higher value of ϵ^δ implies a more

elastic labor supply to the firm and weaker monopsony power.

Matching function in the labor market New matches are formed through a matching technology

$$m_{j,t} = v_{j,t}^{1-\mu} u_{t-1}^{\mu}, \tag{11}$$

where $v_{j,t}$ denotes the number of vacancies posted by firm j and u_{t-1} denotes aggregate unemployment. The parameter $\mu \in (0,1)$ controls the elasticity of matches with respect to unemployment.

4.1.2 Workers

The economy is populated by a representative household (worker) consisting of a continuum of members of measure one. In each period, a member is either employed or unemployed. An employed member supplies one unit of labor inelastically, while an unemployed member supplies no labor. The household provides perfect consumption insurance across its members, so that consumption is identical for all individuals, regardless of their employment status.

Preferences The preferences of the representative household are defined as the equally weighted average of its members' utilities (CRRA):

$$\mathbb{U}_t = \mathbb{E}_t \sum_{s=0}^{\infty} \beta^s u(C_{t+s}). \tag{12}$$

where C_t denotes aggregate consumption, $\beta \in (0,1)$ is the discount factor.

Budget constraint The household can save in risk-free bonds and receives profits from firms, Ψ_t:

$$P_t C_t + \frac{B_t}{R_t} = \int_0^1 P_t w_{j,t} n_{j,t} dj + B_{t-1} + \Psi_t. \tag{13}$$

Here, P_t is the aggregate price level of the final good, B_t denotes nominal holdings of a one-period risk-free bond chosen at time t, and R_t is the gross nominal interest rate. Labor income is obtained from firms, where $w_{j,t}$ denotes the real wage paid by firm j per worker and $n_{j,t}$ denotes firm j's employment level. The term Ψ_t represents aggregate firm profits rebated lump-sum to the household.

4.1.3 Price Setting; final goods producing firms and wholesale firms

Final-good prices are set by monopolistically competitive wholesale firms subject to Rotemberg adjustment costs. Monopolistic wholesale goods firm i chooses its price $p_t(i)$ to maximize the sequence of discounted future profits:

$$\mathbb{E}_t \left[\sum_{s=t}^{\infty} q_s \left(\frac{p_s(i)}{P_s} y_s(i) - P_s^w y_s(i) - AC_s(i) \right) \right], \tag{14}$$

where q_s is the real stochastic discount factor of households. The adjustment cost is:

$$AC_t(i) = \frac{\varphi}{2}\left(\frac{p_t(i)}{p_{t-1}(i)} - \Pi\right)^2 \frac{p_t(i)}{P_t} y_t(i). \tag{15}$$

where $\varphi > 0$ governs price rigidity, Π denotes steady-state gross inflation. The final good is aggregated according to a CES technology:

$$Y_t = \left(\int_0^1 y_t(i)^{\frac{\epsilon^p - 1}{\epsilon^p}} di\right)^{\frac{\epsilon^p}{\epsilon^p - 1}}. \tag{16}$$

where $\epsilon^p > 1$ is the elasticity of substitution across varieties.

The New Keynesian Phillips Curve (NKPC) in this economy is given by:

$$\pi_t = \beta \mathbb{E}_t\left[\pi_{t+1}\right] + \frac{\epsilon^p - 1}{\varphi \Pi^2} p_t^w. \tag{17}$$

where p_t^w denotes real marginal cost.

4.1.4 Central bank

Monetary policy follows a Taylor rule:

$$r_t = \rho^{mp} r_{t-1} + (1 - \rho^{mp})\phi_\pi \pi_t. \tag{18}$$

where r_t is the policy rate, $\rho_{mp} \in [0, 1)$ captures interest-rate smoothing, and $\phi_\pi > 0$ governs the response to inflation.

4.1.5 Equilibrium definition

Equilibrium System *Given a sequence of productivity shocks $\{\epsilon_t^z\}$ and an initial firm distribution $\Phi_0(a, w, n)$, a competitive equilibrium is a sequence of policies $\{w(a, w_-, n_-), n(a, w_-, n_-)\}$ for intermediate-goods-producing firms, distributions $\Phi_t(a, w_-, n_-)$, prices $\{P_t, P_t^W, W_t, R_t\}$, and aggregate quantities $\{Y_t, N_t, C_t\}$ such that all workers optimize, all firms optimize, and the goods market clears.*

4.2 DNWR yields UNWR, and leads to Wage Freeze

Figure 7 illustrates the policy function for wage setting in the current period for heterogeneous firms. The case where DNWR exists ($\lambda > 0$) is represented by the blue surface, while the case without DNWR ($\lambda = 0$) is shown by the yellow surface. Comparing the left side of the figure (firms with low wages in the previous period), it can be observed that wage levels are lower when DNWR exists

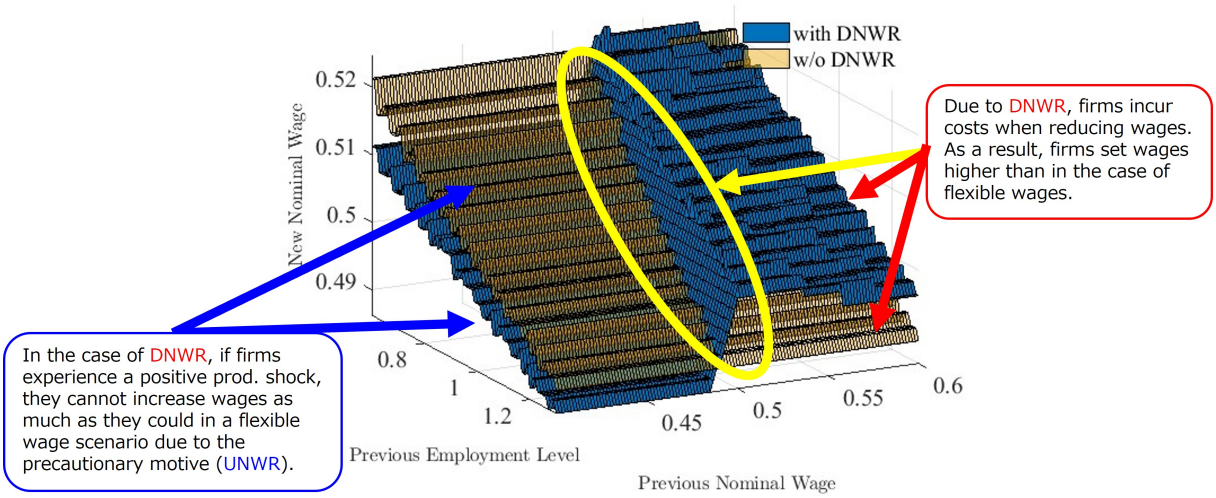


Figure 7: Firms' policy functions for wages in the next period

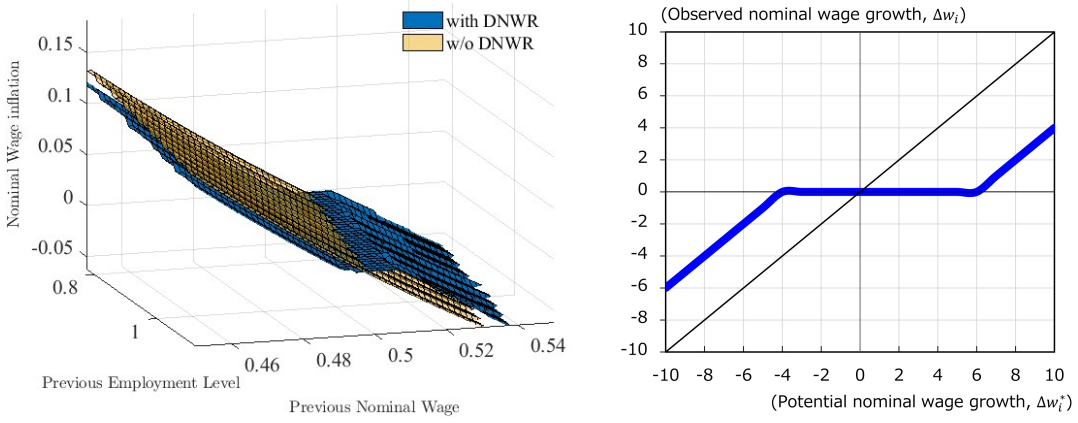


Figure 8: Potential and observed nominal wage growth

than when it does not. This indicates that, in the presence of DNWR, firms have an incentive not to raise wages in the current period due to the high costs associated with reducing wages in the future. In other words, the existence of DNWR indicates the emergence of Upward Nominal Wage Rigidity, making it more difficult to raise wages in the current period.

Conversely, comparing the right side of the figure (firms with high wages in the previous period), it can be observed that wage levels are higher when DNWR exists than when it does not. This is because, in the presence of DNWR, there are costs associated with reducing the previous period's wages in the current period. As a result, wages are set higher compared to the case where DNWR is absent, allowing for more flexible wage reductions.

Figure 8 illustrates the policy functions for wage growth in cases where DNWR (Downward Nominal Wage Rigidity) exists and does not exist. When DNWR is present, it can be observed that some firms face a *Wage Freeze*, where they are unable to either raise or lower wages.

Table 2: *Calibrated* parameters

Parameter		Value	Reason
Heterogeneous firms			
Discount rate	β	0.99	$\beta = (1+\gamma) / (1+\bar{r})$
Returns to scale in production	α	0.70	Ehrlich and Montes (2024)
Hiring cost constant	c^v	0.25	Match unemployment rate in Japan
Posting vacancy paramter	χ	3.45	Bilal et al. (2022)
Average separation rate	$\bar{\delta}$	0.15	Separation rate in Japan excl. retirement
Persistence of individual prod. shock	ρ_u	0.388	Ehrlich and Montes (2024)
S.D. of individual prod. shock	σ_u	0.419	Ehrlich and Montes (2024)
Elasticity of matches w.r.t. job search	μ	0.5	Petrongolo and Pissarides (2001)
Workers, Wholesale firms, and Final goods producing firms			
Relative risk aversion	σ	1	Christiano et al. (2005)
Rotemberg adjustment cost	φ	10	≈ 13.85 in Fueki et al. (2018)
Retail price elasticity	ϵ^p	6.0	Aoki et al. (2023), $\epsilon^p / (\epsilon^p - 1) = 1.2$
Monetary policy			
Taylor rule coefficient	ϕ^π	2.0	standard
Persistence of monetary shock	ρ^{mp}	0.65	0.9^4
Economy wide			
Growth trend	γ	0.01	Annual growth rate in Japan

Furthermore, if we consider the case with DNWR as the observed wages and the case without DNWR as the potential wages, a relationship between the two can be identified, as shown on the right in Figure 8. In other words, the theoretically derived policy functions for wage growth are consistent with the assumptions made in the empirical analysis.

4.3 Calibration

This section describes the calibration of the model parameters. The parameters can be grouped into two broad categories: (i) parameters whose values are taken from the literature or empirical data, and (ii) key parameters that are calibrated to align the model with the data. The latter govern labor market frictions, particularly those related to separations and wage-setting behaviors, including mechanisms governing DNWR.

4.3.1 Parameter calibration from the literature

Table 2 summarizes the baseline parameter values and their empirical or theoretical justifications.

Preferences and production The discount factor β is chosen to be consistent with a growth rate and an annual real interest rate. Production depends solely on labor and features decreasing returns to scale. The assumption of decreasing returns ensures an interior labor demand solution and stabilizes the firm's dynamic wage and employment problem, as in Ehrlich and Montes (2024).

Labor market The hiring cost constant, c^v, is calibrated so that the steady-state unemployment rate aligns with the actual unemployment rate observed in Japan. The degree of convexity of vacancy-posting cost, χ, adopts the estimated value from Bilal et al. (2022). The average separation rate is set to be consistent with the corresponding data for Japan. The elasticity of matching is set at the standard value of 0.5 as commonly used in the literature (Petrongolo and Pissarides (2001)).

4.3.2 Calibration on λ and ϵ^δ

Under the baseline calibration, the model jointly matches key labor market moments, including the distributions of wage changes, separation rates, and hiring rates. In particular, the parameters governing DNWR, λ, play a central role in shaping the composition of employment adjustments across wages and separations. The quantitative exercises that follow use this calibrated model to evaluate how changes in wage rigidity affect firm-level employment dynamics and aggregate labor market outcomes.

Downward nominal wage rigidity, λ DNWR is introduced through an asymmetric wage adjustment cost. Firms incur a cost when cutting nominal wages, consisting of a fixed component and a component that is proportional to the size of the nominal wage cut. These parameters are calibrated using the distribution of wage changes.

Specifically, we simulate wage changes from the model and apply the same kernel-density-based procedure used in the data to measure the fraction of nominal wage cuts that are prevented. The DNWR parameters are chosen such that the model-implied measure of prevented wage cuts matches its empirical counterpart. This indirect calibration strategy ensures that the model reproduces the observed asymmetry in the wage change distribution.

Separation elasticity, ϵ^δ Worker separations arise endogenously through a wage-dependent separation rate. The scale parameter of the separation function is calibrated such that the steady-state separation rate matches the average annual separation rate observed in the data. The elasticity of separations with respect to wages, ϵ^δ, governs the degree of monopsony power in the labor market. A higher value of ϵ^δ implies that worker flows are more sensitive to wages and that firms face a more competitive labor market.

Calibration strategy Choose λ and ϵ^δ to match the distribution of nominal wage growth between the data and the model.

We perform a grid search over 31 grid points for parameter λ (ranging from 0.0 to 3.0) and 40 grid points for parameter $\epsilon^{\delta,T}$ (ranging from 0.1 to 4.0), and select the parameter set that minimizes:

$$\min_{\lambda,\{\epsilon^{\delta,T}\}} \left[\frac{\Sigma_T \left\| \widehat{\mathcal{F}(0 < x < 0.01)} - \mathcal{F}(0 < x < 0.01) \right\|}{\text{number of T}} + \Sigma_T \left\| \widehat{\mathcal{F}(x > 0)} - \mathcal{F}(x > 0) \right\| \right] \tag{19}$$

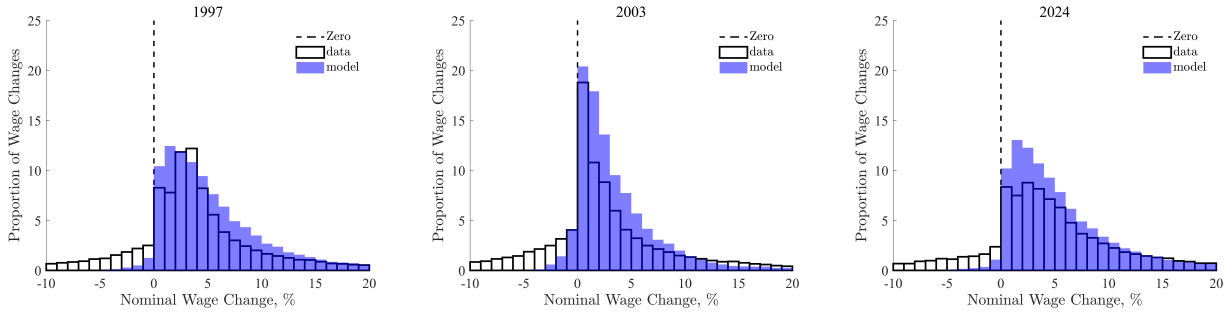


Figure 9: Comparison of nominal wage growth distributions between observed data and the model

Table 3: *Calibrated* parameters

Parameter		Value
Degree of DNWR	λ	0.9
Separation elasticity in 1997	$\epsilon^{\delta,1997}$	1.6
Separation elasticity in 2003	$\epsilon^{\delta,2003}$	1.4
Separation elasticity in 2024	$\epsilon^{\delta,2024}$	1.4

where $\mathcal{F}(\cdot)$ denotes the CDF of nominal wage growth.

Calibration results Figure 9 presents comparisons between the distributions of nominal wage growth as predicted by the model and observed in the data. Table 3 shows the calibrated parameters for λ and ϵ^{δ} for the selected time periods. Our calibrated values for ϵ^{δ} are consistent with the estimates reported by Ehrlich and Montes (2024), which are based on micro-level evidence in Germany regarding the wage-separation elasticity. Moreover, we observe that the model-implied wage markdown gap has widened over time, which aligns with the findings of Aoki et al. (2023).

5 Wage Markdown and Allocative Efficiency

This section characterizes how the steady-state inflation rate and the elasticity of separations with respect to wages affect wage markdowns and aggregate productivity.

5.1 Optimality Conditions for Heterogeneous Firms

5.1.1 Hiring FOC

As a prerequisite, remember that firms produce according to $y_{j,t} = a_{j,t} n_{j,t}^{\alpha}, \alpha \in (0,1)$, and earn revenues $P_t^w y_{j,t}$. The marginal revenue product of labor (MRPL) is therefore:

$$MRPL_{j,t} \equiv \partial_{n_{j,t}}(P_t^w y_{j,t}) = P_t^w MPL_{j,t}. \tag{20}$$

We assume the intermediate-good producing firms have an interior solution.[5] Remember $v_{j,t}$ denote vacancies and $m_{j,t}$ matches. The firm's hiring FOC at an interior optimum is:

$$\partial_v \mathcal{C}(v_{j,t}, n_{j,t-1}) = \big(\partial_n \Pi_{j,t} + \mathcal{M}_{j,t}\big)\partial_v m_{j,t}, \tag{21}$$

where we define the marginal value of employment and the continuation value of employment as:

$$\begin{aligned} \partial_n \Pi_{j,t} &= MRPL_{j,t} - w_{j,t} - \mathcal{D}_{j,t}, \\ \mathcal{M}_{j,t} &\equiv \beta \mathbb{E}_t[\partial_n \mathbb{V}_{j,t+1}]. \end{aligned} \tag{22}$$

Now, define the marginal cost per hire (adding one worker through vacancy posting) as:[6]

$$mc^H_{j,t} \equiv \frac{\partial_v \mathcal{C}(v_{j,t}, n_{j,t-1})}{\partial_v m_{j,t}} = \frac{c^v(\chi+1)}{1-\mu} \frac{v^{\chi+\mu}_{j,t}}{n^{\chi}_{j,t-1} u^{\mu}_{t-1}}. \tag{23}$$

Firms that operate on the hiring margin optimally adjust employment by posting vacancies rather than by retaining incumbent workers through wage increases. For these firms, the marginal cost of expanding employment is given by $mc^H_{j,t}$ in eq. (23). This condition implies that the marginal cost of hiring is increasing in vacancies $v_{j,t}$ and decreasing in firm size $n_{j,t-1}$ and labor market slack u_{t-1}. As a result, firms that are small, expanding rapidly, or operating in tight labor markets face particularly high hiring costs and are more likely to rely on the retention margin, which we describe in the next, instead. Conversely, firms with a large existing workforce or operating in slack labor markets can expand employment relatively cheaply through vacancy posting and thus optimally locate on the hiring margin.

Importantly, for firms on the hiring margin, wages do not directly enter the marginal cost of labor. Instead, wages affect employment only indirectly by determining whether the firm remains on the hiring margin or switches to the retention margin. Consequently, wage changes for these firms have limited immediate effects on employment, and wage markdowns are primarily pinned down by hiring frictions rather than by separation behavior.

To relate the marginal cost per hire with the MRPL, we convert eq. (21) and obtain:

$$mc^H_{j,t} = \partial_n \Pi_{j,t} + \mathcal{M}_{j,t}. \tag{24}$$

[5] $\partial_x y$ denotes $\frac{\partial y}{\partial x}$ in the followings.

[6] The marginal vacancy cost and marginal matches are

$$\frac{\partial \mathcal{C}_{j,t}}{\partial v_{j,t}} = c^v(\chi+1) n^{-\chi}_{j,t-1} v^{\chi}_{j,t} \quad \text{and} \quad \frac{\partial m_{j,t}}{\partial v_{j,t}} = (1-\mu) v^{-\mu}_{j,t} u^{\mu}_{t-1}.$$

Combining eq. (22) and eq. (24) yields the identity for the MRPL:

$$MRPL_{j,t} = w_{j,t} + mc^{H}_{j,t} + \mathcal{D}_{j,t} - \mathcal{M}_{j,t}. \tag{25}$$

We use this expression when we derive the expression for the wage markdown.

5.1.2 Wage FOC

We assume the intermediate-good producing firms have an interior solution as in the case of hiring FOC. The wage FOC at an interior optimum is:

$$-n_{j,t}\left(1+\partial_w \mathcal{D}_{j,t}\right) + \left(\partial_n \Pi_{j,t} + \mathcal{M}_{j,t}\right)\partial_w n_{j,t} + \beta \mathbb{E}_t[\partial_w \mathbb{V}_{j,t+1}] = 0. \tag{26}$$

Dividing the both sides by $\partial_w n_{j,t}$, we obtain:

$$-\frac{n_{j,t}}{\partial_w n_{j,t}} + \left(\partial_n \Pi_{j,t} + \mathcal{M}_{j,t}\right) - \frac{n_{j,t}\partial_w \mathcal{D}_{j,t} - \beta \mathbb{E}_t[\partial_w \mathbb{V}_{j,t+1}]}{\partial_w n_{j,t}} = 0. \tag{27}$$

Now we define the cost of adjusting wage as:

$$\psi_{j,t} \equiv \frac{n_{j,t}\partial_w \mathcal{D}_{j,t} - \beta \mathbb{E}_t[\partial_w \mathbb{V}_{j,t+1}]}{\partial_w n_{j,t}}. \tag{28}$$

In terms of $\partial_w n_{j,t}$, holding W_t fixed, the effect of the wage on contemporaneous employment (via retention) is

$$\frac{\partial n_{j,t}}{\partial w_{j,t}} = -\frac{\partial \delta(w_{j,t}/W_t)}{\partial w_{j,t}}\, n_{j,t-1} = \epsilon^{\delta}\delta\!\left(\tfrac{w_{j,t}}{W_t}\right)\frac{n_{j,t-1}}{w_{j,t}}. \tag{29}$$

Eq. (29) is the core "retention channel": raising wages reduces separations and increases employment by retaining incumbents. Then, we have:

$$\partial_n \Pi_{j,t} + \mathcal{M}_{j,t} = \frac{w_{j,t}}{\epsilon^{\delta}\,\delta\left(\frac{w_{j,t}}{W_t}\right)}\frac{n_{j,t}}{n_{j,t-1}} + \psi_{j,t}\left(\equiv mc^{R}_{j,t}\right). \tag{30}$$

We define the term on the right hand side as $mc^{R}_{j,t}$, which implies marginal cost of retention. DNWR introduces an additional (state-dependent) shadow cost of raising wages today because higher wages are harder to reverse in the future. We summarize this effect as an additive term $\psi_{j,t} \geq 0$, decreasing in inflation and increasing in the strength of DNWR. The term $\psi_{j,t}$ captures the option value of wage flexibility lost under DNWR.

Implications for firms operating on the retention margin. Firms that operate on the retention margin optimally adjust employment by raising wages to reduce worker separations rather than by posting additional vacancies. For these firms, the marginal cost of expanding or maintaining employment is given by $mc^R_{j,t}$ in eq. (30).

Eq. (30) highlights three key determinants of the marginal cost of retention. First, a higher separation elasticity ϵ^δ or a higher baseline separation rate $\delta(\cdot)$ lowers $mc^R_{j,t}$ by making wage increases more effective at retaining workers. As a result, firms in more fluid labor markets find it optimal to rely more heavily on the retention margin and to set higher wages relative to productivity.

Second, the presence of DNWR raises $mc^R_{j,t}$ through the shadow cost $\psi_{j,t}$. When future wage cuts are costly, raising wages today becomes less attractive, shifting firms away from the retention margin and toward vacancy posting. Higher inflation reduces this shadow cost by relaxing DNWR, thereby lowering $mc^R_{j,t}$ and encouraging wage-based adjustment.

Finally, because wages directly enter $mc^R_{j,t}$, firms operating on the retention margin face a tight link between wage setting and employment outcomes. In contrast to firms on the hiring margin, for which wages affect employment only indirectly, retention-margin firms internalize the employment consequences of wage changes one-for-one. This distinction plays a central role in shaping wage markdowns and the distribution of wages across firms.

5.2 Wage Markdown and Aggregate Wage

Given the optimality conditions we derive in the above, we analyze how inflation and the elasticity of separation would matter for both the aggregate wage and productivity. Note that we abbrebiate the time notation, since we are interested in the long-term implication.

Definition of Wage Markdown. We define the wage markdown of the individual firm j as:

$$\Omega_{j,t} \equiv \frac{w_{j,t}}{MRPL_{j,t}}. \tag{31}$$

Then we obtain an exact characterization of the wage markdown:

$$\Omega_{j,t} = \frac{w_{j,t}}{w_{j,t} + mc^H_{j,t} + \mathcal{D}_{j,t} - \mathcal{M}_{j,t}} = \frac{w_{j,t}}{w_{j,t} + mc^R_{j,t} + \mathcal{D}_{j,t} - \mathcal{M}_{j,t}}. \tag{32}$$

$mc^H_{j,t} = mc^R_{j,t}$ implies a following:

$$\frac{c^v(\chi+1)}{1-\mu} \frac{v^{\chi+\mu}_{j,t}}{n^{\chi}_{j,t-1} u^{\mu}_{t-1}} = \frac{w_{j,t}}{\epsilon^\delta \, \delta\left(\frac{w_{j,t}}{W_t}\right)} \frac{n_{j,t}}{n_{j,t-1}} + \psi_{j,t}. \tag{33}$$

Definition of aggregate wages. Let aggregate wages be the employment-weighted mean of firm wages:

$$W_t \equiv \frac{\int w_{j,t} n_{j,t}\, dj}{\int n_{j,t}\, dj} = \int w_{j,t} s_{j,t}\, dj, \qquad s_{j,t} \equiv \frac{n_{j,t}}{N_t}. \tag{34}$$

Assuming $w_{j,t}$ and $s_{j,t}$ are differentiable in time and $\int s_{j,t} dj = 1$, we have the exact identity

$$d \ln W_t = \underbrace{\int s_{j,t}\, d \ln w_{j,t}\, dj}_{\text{within (wage changes)}} + \underbrace{\int \left(\frac{w_{j,t}}{W_t} - 1 \right) ds_{j,t}\, dj}_{\text{between (composition)}}. \tag{35}$$

Equation (35) is an accounting identity (no approximation).

5.2.1 Role of inflation

Higher inflation relaxes downward nominal wage rigidity (DNWR), lowering the shadow cost of wage adjustment,

$$\pi_t \uparrow \quad \Rightarrow \quad \psi_{j,t}(\pi_t) \downarrow \quad \text{and} \quad \mathcal{D}_{j,t}(\pi_t) \downarrow \tag{36}$$

and thereby reducing the marginal cost of retention,

$$mc^R_{j,t} = \frac{w_{j,t}}{\epsilon^\delta\, \delta(w_{j,t}/W_t)} \frac{n_{j,t}}{n_{j,t-1}} + \psi_{j,t}. \tag{37}$$

By lowering the cost of wage-based adjustment, higher inflation shifts firms toward the retention margin. Firms raise wages to reduce separations, compressing wage markdowns. As a result, firm-level wages increase and aggregate wages rise.

5.2.2 Role of separation elasticity

An increase in the quit elasticity ϵ^δ captures greater labor market fluidity. From the employment law of motion,

$$\frac{\partial n_{j,t}}{\partial w_{j,t}} = \epsilon^\delta \delta\Big(\tfrac{w_{j,t}}{W_t} \Big) \frac{n_{j,t-1}}{w_{j,t}}, \tag{38}$$

a higher ϵ^δ strengthens the retention channel by making separations more responsive to relative wages. A higher separation elasticity lowers the marginal cost of retention, inducing firms to rely more on wage increases to stabilize employment. Wage markdowns shrink, firm-level wages rise, and aggregate wages increase.

5.3 Aggregate Labor Productivity

Remember that each firm j produces using labor only, $y_{j,t} = a_{j,t} n_{j,t}^{\alpha}$, where $a_{j,t}$ denotes firm-level productivity. Aggregate output and employment are given by:

$$Y_t = \int y_{j,t}\, dj, \qquad N_t = \int n_{j,t}\, dj. \tag{39}$$

Remember that $s_{j,t} \equiv n_{j,t}/N_t$ denote firm j's employment share. The aggregate labor productivity is defined as

$$P_t \equiv \frac{Y_t}{N_t} = \int a_{j,t} s_{j,t}^{\alpha} N_t^{\alpha-1}\, dj. \tag{40}$$

Let $\omega_{j,t} \equiv y_{j,t}/Y_t$ denote the output share of firm j. By the Divisia index identity, aggregate output growth satisfies

$$d\ln Y_t = \int \omega_{j,t}\, d\ln y_{j,t}\, dj. \tag{41}$$

Using $d\ln y_{j,t} = d\ln a_{j,t} + \alpha\, d\ln n_{j,t}$ and subtracting $d\ln N_t$, aggregate labor productivity growth can be written as

$$d\ln P_t = \underbrace{\int \omega_{j,t}\, d\ln a_{j,t}\, dj}_{\text{within-firm effect}} + \underbrace{\alpha \int \omega_{j,t}\, d\ln s_{j,t}\, dj}_{\text{between-firm (reallocation) effect}} + \underbrace{(\alpha-1) d\ln N_t}_{\text{employment dilution}}, \tag{42}$$

Eq. (42) is an exact identity of the aggregate productivity. Inflation and the separation elasticity primarily affect productivity through the between-firm term, by altering firms' incentives to retain or shed workers and hence the allocation of labor across firms as we examine in the followings.

Aggregate labor productivity responds to shocks through three distinct channels. The first term captures the direct (within-firm) effect of productivity changes. The second term reflects reallocation of employment toward more productive firms. The third term captures a mechanical dilution effect: because $\alpha < 1$, an expansion in aggregate employment reduces output per worker ceteris paribus.

5.3.1 Role of inflation

The reduction in DNWR also improves allocative efficiency. High-productivity firms can raise wages to attract and retain workers, while low-productivity firms are less able to hoard labor. Employment shares $s_{j,t}$ therefore become more positively correlated with productivity $a_{j,t}$, increasing the between-firm term in (??) and raising labor productivity Y_t/N_t.

5.3.2 Role of separation elasticity

Greater labor market fluidity enhances worker mobility across firms. Employment responds more strongly to wage differentials, allowing high-productivity firms to expand and low-productivity

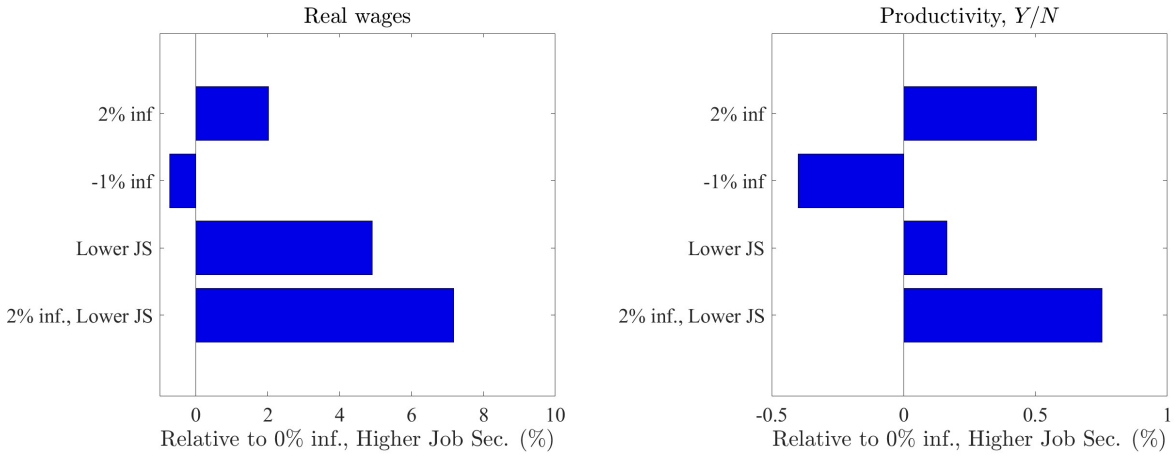


Figure 10: Steady state comparison with the benchmark of 0% inflation and a higher preference for job security ($\epsilon^{\delta} = 1.4$), where a lower preference for job security corresponds to $\epsilon^{\delta} = 1.6$

firms to contract. This increases the covariance between productivity and employment shares, amplifying the between-firm reallocation effect in (??) and raising aggregate labor productivity.

Overall, separations discipline firms' wage-setting even without bargaining: lowering wages increases separations, forcing the firm to either incur hiring costs or raise wages to retain workers. DNWR amplifies these forces by making future wage reductions costly, while inflation mitigates DNWR and thereby reduces the wage markdown.

5.4 Quantitative Analysis

Although higher inflation and a higher separation elasticity operate through different mechanisms, their macroeconomic effects are complementary. Inflation lowers the cost of adjusting wages by mitigating DNWR, while a higher ϵ^{δ} increases the effectiveness of wage adjustments in influencing worker flows. In both cases, aggregate wages and labor productivity rise because labor is reallocated toward more productive firms, strengthening the between-firm component of productivity.

Figure 10 presents quantitative results from our model. We consider 0% inflation and a stronger preference for job security as the baseline. We then compare the aggregate wage and productivity against scenarios with other inflation levels and variations in job security preferences. The findings indicate that both higher inflation and less emphasis on job security lead to increases in aggregate wage and productivity, aligning with theoretical predictions. When both higher inflation and reduced importance of job security coincide, the effects on wages and productivity become amplified in a non-linear manner.

6 The Propagation of Productivity Shocks

This section provides an analytical characterization of how a positive aggregate productivity shock affects aggregate wages and aggregate labor productivity in the presence of DNWR. The key ob-

jective is to show how these responses depend on (i) the steady-state inflation rate π and (ii) the elasticity of separations with respect to relative wages, denoted by ϵ^{δ}. The analysis proceeds in three steps. First, we derive an exact decomposition of aggregate labor productivity. Second, we characterize firms' employment adjustment margins using first-order conditions. Third, we combine these results to derive comparative statics with respect to π and ϵ^{δ}.

Aggregate Productivity Shocks We assume that firm-level productivity can be decomposed as:

$$a_{j,t} = z_t u_j, \tag{43}$$

where z_t is an aggregate productivity component and u_j is a time-invariant idiosyncratic component. A positive aggregate productivity shock corresponds to $d \ln z_t > 0$. Since u_j is fixed, it follows immediately that

$$\int \omega_{j,t} \, d \ln a_{j,t} \, dj = d \ln z_t. \tag{44}$$

The within-firm component of aggregate productivity growth is pinned down by the aggregate shock itself and is independent of labor market frictions, inflation, or separation elasticities. All interaction between productivity shocks and labor market institutions must therefore operate through the reallocation and dilution terms in (42).

6.1 The response of aggregate wage with respect to π and ϵ^{δ}

This subsection derives how the response of aggregate wages to a productivity shock varies with the steady-state inflation rate π and the separation elasticity ϵ^{δ}. We proceed in two steps: (i) an exact chain-rule decomposition of $\partial_{\pi} d \ln W_t$ and $\partial_{\epsilon^{\delta}} d \ln W_t$, and (ii) a first-order (local) approximation that maps these derivatives to objects implied by firms' policies and marginal-cost terms.

We consider the mapping from parameters (π, ϵ^{δ}) to the equilibrium objects $\{w_{j,t}, s_{j,t}, W_t\}$ along the IRF to a productivity shock. For notational economy, define the within and between terms in (35) as:

$$\mathcal{W}_t \equiv \int s_{j,t} \, d \ln w_{j,t} \, dj, \qquad \mathcal{B}_t \equiv \int \left(\frac{w_{j,t}}{W_t} - 1 \right) ds_{j,t} \, dj, \tag{45}$$

so that $d \ln W_t = \mathcal{W}_t + \mathcal{B}_t$.

To differentiate under the integral sign, we assume: (i) for each j, $w_{j,t}(\pi, \epsilon^{\delta})$ and $s_{j,t}(\pi, \epsilon^{\delta})$ are continuously differentiable in (π, ϵ^{δ}) in a neighborhood of interest; (ii) there exist integrable envelopes that dominate the partial derivatives.

Under the regularity conditions, we obtain:

$$\frac{\partial}{\partial x} d \ln W_t = \frac{\partial \mathcal{W}_t}{\partial x} + \frac{\partial \mathcal{B}_t}{\partial x}, \quad x \in \{\pi, \epsilon^{\delta}\} \tag{46}$$

with

$$\begin{aligned}\frac{\partial \mathcal{W}_t}{\partial x} &= \underbrace{\int \left(\frac{\partial s_{j,t}}{\partial x}\right) d\ln w_{j,t}\, dj}_{J_1^x} + \underbrace{\int s_{j,t}\, \frac{\partial}{\partial x}\Big(d\ln w_{j,t}\Big)\, dj}_{J_2^x}, \\ \frac{\partial \mathcal{B}_t}{\partial x} &= \underbrace{\int \frac{\partial}{\partial x}\Big(\frac{w_{j,t}}{W_t} - 1\Big)\, ds_{j,t}\, dj}_{J_3^x} + \underbrace{\int \Big(\frac{w_{j,t}}{W_t} - 1\Big)\, \frac{\partial}{\partial x}\Big(ds_{j,t}\Big)\, dj}_{J_4^x}.\end{aligned} \tag{47}$$

J_1^x: **Composition–growth interaction (reweighting).** These terms capture reweighting: the parameter changes employment shares toward firms that experience larger wage growth under the productivity shock. They are driven by the covariance between parameter-induced share shifts and wage growth across firms.

J_2^x: **Pure within effects (pass-through in wage setting).** These measure how the parameter changes firm-level wage growth holding shares fixed. Under DNWR, higher π reduces the shadow cost of potential future wage cuts, increasing within pass-through. A larger ϵ^δ raises the marginal benefit of wage increases for retention (since separations become more wage-sensitive), which can also increase within wage adjustments when firms are on the retention margin.

J_3^x: **Relative-wage general-equilibrium interaction.** Because $w_{j,t}/W_t$ depends on the aggregate index W_t, these are inherently general-equilibrium "quotient" effects. Their sign is typically not robust without additional structure.

J_4^x: **Parameter amplification of reallocation (between-gain).** These terms capture how the parameter changes the magnitude of employment reallocation toward high-wage firms. They are large when the retention margin maps wage changes into employment strongly, which is precisely controlled by ϵ^δ.

6.1.1 Two-type representation: sign conditions in closed form

To make sign conditions transparent, consider two types $k \in \{H, L\}$ (high/low productivity), with shares $s_H + s_L = 1$ and relative wages $\widetilde{w}_k \equiv w_k/W$. Let the shock of interest be a positive productivity shock ("z-shock").

In this setting, the four terms become:

$$J_1^x = (\partial_x s_H)\,(d\ln w_H - d\ln w_L), \tag{48}$$

$$J_2^x = s_H\, \partial_x(d\ln w_H) + s_L\, \partial_x(d\ln w_L), \tag{49}$$

$$J_3^x = (\partial_x \widetilde{w}_H - \partial_x \widetilde{w}_L)\, ds_H, \tag{50}$$

$$J_4^x = (\widetilde{w}_H - \widetilde{w}_L)\, \partial_x(ds_H). \tag{51}$$

6.1.2 Role of ϵ^δ.

Under iso-elastic separations of the form:

$$\delta\Big(\frac{w_{j,t}}{W_t}\Big) = \bar{\delta}\Big(\frac{w_{j,t}}{W_t}\Big)^{-\epsilon^\delta}, \tag{52}$$

the wage sensitivity of employment through retention scales with ϵ^δ:

$$\partial_w n_{j,t} \;\propto\; \epsilon^\delta\, \delta(\cdot)\, \frac{n_{j,t-1}}{w_{j,t}}. \tag{53}$$

Thus, higher ϵ^δ strengthens the mapping from wage changes to employment changes, amplifying ds_H and therefore the between terms ($J_4^{\epsilon^\delta}$, and indirectly $J_1^{\epsilon^\delta}$ via reweighting). In addition, when firms choose wages to retain workers, a larger ϵ^δ increases the marginal benefit of raising wages, which tends to raise $\partial_{\epsilon^\delta}(d \ln w_k)$ (hence $J_2^{\epsilon^\delta}$) on the retention margin.

Higher ϵ^δ raises the responsiveness of employment to wages (53), strengthening reallocation amplification (J_4^π) and often within adjustments (J_2^π) on the retention margin. Therefore, for given DNWR strength, higher ϵ^δ expands the parameter region where $\partial_\pi d \ln W_t > 0$ holds.

6.1.3 Role of π.

The quantitative importance of $(J_1^\pi, \ldots, J_4^\pi)$ is governed by three structural elements:

(i) Retention vs hiring: making retention dominant When employment adjustment is mainly via retention, ϵ^δ is the key amplifier. A sufficient set of restrictions making retention dominant is:

$$\begin{aligned} &\text{(a) high vacancy posting cost level/curvature;} \\ &\text{(b) tight matching so hires are expensive;} \\ &\text{(c) DNWR binds for some firms.} \end{aligned} \tag{54}$$

Under (54), within pass-through (J_2^π) and reallocation amplification (J_4^π) become the dominant inflation channels, while J_3^π is typically small and fragile (GE quotient effect).

(ii) DNWR strength and inflation sensitivity Let $\psi_{j,t}(\pi)$ denote the shadow (or adjustment) cost associated with downward nominal wage rigidity. The inflation effect on within wages is stronger when

$$-\,\partial_\pi \psi_{j,t}(\pi) \text{ is large for a non-negligible mass of firms.} \tag{55}$$

This restriction raises J_2^π directly and tends to increase J_4^π indirectly through larger wage-induced employment adjustments.

(iii) Cross-sectional productivity dispersion Reallocation channels require heterogeneity:

$$\text{Productivity dispersion is large so that } d\ln w_H - d\ln w_L \text{ is sizable under the } z\text{-shock.} \tag{56}$$

Under (56), J_1^π and J_4^π become larger whenever π or ϵ^δ tilt shares toward H.

Dominance map. Under the combination of (54)–(56): J_2^π is large when (55) holds (inflation relaxes DNWR strongly). J_4^π is large when (53) is strong (large ϵ^δ) and (56) holds. J_1^π is typically positive but smaller, acting as a covariance/reweighting correction. J_3^π is the least robust (GE quotient effect) and is often small. The separation elasticity ϵ^δ governs the strength of retention and thus the size of reallocation-based contributions. Under retention dominance, strong DNWR sensitivity to inflation, and sufficient productivity dispersion, higher inflation increases aggregate wage growth.

6.2 Aggregate Productivity

This section builds on the analytical results derived above and provides (i) sufficient conditions under which the reallocation effect dominates employment dilution following a positive productivity shock, and (ii) a quantitative interpretation of these conditions using empirically plausible values of the separation elasticity ϵ^δ, with particular attention to the Japanese labor market.

We show that higher inflation and greater separation elasticity unambiguously strengthen the reallocation channel, while their effect on aggregate employment introduces a countervailing dilution force. Whether aggregate labor productivity responds more strongly to a productivity shock in a high-inflation or high-mobility environment depends on the balance between these two forces. This trade-off provides a transparent mechanism linking monetary conditions, labor market fluidity, and the propagation of productivity shocks.

6.2.1 Reallocation effect and dilution effect

Recall that aggregate labor productivity growth satisfies:

$$d\ln P_t = d\ln z_t + \alpha\mathcal{R}_t + (\alpha-1)d\ln N_t, \tag{57}$$

where $\mathcal{R}_t \equiv \int \omega_{j,t} d\ln s_{j,t} dj$ denotes the reallocation term. Here, we use the condition (44) that the within-firm component of aggregate productivity growth is pinned down by the aggregate shock itself and is independent of labor market frictions, inflation, or separation elasticities.

Proposition 1 (Reallocation-Dominance Condition)**.** *Following a positive aggregate productivity shock* $d\ln z_t > 0$*, aggregate labor productivity responds more strongly in environments with higher inflation*

π or higher separation elasticity ϵ^δ if the following condition holds:

$$\alpha \frac{\partial \mathcal{R}_t}{\partial x} > (1-\alpha) \frac{\partial d \ln N_t}{\partial x}, \qquad x \in \{\pi, \epsilon^\delta\}. \tag{58}$$

Condition (58) states that the marginal increase in the reallocation component induced by x must outweigh the marginal employment dilution effect. Since $(1-\alpha) > 0$, the right-hand side captures the mechanical reduction in output per worker associated with aggregate employment expansion. The left-hand side reflects the efficiency gain from reallocating labor toward more productive firms.

6.2.2 Why π and ϵ^δ Increase Reallocation

From the firm-level analysis, the marginal cost of retention is given by

$$MC^R_{j,t} = \frac{w_{j,t}}{\epsilon^\delta \delta(w_{j,t}/W_t)} + \psi_{j,t}(\pi), \tag{59}$$

with $\partial MC^R_{j,t}/\partial \pi < 0$ and $\partial MC^R_{j,t}/\partial \epsilon^\delta < 0$.

Hence, an increase in either π or ϵ^δ expands the set of firms for which retention is the dominant adjustment margin,

$$\mathcal{J}^R_t \equiv \{j : MC^R_{j,t} \leq MC^H_{j,t}\}. \tag{60}$$

Lemma 1 (Reallocation Amplification)**.** *Suppose that firm-level productivity $a_{j,t}$ is positively correlated with the marginal revenue product of labor $MPL_{j,t}$. Then an increase in π or ϵ^δ weakly increases $\mathcal{R}_t$ following a productivity shock.*

Proof of sketch. Higher π or ϵ^δ reduces $MC^R_{j,t}$, shifting high-productivity firms into $\mathcal{J}^R_t$. These firms optimally raise wages by more in response to a productivity shock, leading to larger increases in employment through the retention channel:

$$dn^{(R)}_{j,t} = \epsilon^\delta \delta(\cdot) \frac{n_{j,t-1}}{w_{j,t}} dw_{j,t}.$$

Since $dw_{j,t}$ is increasing in $a_{j,t}$ under retention, employment shares shift toward high-productivity firms, implying $\mathcal{R}_t$ increases. □

6.2.3 Employment Dilution and Its Dependence on π and ϵ^δ

Aggregate employment growth can be written as

$$d \ln N_t = \int s_{j,t} d \ln n_{j,t} dj. \tag{61}$$

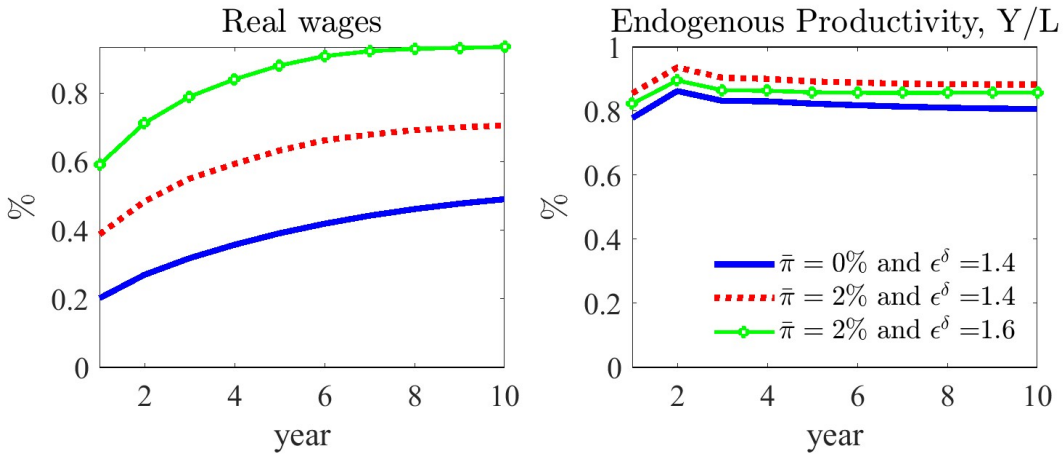


Figure 11: IRFs to an exogeneous 1% productivity shock with persistence ≈ 1

Both higher inflation and higher separation elasticity increase firms' incentives to expand employment, either through retention or hiring. Therefore,

$$\frac{\partial d \ln N_t}{\partial \pi} \geq 0, \qquad \frac{\partial d \ln N_t}{\partial \epsilon^{\delta}} \geq 0. \tag{62}$$

While π and ϵ^{δ} unambiguously strengthen reallocation, their net effect on aggregate productivity is ambiguous ex ante, as employment dilution works in the opposite direction. This motivates a quantitative evaluation.

6.3 Dynamic Analysis

Figure 11 shows the IRFs of macro real wages and endogenous productivity, which is defined as Y/L. We assume that the economy is exposed to a 1% permanent increase in productivity due to a shock. These results indicate that when the inflation rate is consistently high and the elasticity of separation is elevated, a productivity-increasing shock leads to greater increases in real wages and macro-level productivity. These findings are consistent with the aforementioned theory, and importantly, they highlight that minimal labor market frictions, including downward nominal wage rigidity (DNWR), play a crucial role in facilitating resource reallocation across firms.

7 Conclusion

This paper empirically examines whether firms' rigid wage-setting practices have changed over the last decades and analyzes the impact of relaxing such rigidities on aggregate productivity using a theoretical quantitative model.

In the empirical analysis, administrative microdata on workers in Japan were utilized to quantitatively examine nominal wage rigidity through the application of a friction model. The results

suggest that, since the late 1990s, the number of workers affected by rigid wage-setting practices has increased under deflationary regime, potentially suppressing wage growth rates. However, in recent years, during a period of rising inflation, potential wage growth rates have been driven upward by increasing inflation rates. Consequently, the share of workers facing upward nominal wage rigidity has decreased, indicating that the impact of rigid wage-setting practices has been alleviated.

In the theoretical analysis, we examined the impact of nominal wage rigidity on productivity using a general equilibrium model that incorporates two key assumptions: (i) downward nominal wage rigidity and (ii) the trade-off between job security and wage increases in labor-management negotiations. The results indicate that rigid wage-setting practices, combined with factors such as declining trend inflation and the practice of prioritizing job security in labor-management negotiations — both of which exert downward pressure on nominal wage growth rates — can hinder efficient resource allocation across firms and theoretically suppress aggregate productivity.

References

Aktug, Emrehan. 2025. "Downward nominal wage rigidity and the optimal inflation target." *Macroeconomic Dynamics* 29 e108.

Altonji, Josepth, and Paul J Devereux. 1999. "The extent and consequences of downward nominal wage rigidity."

Aoki, Kōsuke, Yoshihiko Hogen, Yojiro Ito, Kenji Kanai, and Kosuke Takatomi. 2024. "Determinants of Price Markups at Japanese Firms and Implications for Productivity."Technical report, Bank of Japan.

Aoki, Kōsuke, Yoshihiko Hogen, and Kosuke Takatomi. 2023. "Price markups and wage setting behavior of Japanese firms."Technical report, Bank of Japan.

Auclert, Adrien, Bence Bardóczy, Matthew Rognlie, and Ludwig Straub. 2021. "Using the sequence-space Jacobian to solve and estimate heterogeneous-agent models." *Econometrica* 89 (5): 2375–2408.

Autor, David, David Dorn, Lawrence F Katz, Christina Patterson, and John Van Reenen. 2020. "The fall of the labor share and the rise of superstar firms." *The Quarterly journal of economics* 135 (2): 645–709.

Baqaee, David Rezza, and Emmanuel Farhi. 2020. "Productivity and misallocation in general equilibrium." *The quarterly journal of economics* 135 (1): 105–163.

de la Barrera, Marc, and Masao Fukui. 2025. "Firm Dynamics with Downward Nominal Wage Rigidity."

Berger, David, Kyle Herkenhoff, and Simon Mongey. 2022. "Labor market power." *American Economic Review* 112 (4): 1147–93.

Bergholt, Drago, Francesco Furlanetto, and Nicolò Maffei-Faccioli. 2022. "The decline of the labor share: new empirical evidence." *American Economic Journal: Macroeconomics* 14 (3): 163–198.

Bilal, Adrien, Niklas Engbom, Simon Mongey, and Giovanni L Violante. 2022. "Firm and worker dynamics in a frictional labor market." *Econometrica* 90 (4): 1425–1462.

Bloesch, Justin, Birthe Larsen, and Anders Yding. 2024. "Monopsony with Recruiting." *Available at SSRN.*

Branten, Eva, Ana Lamo, and Tairi Room. 2018. "Nominal wage rigidity in the EU countries before and after the Great Recession: evidence from the WDN surveys."

Burdett, Kenneth, and Dale T Mortensen. 1998. "Wage differentials, employer size, and unemployment." *International economic review* 257–273.

Chib, Siddhartha. 1992. "Bayes inference in the Tobit censored regression model." *Journal of Econometrics* 51 (1-2): 79–99.

Christiano, Lawrence J, Martin Eichenbaum, and Charles L Evans. 2005. "Nominal rigidities and the dynamic effects of a shock to monetary policy." *Journal of political Economy* 113 (1): 1–45.

Christofides, Louis N, and Dingding Li. 2005. "Nominal and real wage rigidity in a friction model." *Economics Letters* 87 (2): 235–241.

Daly, Mary C, and Bart Hobijn. 2014. "Downward nominal wage rigidities bend the Phillips curve." *Journal of Money, Credit and Banking* 46 (S2): 51–93.

Date, Daiki, Takushi Kurozumi, Takashi Nakazawa, and Yu Sugioka. 2024. "Heterogeneity and wage growth of full-time workers in Japan: An empirical analysis using micro data." *Journal of the Japanese and International Economies* 73 101324.

Dickens, William T, Lorenz Goette, Erica L Groshen, Steinar Holden, Julian Messina, Mark E Schweitzer, Jarkko Turunen, and Melanie E Ward. 2007. "How wages change: micro evidence from the International Wage Flexibility Project." *Journal of Economic Perspectives* 21 (2): 195–214.

Dobbelaere, Sabien, and Kozo Kiyota. 2018. "Labor market imperfections, markups and productivity in multinationals and exporters." *Labour Economics* 53 198–212.

Ehrlich, Gabriel, and Joshua Montes. 2024. "Wage rigidity and employment outcomes: Evidence from administrative data." *American Economic Journal: Macroeconomics* 16 (1): 147–206.

Elsby, Michael WL. 2009. "Evaluating the economic significance of downward nominal wage rigidity." *Journal of Monetary Economics* 56 (2): 154–169.

Fehr, Ernst, and Lorenz Goette. 2005. "Robustness and real consequences of nominal wage rigidity." *Journal of Monetary Economics* 52 (4): 779–804.

Foster, Lucia, John C Haltiwanger, and Cornell John Krizan. 2001. "Aggregate productivity growth: Lessons from microeconomic evidence." In *New developments in productivity analysis*, 303–372, University of Chicago Press.

Fueki, Takuji, Ichiro Fukunaga, Hibiki Ichiue, and Toyoichiro Shirota. 2018. "Measuring potential growth with an estimated DSGE model of Japan's economy." *43rd issue (March 2016) of the International Journal of Central Banking*.

Fukao, Keiji, Young Gak Kim, Hyuck Uk Kwon, and Kenta Ikeuchi. 2021. "Business Dynamics and Productivity Growth Under Abenomics: Analysis Using the 'Economic Census Activity Survey' Questionnaire Data." in Japanese.

Fukao, Kyoji, and Cristiano Perugini. 2021. "The long-run dynamics of the labor share in Japan." *Review of Income and Wealth* 67 (2): 445–480.

Goldin, Ian, Pantelis Koutroumpis, François Lafond, and Julian Winkler. 2024. "Why is productivity slowing down?" *Journal of Economic Literature* 62 (1): 196–268.

Goodhart, Charles Albert Eric, and Manoj Pradhan. 2020. *The great demographic reversal: Ageing societies, waning inequality, and an inflation revival*. Volume 10. Springer.

Grigsby, John, Erik Hurst, and Ahu Yildirmaz. 2021. "Aggregate nominal wage adjustments: New evidence from administrative payroll data." *American Economic Review* 111 (2): 428–471.

Grossman, Gene M, and Ezra Oberfield. 2022. "The elusive explanation for the declining labor share." *Annual Review of Economics* 14 (1): 93–124.

Hirata, Wataru, Toshitaka Maruyama, and Tomohide Mineyama. 2020. "Flattening of the Wage Phillips Curve and Downward Nominal Wage Rigidity: The Japanese Experience in the 2010s."Technical report, Bank of Japan.

Hopenhayn, Hugo, and Richard Rogerson. 1993. "Job turnover and policy evaluation: A general equilibrium analysis." *Journal of political Economy* 101 (5): 915–938.

Ikeuchi, Kenta, Young Gak Kim, Hyuck Uk Kwon, Kyoji Fukao et al. 2018. *Productivity Dynamics in Small and Medium Enterprises: Empirical Analysis Using the Credit Risk Database (CRD)*. Institute of Economic Research, Hitotsubashi University, , in Japanese.

Kambayashi, Ryo. 2011. "Nominal Wage Rigidity in Japan (1993-2006): An Approach Using Pseudo Panel Data." 62 (4): 301–307, in Japanese.

Karabarbounis, Loukas. 2024. "Perspectives on the labor share." *Journal of Economic Perspectives* 38 (2): 107–136.

Kuroda, Sachiko, Isamu Yamamoto Sachico Kuroda, and Isamu Yamamoto. 2003. *Are Japanese nominal wages downwardly rigid?(Part II): Examinations using a friction model*. Institute for Monetary and Economic Studies, Bank of Japan.

Lodge, David, Mary Everett, Olivier De Bandt, Georgios Georgiadis, Michele Ca'Zorzi, Povilas Lastauskas, Juan Carluccio, Susana Parrága, Pietro Cova, Maria Grazia Attinasi et al. 2021. "The implications of globalisation for the ECB monetary policy strategy."

Mineyama, Tomohide. 2022. "Revisiting the optimal inflation rate with downward nominal wage rigidity: The role of heterogeneity." *Journal of Economic Dynamics and Control* 139 104350.

Ogura, Kazuya. 2017. "Management's Perspective on Wage Increase and Its Background." In *Why Are Wages Not Rising despite the Labor Shortage?*, edited by Genda, Yuji 44–58, Keio University Press, , in Japanese.

Omori, Yasuhiro, and Koji Miyawaki. 2010. "Tobit model with covariate dependent thresholds." *Computational statistics & data analysis* 54 (11): 2736–2752.

Ottonello, Pablo, and Thomas Winberry. 2020. "Financial heterogeneity and the investment channel of monetary policy." *Econometrica* 88 (6): 2473–2502.

Petrongolo, Barbara, and Christopher A Pissarides. 2001. "Looking into the black box: A survey of the matching function." *Journal of Economic literature* 39 (2): 390–431.

Schaefer, Daniel, and Carl Singleton. 2023. "The extent of downward nominal wage rigidity: New evidence from payroll data." *Review of Economic Dynamics* 51 60–76.

Tobin, J. 1972. "Inflation and Unempolyment." *AMER. ECON. REV* 62 (1): 1–18.

Watanabe, Tsutomu. 2024. *Thinking about Prices: The Mystery of Deflation and Inflation*. Nikkei Publishing, , in Japanese.

Yamamoto, Isao. 2007. "Wage Flexibility in Deflationary Periods: An International Comparative Perspective." *Mita Business Review* 50 (5): 1–14, in Japanese.

Yamamoto, Isao, and Sachiko Kuroda. 2016. "Does Past Wage Cuts Enhance Wage Flexibility?: Evidence from Firm-Level Panel Data." *RIETI Discussion Paper*, in Japanese.

Yeh, Chen, Claudia Macaluso, and Brad Hershbein. 2022. "Monopsony in the US labor market." *American Economic Review* 112 (7): 2099–2138.