UC Santa Cruz

UC Santa Cruz Electronic Theses and Dissertations

Title

Essays in Macroeconomics: Measuring Economic Slack in Labor Markets and Business Cycles

Permalink

https://escholarship.org/uc/item/0pf595rq

ISBN

9798190914665

Author

Zhang, Rongjin

Publication Date

2026-08-01

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA
SANTA CRUZ

ESSAYS IN MACROECONOMICS: MEASURING ECONOMIC SLACK IN LABOR MARKETS AND BUSINESS CYCLES

A dissertation submitted in partial satisfaction of the
requirements for the degree of

Doctor of Philosophy

in

ECONOMICS

by

Rongjin Zhang

August 2026

The Dissertation of Rongjin Zhang
is approved:

Pascal Michaillat, Chair

Kenneth Kletzer

Alonso Villacorta

Donald Smith
Interim Vice Provost and Dean of Graduate Studies

Copyright © by

Rongjin Zhang

2026

Table of Contents

List of Figures **vi**

List of Tables **xi**

Abstract **xiii**

Dedication **xvi**

Acknowledgments **xvii**

1 Beveridgean Unemployment Gap with Part-time Employment **1**
1.1 Introduction 1
1.2 Background and existing measures of labor-market slack 5
1.3 Review of the Beveridgean unemployment gap 9
1.3.1 Three sufficient statistics in Michaillat and Saez (2021a) . 9
1.4 Incorporating part-time employment 13
1.4.1 Flow social welfare with part-time employment 13
1.4.2 Two additional sufficient statistics 15
1.4.3 Estimation of remaining sufficient statistics 28
1.5 Application to the United States 31
1.5.1 Baseline results 33
1.5.2 Incorporation of part-time employment 36
1.5.3 Beveridgean unemployment gap in the United States, 1951–2019, 2020–2026 44
1.5.4 Robustness of U.S. results 46
1.6 Application to Japan 53
1.6.1 2008 Financial Crisis 65
1.6.2 2011 Earthquake Aftermath 66
1.6.3 Pandemic Period 67
1.6.4 Involuntary part-time 69
1.6.5 Inverse-optimum sufficient statistics 70
1.7 Conclusion 73

2 Recession Detection in Japan Using Labor Market Data 76
2.1 Introduction . . . 76
2.2 Literature Review . . . 79
2.3 Data . . . 82
2.3.1 Japan Time Series Data . . . 82
2.3.2 Business Cycle Dating . . . 82
2.3.3 GDP Data . . . 83
2.3.4 Unemployment and Vacancy Data . . . 84
2.3.5 Labor Market Dynamics . . . 85
2.4 Background: Sahm Rule and Michez Rule in Japan . . . 87
2.4.1 Sahm Rule in Japan . . . 88
2.4.2 Michez Rule in Japan . . . 91
2.5 Construction of the Recession Indicators . . . 92
2.5.1 Smoothing the Data . . . 94
2.5.2 Detecting Turning Points . . . 95
2.5.3 Scaling Variations . . . 96
2.5.4 Combining Single-Variable Indicators . . . 97
2.5.5 Summary . . . 100
2.6 Results . . . 101
2.6.1 Detection Methodology . . . 102
2.6.2 Selecting Perfect Classifiers . . . 102
2.6.3 Evaluating Perfect Classifiers . . . 103
2.6.4 Finding the Anticipation-Precision Frontier . . . 104
2.6.5 Selecting the Frontier Classifiers with Highest Precision . . . 107
2.7 Detecting Recessions . . . 114
2.7.1 Empirical Construction of Recession Probabilities . . . 114
2.7.2 Recession Probability from Classifier Ensemble . . . 116
2.7.3 Application of Recession Detection to Current Data . . . 117
2.8 Evaluating the recession detection algorithm with backtests . . . 118
2.8.1 Backtesting from 2015 . . . 119
2.8.2 Backtesting from 2005 . . . 119
2.8.3 Backtesting from 1995 . . . 120
2.8.4 Sensitivity Test: Excluding the 2012 recession episode . . . 121
2.9 Conclusion . . . 125

3 Recession Detection Using Real-time GDP Data 127
3.1 Introduction . . . 128
3.2 Related literature . . . 129
3.3 Data . . . 130
3.3.1 real-time GDP Data . . . 130
3.3.2 Business cycle dating . . . 131
3.4 Real-time GDP . . . 131

3.5 Constructing and evaluating recession classifiers 134
3.5.1 Detection methodology 134
3.5.2 Selecting perfect classifiers 134
3.5.3 Evaluating perfect classifiers 135
3.5.4 Finding the anticipation-precision frontier 136
3.5.5 Selecting the frontier classifiers with highest precision . . . 138
3.6 Detecting recessions 138
3.6.1 Empirical Construction of Recession Probabilities 139
3.6.2 Application of Recession Detection to Current Data 141
3.7 Evaluating the recession detection algorithm with backtests 144
3.7.1 Backtesting from 2015 144
3.7.2 Backtesting from 2005 146
3.8 Conclusion 147

A **149**
A.1 Notation relative to Michaillat and Saez (2021) 149
A.2 Additional results 151
A.2.1 Additional results for the United States 151
A.2.2 Additional Japanese results 157
A.2.3 Robustness test tables 159

B **162**
B.1 Additional Results for 1985-2015 Backtests 164
B.2 Additional results for the 1985–2015 backtests excluding the 2012 recession episode 168

C **170**
C.1 Additional results 170

List of Figures

1.1 Part-time employment in some OECD countries 3
1.2 Involuntary part-time employment in the U.S. *Source*: U.S. Bureau of Labor Statistics 4
1.3 OECD part-time employment and Japanese part-time employment. *Source*: OECD Database. 5
1.4 Efficient unemployment rate 12
1.5 Share of full-time and part-time employment in the United States, 1976–2025 18
1.6 Share of full-time and involuntary part-time employment in the United States, 1988–2025 19
1.7 Average weekly hours worked by full-time and part-time workers in the United States, 1976–2025 22
1.8 Average weekly hours worked by full-time and involuntary part-time workers in the United States, 1988–2025 23
1.9 Efficient unemployment rate with and without part-time employment 28
1.10 Beveridge curve estimation in the United States, 1951–2026 32
1.11 Efficient labor-market tightness, efficient unemployment, and unemployment gap in the United States under the full-time baseline, 1951–2026 35
1.12 Efficient labor-market tightness, efficient unemployment, and unemployment gap with part-time employment in the United States, 1951–2026 38

1.13 Comparison of efficient labor-market tightness, efficient unemployment, and unemployment gaps with and without total part-time employment in the United States 39
1.14 Efficient labor-market tightness, efficient unemployment, and unemployment gap in the United States 42
1.15 Comparison of efficient labor-market tightness, efficient unemployment, and unemployment gaps with and without involuntary part-time employment in the United States 43
1.16 Efficient unemployment rates for alternative sufficient statistics in the United States, 1951–2019 48
1.17 Efficient unemployment rates for alternative sufficient statistics in the United States, 2020–2026 49
1.18 Inverse-optimum sufficient statistics in the United States, 1951–2019 51
1.19 Inverse-optimum sufficient statistics in the United States, 2020–2026 52
1.20 Beveridge curve estimation in Japan, 1970–2025 54
1.21 Full-time and part-time employment in Japan, 2002–2025 57
1.22 Involuntary part-time and full-time employment in Japan, 2002–2025 58
1.23 Involuntary part-time employment as a share of total part-time employment in Japan, 2002–2025 59
1.24 Average weekly hours for part-time and full-time workers in Japan, 2013–2026 60
1.25 Efficient labor-market tightness, efficient unemployment rate, and unemployment gap in Japan, 1970–2025 64
1.26 Efficient unemployment in Japan using smoothed Beveridge elasticities estimated over Bai–Perron regimes 65
1.27 Labor-market tightness, efficient unemployment rate, and unemployment gap in Japan, 1970–2025 68
1.28 Efficient unemployment rates for alternative sufficient statistics in Japan, 1970–2025 72
1.29 Inverse-optimum sufficient statistics in Japan, 1970–2025 73

2.1 Japanese real GDP deviations from trend, 1970–2025 80
2.2 United States: Unemployment and Vacancy Rate 86
2.3 Japan: Unemployment Rate and Vacancy Rate 87
2.4 Sahm Rule Results for the United States 90
2.5 Sahm Rule Results for Japan 90
2.6 Michez Rule: Minimum of Two Indicators 92
2.7 Michez Rule Results 93
2.8 Increase in the Japanese Unemployment Rate (1970–2025) 98
2.9 Decrease in the Japanese Vacancy Rate (1970–2025) 99
2.10 4,356 Unemployment Indicators and 4,356 Vacancy Indicators . . 101
2.11 Anticipation and Precision of 1,844,564 Perfect Recession Classifiers in Japan, January 1970–December 2021 109
2.12 Classifier 1: Threshold 0.0005 110
2.13 Classifier 2: Threshold 0.0056 110
2.14 Classifier 3: Threshold 0.0081 111
2.15 Classifier 4: Threshold 0.0054 111
2.16 Classifier 5: Threshold 0.0073 112
2.17 Classifier 6: Threshold 0.0072 113
2.18 Full Set of All Six Recession Classifiers 113
2.19 In-sample recession probability, January 1970–December 2021 . . 116
2.20 Out-of-sample recession probability, January 2022–December 2025 117
2.21 Out-of-sample Japan Recession Probability, January 2015–December 2025 119
2.22 Out-of-sample Japan Recession Probability, January 2005–December 2025 120
2.23 Out-of-sample Japan Recession Probability, January 1995–December 2025 121
2.24 In-sample Recession Probability (excl.2012 recession), January 1970–December 2021 123
2.25 Out-of-sample Recession Probability, January 2022–December 2025 123
2.26 Out-of-sample Recession Probability, January 2015–December 2025 124

2.27 Out-of-sample Recession Probability (excl.2012 recession), January 2005–December 2025 124
2.28 Out-of-sample Recession Probability (excl.2012 recession), January 1995–December 2025 125

3.1 Real-Time Log GDP Gap from First-Release Real GDP 133
3.2 real-time GDP Indicator 133
3.3 Anticipation and Precision of 137,457 Perfect Recession Classifiers in the United States, January 1947–December 2021 137
3.4 Classifier 1: Threshold 0.0015 142
3.5 Classifier 2: Threshold 0.0027 142
3.6 In-sample Recession Probability, January 1947–December 2021. . 143
3.7 Out-of-sample Recession Probability, January 2022–December 2025 143
3.8 Out-of-sample U.S. Recession Probability, January 2015–December 2025 146
3.9 Out-of-sample U.S. Recession Probability, January 2005–December 2025 147

A.1 Part-time employment shares based on actual and usual hours in the CPS ASEC 151
A.2 Robustness to actual-hours and usual-hours part-time classifications in the United States 152
A.3 Efficient labor-market tightness, efficient unemployment, and unemployment gap in the United States, 1951–2026 153
A.4 Comparison of efficient labor-market tightness, efficient unemployment, and unemployment gaps with and without total part-time employment in the United States, 1951–2026 154
A.5 Efficient labor-market tightness, efficient unemployment, and unemployment gap in the United States, 1951–2026 155
A.6 Comparison of efficient labor-market tightness, efficient unemployment, and unemployment gaps with and without involuntary part-time employment in the United States, 1951–2026 156

A.7 Robustness to the pre-2002 full-time employment share in Japan . 157

B.1 Out-of-sample Japan recession probability, January 2015–December 2025 164
B.2 Out-of-sample Japan recession probability, January 2005–December 2025 164
B.3 Out-of-sample Japan recession probability, January 1995–December 2025 165
B.4 Out-of-sample Japan recession probability, January 1985–December 2025 165
B.5 In-sample recession probability, January 1970–December 2014 . . 168
B.6 In-sample recession probability, January 1970–December 2004 . . 168
B.7 In-sample recession probability, January 1970–December 1994 . . 169

C.1 Anticipation Precision Frontier 2015 170
C.2 Anticipation Precision Frontier 2005 171
C.3 In Sample Recession Probability, January 1947 - December 2014 . 172
C.4 In Sample Recession Probability, January 1947 - December 2004 . 172

List of Tables

1.1 Summary Statistics of Full-Time and Total Part-Time Employment Shares 20
1.2 Summary Statistics of Full-Time and Involuntary Part-Time Employment Shares 21
1.3 Summary Statistics of Average Weekly Hours: Total Part-Time (γ^P) 24
1.4 Summary Statistics of Average Weekly Hours: Involuntary Part-Time (γ^{IP}) 24
1.5 Beveridge elasticity estimates by regime, 1951Q1–2026Q1 30
1.6 Summary Statistics of Full-Time and Part-Time Employment . . 57
1.7 Summary Statistics of Full-Time and Involuntary Part-Time Employment 58
1.8 Summary Statistics of Part-Time and Involuntary Part-Time Employment 59
1.9 Summary Statistics of Weekly Hours and Relative Hours in Japan 60

2.1 ESRI Recession Start Announcements 80
2.2 Start Dates of Japanese Recessions (1970-2025) 84
2.3 Detected Recession Starts by Threshold for the Main Ensemble . 106
2.4 Classifier Ensemble Selected from the Anticipation–Precision Frontier 108
2.5 Performance of the Algorithm on Backtests 118

3.1 Classifier Performance Metrics 139
3.2 Out-of-Sample Backtesting Results 145

A.1 Summary Statistics of Labor-Market Tightness in Japan 158
A.2 Summary Statistics of Actual and Efficient Unemployment Rates in Japan 158
A.3 Summary Statistics of the Unemployment Gap in Japan 158
A.4 Summary Statistics for Efficient Unemployment and Beveridge-Elasticity Uncertainty 159
A.5 Summary Statistics for Efficient Unemployment and Social Value of Nonwork Uncertainty 159
A.6 Summary Statistics for Efficient Unemployment and Recruiting-Cost Uncertainty 160
A.7 Summary Statistics for Inverse-Optimum Beveridge Elasticity . . 160
A.8 Summary Statistics for Inverse-Optimum Social Value of Nonwork 161
A.9 Summary Statistics for Inverse-Optimum Recruiting Cost 161

B.1 Detected Recession Starts by Threshold for the Main Ensemble (continued) 163
B.2 Classifier Ensemble Selected in the 2015 Backtest 166
B.3 Classifier Ensemble Selected in the 2005 Backtest 167

C.1 Classifier Ensemble Selected in the 2015 Backtest 171
C.2 Classifier Ensemble Selected in the 2005 Backtest 171

Abstract

Essays in Macroeconomics: Measuring Economic Slack in Labor Markets and Business Cycles

by

Rongjin Zhang

This dissertation consists of three essays in macroeconomics. The first essay examines labor-market efficiency by incorporating part-time employment. The second and third essays study real-time recession detection, using labor market data in Japan and real-time GDP data in the United States.

Chapter 1 extends the sufficient-statistics formula for efficient unemployment developed by Michaillat and Saez (2021a) to account for part-time employment. I introduce two additional sufficient statistics that measure the share of part-time employment and part-time hours relative to full-time hours. Applying the framework to the United States (1951–2026) and Japan (1970–2025), I compare the effects of total part-time employment and involuntary part-time employment on efficient unemployment. Total part-time employment has substantially larger effects than involuntary part-time employment. While involuntary part-time employment provides information about labor-market slack, the main change in efficient unemployment comes from part-time work itself because part-time workers supply fewer market hours than full-time workers. Under the total part-time calibration, efficient unemployment averages 4.7% in the United States before COVID and 4.2% after COVID. In the Japanese application, the full-sample average is 2.7%. The distinction is especially important in Japan, where part-time employment is widespread and often reflects flexible work arrangements. These findings suggest that aggregate labor input, rather than involuntary part-time employment alone, is an important determinant of labor-market efficiency. Chapter 2,

a joint work with Neha Sikand, examines whether labor market–based recession indicators developed for the United States can be adapted to Japan. Recession indicators are often viewed as U.S.-specific, raising the question of whether labor market–based rules such as the Sahm Rule and the Michez Rule can reliably detect recessions in other countries. To answer this, we evaluate whether such rules can be adapted to Japan by calibrating thresholds and smoothing parameters to Japanese labor market data. We construct a large set of 95,832 recession indicators combining unemployment and vacancy data. The selected classifiers are statistically perfect as they identify all 11 historical recessions in the 1970–2021 training period without generating any false positives. Among these, 193 classifiers lie on the anticipation–precision frontier. Restricting attention to the high-precision segment yields six classifiers with a standard deviation of detection errors below 3 months. The selected classifier ensemble signals recessions, on average, 0.06 months after their true onset. Overall, these findings suggest that slack-based labor market rules provide a general framework for improving real-time recession detection across countries.

Chapter 3, a joint work with Neha Sikand, examines whether real-time GDP announcements can reliably identify business-cycle turning points. Using U.S. real-time GDP vintages from 1947 to 2021, we construct 4,356 recession indicators based on alternative smoothing methods and scaling variations. We then combine these recession indicators with different thresholds to obtain 137,457 perfect recession classifiers. The selected classifiers identify all 12 historical recessions without generating false positives. Restricting attention to the high-precision segment yields two classifiers with a standard deviation of detection errors below three months, while the selected ensemble signals recessions, on average, 3.04 months after official onset. Our framework accurately identifies recession episodes across

vintages, suggesting that real-time data revisions alone cannot fully account for discrepancies reported in prior work. These results suggest that real-time GDP announcements provide a practical real-time proxy for NBER-style recession dating.

To My Family

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my advisor, Professor Pascal Michaillat, who has truly been a light throughout my Ph.D. journey. I have always admired his work, his thinking, and the clarity with which he approaches economics. His guidance helped me build my path in macroeconomics and taught me to seek truth with discipline, simplicity, and clarity. Working closely with him has been one of the greatest privileges of my doctoral studies, and I am profoundly grateful for his wisdom, generosity, and belief in me.

I am also grateful to Professor Alonso Villacorta and Professor Kenneth Kletzer for their generous support, thoughtful guidance, and encouragement throughout this journey. I thank Professor Villacorta for the many hours of discussion he devoted to my projects, and for the trust, insight, and patience he offered along the way. I also thank Professor Kletzer for his kindness, advice, and encouragement, which made this journey lighter, more enjoyable, and more meaningful.

I am deeply grateful to my parents for their unconditional love and support. They have always respected my decisions and given me the freedom to pursue my own path. I thank my father for giving me the courage to keep moving forward and for passing on to me his entrepreneurial spirit, which made me bold enough to become the first person in our business family to earn a doctorate. I thank my mother for teaching me to believe in myself and not to be shaken by others' opinions. Above all, I thank them both for teaching me that being a good person matters more than anything else. I am grateful to stand where I am today with the character, strength, and values they have cultivated in me.

My grandparents also hold a special place in this journey. They raised me with unconditional love and quiet strength. My grandmother passed away in July 2019, while I was on my flight to the United States to pursue my academic dream,

and my grandfather passed away in 2025, shortly after I passed my qualifying exam. My grandmother on my mother's side also passed away in 2026, during the time of my graduation, making this moment of completion inseparable from grief. Their love shaped who I am long before this dissertation took form. Although they are no longer here to see this moment, everything I have achieved carries their presence.

Lastly, I would like to thank my fellow Ph.D. colleagues and friends, with whom I was fortunate to share this journey. I am deeply grateful to Neha Sikand, Drishan Banerjee, Yuchao Li, Dizhi Wang, Yifan Si, Yoshihisa Shono, Bhavyaa Sharma, Giancarlo Barcia, Yiqing Feng, Sanjana Gupta, Ayden Young, Taisuke Kameda, Ananya Pimpley, Gonzalo Martín Respighi Grasso, Kirill Zhazhin and many others in the department for the conversations, encouragement, laughter, and companionship that carried me through these years. Their presence made graduate school not only more manageable, but also far more meaningful and memorable.

Chapter 1

Beveridgean Unemployment Gap with Part-time Employment

1.1 Introduction

Workers who are classified as employed can differ substantially in the amount of labor they supply. Part-time workers typically work fewer hours than full-time workers and therefore contribute less market production. As a result, the same unemployment rate can correspond to different levels of aggregate labor input depending on the composition of employment. This distinction matters for welfare-based measures of labor-market efficiency, which depend on labor input rather than employment status alone.

Part-time employment has become increasingly common in many advanced economies. Figure 1.1 illustrates this trend across several OECD countries, showing a persistent rise in the part-time employment share over the past several decades. In several countries, part-time workers now account for a substantial fraction of total employment. The growing prevalence of part-time work therefore

makes the distinction between employment counts and labor input empirically relevant.

Part-time employment is generally defined as employment involving fewer than 35 hours of work per week. According to the U.S. Bureau of Labor Statistics, part-time workers can be classified as working part time for non-economic reasons or for economic reasons. The latter group consists of workers who would prefer additional hours of work but are unable to obtain full-time employment.

Interest in involuntary part-time employment has grown substantially because it is closely linked to business-cycle conditions. The cyclical behavior of part-time employment differs across worker groups. Total part-time employment combines voluntary and involuntary part-time work and therefore reflects both structural labor-supply choices and business-cycle conditions. Involuntary part-time employment is more directly linked to business-cycle conditions than total part-time employment. As shown in Figure 1.2, involuntary part-time employment in the United States rises sharply during recessions and often remains elevated well into subsequent recoveries (Valletta et al., 2020). These patterns suggest that involuntary part-time employment contains valuable information about labor-market fluctuations over the business cycle.

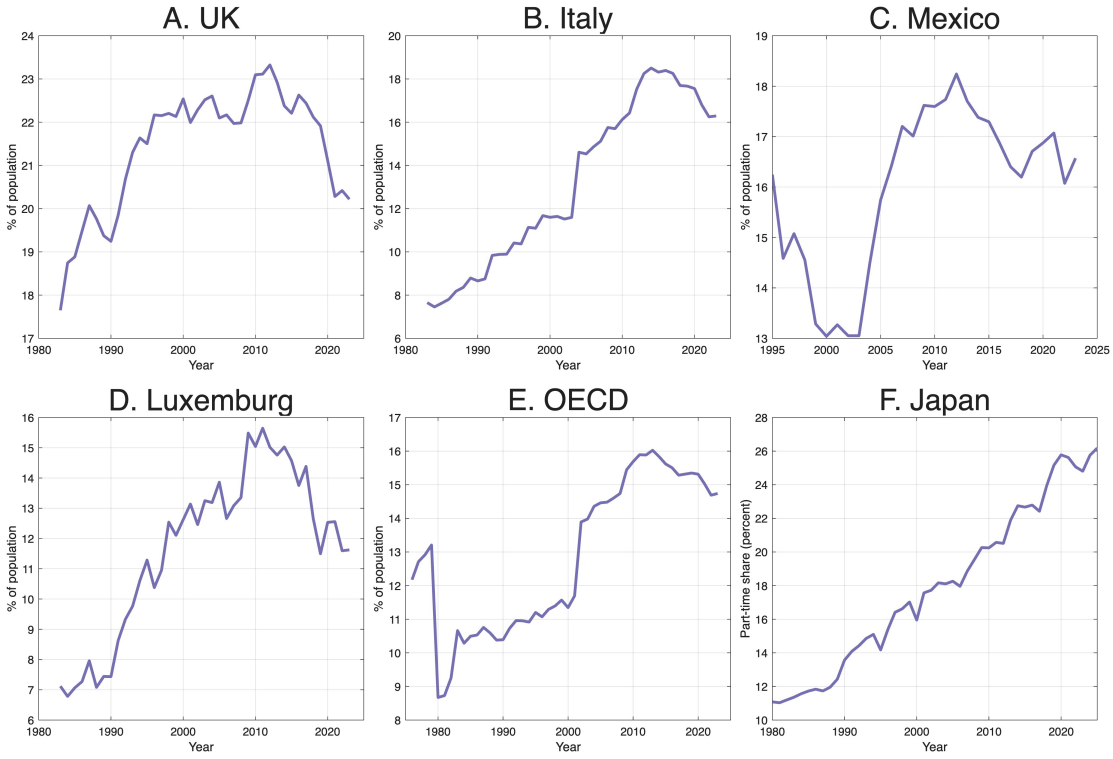


Figure 1.1: Part-time employment in some OECD countries
Source: OECD Database.

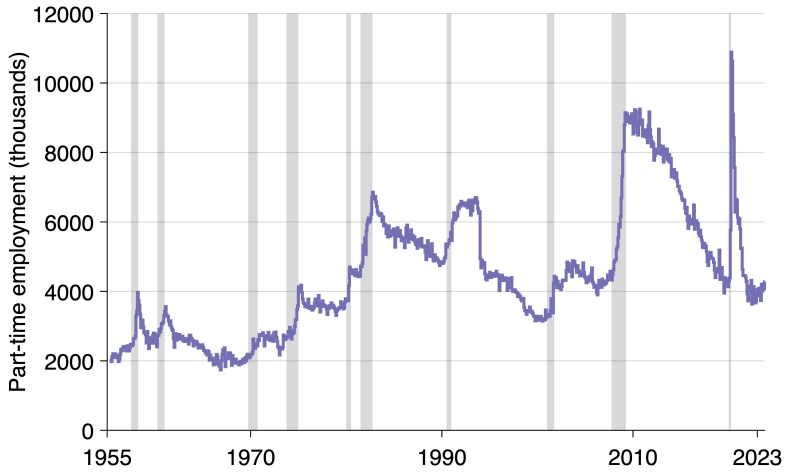


Figure 1.2: Involuntary part-time employment in the U.S. *Source*: U.S. Bureau of Labor Statistics

However, from a welfare perspective, labor-market conditions depend not only on whether workers are employed, but also on how much labor they supply. Even voluntarily chosen part-time employment contributes fewer market hours than full-time employment and therefore reduces effective labor input. As a result, two economies with identical unemployment rates may differ substantially in market production and labor-market efficiency if their employment structures differ.

To account for this dimension of employment heterogeneity, I introduce part-time employment into the Beveridgean unemployment-gap framework of Michaillat and Saez (2021a). In the extended framework, part-time workers contribute less market production than full-time workers but generate greater social value than unemployed workers. This extension yields a generalized measure of efficient unemployment that incorporates differences in hours worked among employed workers. I apply the framework to data from the United States and Japan to examine how part-time employment changes estimates of efficient unemployment and unemployment gaps. Figure 1.3 motivates the Japan application by showing that part-time employment is much more prevalent in Japan than in the OECD

average. While the OECD part-time employment rate rises after the early 2000s and stabilizes below 16 %, Japan's part-time employment rate increases steadily from about 16 % in 2002 to above 22 % around 2020 before declining slightly. This pattern suggests that part-time employment is an important feature of Japanese labor utilization and should be incorporated into the measurement of efficient unemployment.

I also separate involuntary part-time employment from total part-time employment and repeat the analysis to distinguish the effect of reduced labor input from the effect of underemployment.

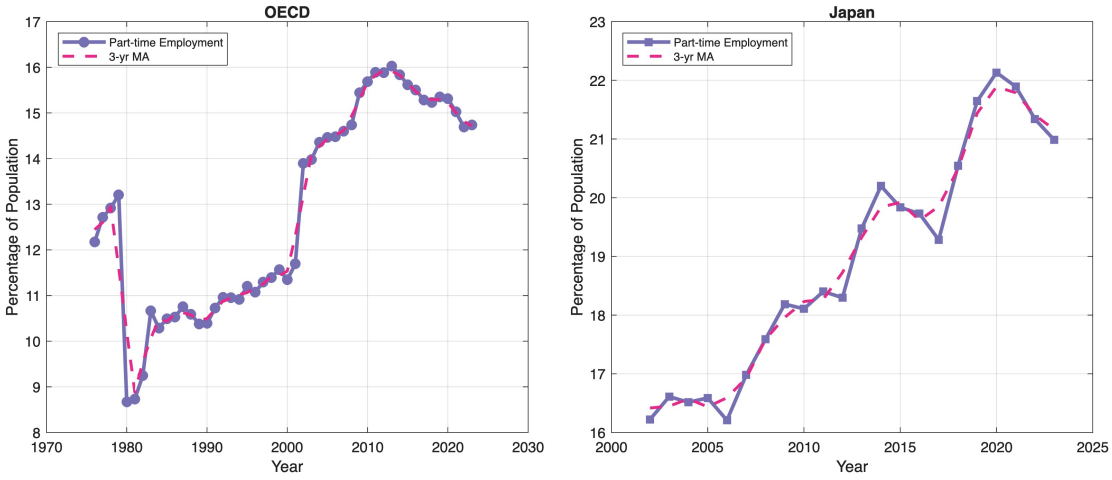


Figure 1.3: OECD part-time employment and Japanese part-time employment. *Source*: OECD Database.

1.2 Background and existing measures of labor-market slack

As discussed by Michaillat and Saez (2021a), two approaches are commonly used to assess unemployment gaps. The first measures deviations of the unemployment rate from its long-run trend, while the second compares the unemployment

rate to the non-accelerating inflation rate of unemployment (NAIRU). Although widely used, neither approach is derived from a social welfare framework and therefore neither directly measures the distance between observed labor-market outcomes and the welfare-maximizing allocation.

Discussions of part-time employment often focus on involuntary part-time employment, which is commonly interpreted as a form of underemployment. Statistical agencies such as the U.S. Bureau of Labor Statistics (BLS) measure underemployment using indicators such as the involuntary part-time rate (IPTR), defined as the share of workers employed part-time for economic reasons. Another widely used measure is the U-6 unemployment rate, which includes unemployed workers, marginally attached workers, and workers employed part-time for economic reasons. Because U-6 extends U-5 by incorporating involuntary part-time workers, the gap between the two measures is often interpreted as reflecting the contribution of involuntary part-time employment to labor-market slack. While informative, these indicators have important limitations. First, they identify the number of involuntary part-time workers but do not measure the intensity of underemployment, such as the extent to which these workers would like to increase their hours. Second, because they are based on worker counts rather than hours worked, they may not fully capture differences in labor utilization across economies. As emphasized by Bell and Blanchflower (2019), comparisons based solely on employment status can be misleading when working-time arrangements differ substantially across labor markets.

This paper takes a different approach by extending the efficient-unemployment framework of Michaillat and Saez (2021a). The original framework uses the official unemployment rate, U-3, which measures total unemployed persons as a share of the civilian labor force. While U-3 is the standard measure of unemployment,

it treats all employed workers as equivalent and therefore does not distinguish between full-time and part-time employment. As a result, two labor markets with the same unemployment rate may differ substantially in their effective labor input if one has a larger share of part-time employment.

One possible alternative is U-6, which includes not only unemployed workers but also marginally attached workers and persons employed part time for economic reasons (U.S. Bureau of Labor Statistics, 2026). However, U-6 remains a worker-count measure and does not account for differences in hours worked across employed workers. Consequently, simply replacing U-3 with U-6 does not provide a welfare-based measure of labor-market efficiency.

The distinction between full-time and part-time employment is particularly important from a welfare perspective. Part-time workers contribute fewer market hours on average than full-time workers and therefore generate less market production. For a given unemployment rate, an increase in part-time employment reduces effective labor input and may alter firms' incentives to post vacancies. Thus, labor markets with identical unemployment rates may differ in their degree of efficiency depending on the composition of employment.

To capture this dimension of labor-market heterogeneity, I extend the sufficient-statistics framework of Michaillat and Saez (2021a) by decomposing employment into full-time and part-time workers and incorporating differences in working hours directly into the welfare analysis. The extension introduces two additional sufficient statistics: the employment composition between full-time and part-time workers and the relative hours worked by part-time workers. As a result, the efficient unemployment rate and the unemployment gap depend not only on the number of employed workers but also on the effective labor input they provide. This approach recognizes that two economies with identical unemployment rates

may differ substantially in labor-market efficiency if their employment structures and working-time arrangements differ.

I first extend the U.S. analysis of Michaillat and Saez (2021a) through 2026, thereby incorporating the COVID-19 period. Because the U.S. Beveridge curve shifts unusually during and after the pandemic, I present the results separately for the pre-pandemic and post-pandemic periods. The results show that incorporating involuntary part-time employment has only a modest effect on estimates of efficient unemployment, whereas incorporating total part-time employment leads to somewhat larger changes in the measured unemployment gap and efficient unemployment rate. Nevertheless, the unemployment gap remains strongly countercyclical, and the labor market continues to exhibit periods of inefficiency. These findings suggest that the efficient-unemployment framework of Michaillat and Saez (2021a) is largely robust to the inclusion of involuntary part-time employment in the United States because workers employed part time for economic reasons account for only a small share of total labor input.

I then apply the extended framework to Japan, where part-time employment represents a substantially larger share of total employment. In contrast to the United States, incorporating total part-time employment has quantitatively important implications for the measurement of labor-market efficiency. Accounting for heterogeneity in working hours raises the efficient unemployment rate and materially alters the assessment of labor-market conditions. By contrast, incorporating involuntary part-time employment produces only small changes in the results. This finding reflects the fact that involuntary part-time employment constitutes a relatively small and declining share of total part-time employment in Japan. More broadly, the results indicate that the welfare implications of part-time work are driven primarily by differences in effective labor input rather than

by the involuntary nature of part-time employment.

1.3 Review of the Beveridgean unemployment gap

The Beveridgean unemployment-gap framework of Michaillat and Saez (2021a) measures labor-market efficiency using three sufficient statistics: the Beveridge elasticity, the social value of nonwork, and the recruiting cost. These statistics jointly determine the welfare-maximizing point on the Beveridge curve and yield closed-form expressions for efficient labor-market tightness and efficient unemployment, which serve as benchmarks for measuring unemployment gaps.

1.3.1 Three sufficient statistics in Michaillat and Saez (2021a)

The measurement of unemployment proposed by Michaillat and Saez (2021a) is grounded in the theory of efficiency within modern labor market models, which are governed by the Beveridge curve—a negative relationship between unemployment and vacancies. This implies that vacancies and unemployment cannot be simultaneously reduced, making the resolution of the unemployment-vacancy tradeoff, the elasticity of the Beveridge curve, central to their analysis.

As in Michaillat and Saez (2021a), the Beveridge elasticity is the elasticity of the vacancy rate with respect to the unemployment rate along the Beveridge curve, normalized to be positive:

$$\epsilon = -\frac{dln(v(u))}{dln(u)} = -\frac{u}{v} \cdot v'(u) \tag{1.1}$$

Both unemployment and vacancies induce welfare costs. Therefore, consider a

flow social welfare that is given by the linear function:

$$W(n, u, v) = p(n + zu - cv)L \tag{1.2}$$

where n is the employment rate, the number of employed workers divided by the size of the labor force. L denotes the total number of workers in the labor force. Unemployed workers engage in home production or other nonmarket activities valued at productivity zp, where $0 < z < 1$. Thus, z measures the social value of nonwork relative to market production.

A more general form can be expressed as:

$$\widehat{W}(u, v) = W(1 - u, u, v) \tag{1.3}$$

where $\widehat{W}$ is strictly decreasing in u and v.

Employed workers have a productivity level $p > 0$, which is greater than that of home production, where $pz < p$. Consequently, from the welfare analysis equation (1.2), the social value of nonwork z is defined as the marginal rate of substitution between unemployment and employment:

$$\frac{\partial W}{\partial u} / \frac{\partial W}{\partial n} = z < 1$$

Therefore, the social cost of unemployment is the marginal contribution of unemployed workers to welfare, relative to the marginal contribution of employed workers:

$$(\frac{\partial W}{\partial u} - \frac{\partial W}{\partial n}) / \frac{\partial W}{\partial n} = 1 - z > 0$$

The last sufficient statistic is the recruiting cost, which equals c in equation (1.2), representing the marginal rate of substitution between vacancies and employment in the welfare function:

$$-\frac{\partial W}{\partial v}/\frac{\partial W}{\partial n} = c > 0$$

According to Proposition 1 in Michaillat and Saez (2021a), efficiency is achieved at the point at which the Beveridge curve is tangent to an isowelfare curve; therefore, the efficient unemployment rate is implicitly defined by:

$$v'(u) = -\frac{\frac{\partial W}{\partial u} - \frac{\partial W}{\partial n}}{\frac{\partial W}{\partial v}} = -\frac{1-z}{c} \tag{1.4}$$

Note: Appendix A.1 shows the correspondence between the notation used in Michaillat and Saez (2021a) and the notation adopted in this paper, and derives the part-time extension.

A typical limitation for sufficient statistics as pointed out in Michaillat and Saez (2021a) is that the parameters may depend on each other (Chetty, 2009). Thus, a key assumption to address the endogeneity of the sufficient statistics as shown above to compute the efficient unemployment rate (Assumption 3 in Michaillat and Saez (2021a)):
The Beveridge elasticity ϵ, social value of nonwork z, and recruiting cost c do not depend on the unemployment and vacancy rates.

This assumption implies an isoelastic Beveridge curve, $v(u) = \hat{a} \cdot u^{-\epsilon}$, then an efficient tightness from the elasticity of Beveridge curve can be derived:

$$\theta^* = \frac{1-z}{c\epsilon} \tag{1.5}$$

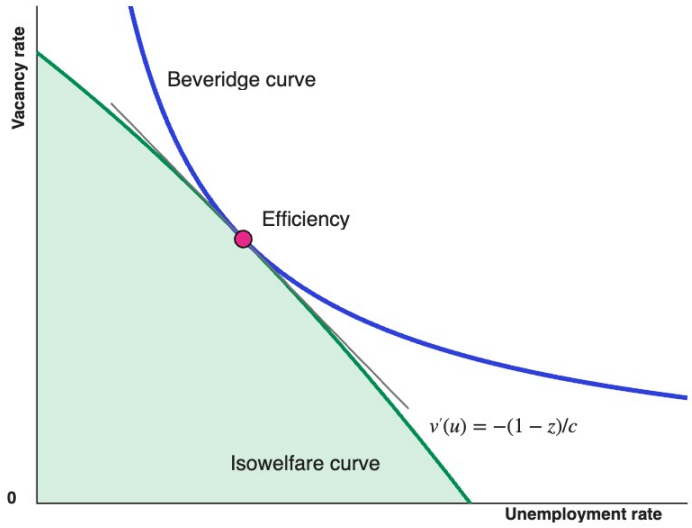


Figure 1.4: Efficient unemployment rate.

Notes: The Beveridge curve has slope $v'(u)$. The isowelfare curve has slope $-\frac{1-z}{c}$, where z is the social value of nonwork and c is the recruiting cost. The tangency of the Beveridge and isowelfare curves gives the efficient labor-market allocation. *Source:* Michaillat and Saez (2021a).

Hence, the efficient unemployment can be expressed in sufficient statistics:

$$u^* = \left(\frac{c\epsilon}{1-z} \cdot \frac{v}{u^{-\epsilon}}\right)^{1/(1+\epsilon)} \tag{1.6}$$

Figure 1.4 illustrates how the three sufficient statistics jointly determine labor-market efficiency in the Beveridgean framework of Michaillat and Saez (2021a). The Beveridge elasticity governs the slope of the Beveridge curve, while the social value of nonwork and recruiting cost determine the slope of the isowelfare curve. Labor-market efficiency is achieved at the tangency point between the Beveridge curve and an isowelfare curve. At this point, the marginal welfare gain from reducing unemployment is exactly offset by the marginal welfare cost of creating additional vacancies. This tangency condition forms the basis for the sufficient-statistics formula used to compute efficient labor-market tightness and efficient

unemployment.

1.4 Incorporating part-time employment

This section extends the Beveridgean unemployment-gap framework by distinguishing between full-time and part-time employment. In the standard framework, all employed workers are treated identically regardless of hours worked. Yet part-time workers supply fewer market hours than full-time workers and therefore contribute less market production. Ignoring this distinction may overstate effective labor utilization in economies where part-time employment represents a substantial share of total employment.

To account for differences in labor input, I decompose employment into full-time and part-time workers. Part-time workers are assumed to generate a combination of market and nonmarket production, implying a productivity level between that of full-time workers and unemployed workers. This extension introduces two additional sufficient statistics: the share of full-time employment in total employment, α, and the share of working hours supplied by part-time workers relative to full-time workers, γ. Together, these statistics allow the efficient-unemployment measure to account for both the composition of employment and differences in hours worked.

1.4.1 Flow social welfare with part-time employment

Following the social-welfare analysis of Michaillat and Saez (2021a), I decompose employment into full-time employment, n_F, and part-time employment, n_P, with unemployment rate u and vacancy rate v. Full-time workers have productivity $p > 0$. A part-time worker allocates a fraction γ of time to market production

and the remaining fraction $1-\gamma$ to nonmarket activities with productivity zp, where z denotes the social value of nonwork relative to market production. The effective productivity of a part-time worker is therefore

$$p\gamma + pz(1-\gamma) = p[\gamma + (1-\gamma)z].$$

Since part-time workers supply fewer market hours than full-time workers but more than unemployed workers engaged in home production,

$$pz < p[\gamma + (1-\gamma)z] < p.$$

The flow social welfare function becomes

$$W(n_F, n_P, u, v) = p\Big(n_F + [\gamma + (1-\gamma)z]n_P - cv + zu\Big)L. \tag{1.7}$$

Equation (1.7) extends the Beveridgean framework by allowing different employment states to contribute different amounts of effective labor input. As shown below, the extension introduces two additional sufficient statistics: the employment share of full-time workers, α, and the relative hours supplied by part-time workers, γ. In the model, γ denotes the fraction of time that part-time workers allocate to market production, with the remaining fraction $1-\gamma$ allocated to nonmarket production valued at zp. In the empirical implementation, I measure this fraction using the ratio of average weekly hours worked by part-time and full-time workers. This mapping assumes that differences in weekly hours reflect differences in market time relative to full-time workers. Under this assumption, the remaining time of part-time workers is valued consistently with the social value of nonwork used elsewhere in the model.

1.4.2 Two additional sufficient statistics

The extended framework requires information on both the composition of employment and the relative labor input supplied by part-time workers. These dimensions are summarized by two additional sufficient statistics, α and γ.

Population share of full-time and part-time employment

Let α^P denote the share of full-time employment in total employment:

$$\alpha^P = \frac{n_F}{n_F + n_P}, \tag{1.8}$$

where n_F and n_P denote the employment rates of full-time and part-time workers, respectively. The share of part-time employment is therefore

$$1 - \alpha^P = \frac{n_P}{n_F + n_P}. \tag{1.9}$$

The aggregate unemployment rate satisfies

$$u = 1 - (n_F + n_P). \tag{1.10}$$

In the baseline analysis, I classify workers as part-time if their actual hours worked (ahrsworkt) are fewer than 35 hours per week and as full-time if their actual hours worked are at least 35 hours per week. The 35-hour threshold is consistent with the conventional distinction between full-time and part-time employment used by the U.S. Bureau of Labor Statistics (U.S. Bureau of Labor Statistics, 2024a), although the BLS applies the threshold to usual rather than actual hours worked. I use actual hours in the baseline to avoid known comparability problems in usual-hours (uhrswork1) reporting in the ASEC microdata (notably in 1998) as shown

in Appendix Figure A.1, the part-time employment share constructed from usual hours exhibits a large spike in 1998 that is absent from the actual-hours series and difficult to reconcile with broader labor-market developments. Outside this episode, the two measures generate very similar part-time employment shares. Using actual hours therefore avoids spurious variation in the classification of workers while maintaining consistency with the construction of the relative-hours statistic γ.

As a robustness check, I also construct full-time and part-time employment shares using usual weekly hours, excluding the anomalous 1998 observation. The resulting employment shares are close to those obtained from actual hours (for example, mean α is 0.75 under actual hours and 0.78 under usual hours over 1994–2025), and the efficient unemployment rate and unemployment gap change by at most 0.13 percentage point relative to the baseline. Appendix Figure A.2 demonstrates that the choice between actual and usual hours has little effect on the results. Panels (a) and (b) show that efficient unemployment rates constructed from the two measures are almost indistinguishable in both the pre-COVID and post-COVID samples. Likewise, panels (c) and (d) show that unemployment gaps are virtually identical under the two classifications. Therefore, the use of actual hours worked in the baseline analysis is driven by data consistency rather than by any material effect on the estimated efficient unemployment rate or unemployment gap.

A rise in voluntary part-time employment need not indicate greater labor-market slack. Nevertheless, even voluntarily chosen part-time work reduces market hours supplied and therefore affects effective labor utilization. Consequently, two economies with identical unemployment rates may differ substantially in aggregate labor input and market production. The purpose of the extension is not

to classify voluntary part-time work as inefficient, but rather to account for differences in labor input when measuring efficient unemployment.

Total part-time employment includes both voluntary and involuntary part-time workers. The involuntary group is especially relevant for measuring labor-market slack because these workers are employed part time for economic reasons and would prefer to work more hours at the prevailing wage (U.S. Bureau of Labor Statistics, 2024b). I therefore also examine involuntary part-time employment separately from total part-time employment to distinguish between the role of part-time employment more broadly and the role of underemployment.

Let n_{IP} denote involuntary part-time employment. When focusing on this margin, the full-time employment share is defined as

$$\alpha^{IP} = \frac{n_F}{n_F + n_{IP}}, \tag{1.11}$$

where n_F denotes full-time employment. The corresponding involuntary part-time employment share is

$$1 - \alpha^{IP} = \frac{n_{IP}}{n_F + n_{IP}}. \tag{1.12}$$

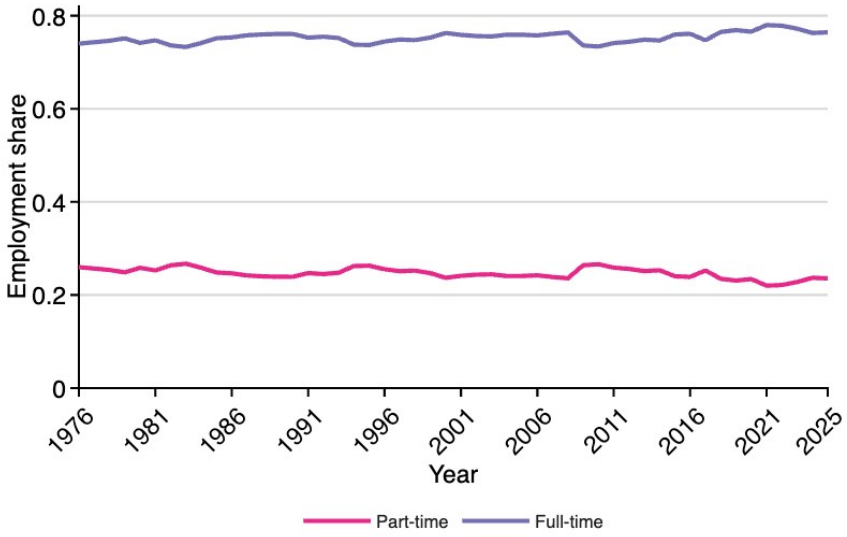


Figure 1.5: Share of full-time and part-time employment in the United States, 1976–2025.

Notes: The figure reports the share of full-time employment, α^P, and the share of part-time employment, $1 - \alpha^P$, in total employment. Full-time workers are defined as employed individuals with actual weekly hours worked of at least 35 (`ahrsworkt` ≥ 35); part-time workers are defined as employed individuals with actual weekly hours worked of fewer than 35 (`ahrsworkt` < 35). The sample includes employed workers (`EMPSTAT` = 10, 12) with valid actual hours and positive person weights. Employment shares are computed using `ASECWT`, with the 2014 ASEC split-sample adjustment (`ASECWT` $\times 5/8$ if `HFLAG`=0; $\times 3/8$ if `HFLAG`=1). Population counts are $n_t^F = \sum_i w_{it} \mathbf{1}\{\text{FT}\}$ and $n_t^P = \sum_i w_{it} \mathbf{1}\{\text{PT}\}$; $\alpha_t^P = n_t^F/(n_t^F + n_t^P)$. *Source:* IPUMS CPS ASEC.

Figure 1.5 plots the evolution of full-time and total part-time employment shares in the United States from 1976 to 2025. The share of part-time employment fluctuates over time but remains within a relatively narrow range. Table 1.1 shows that full-time employment accounts for approximately 75 % of total employment on average, while part-time employment accounts for the remaining 25 %. This stability motivates calibrating the employment-share parameter at its sample average, $\alpha = 0.75$.

For involuntary part-time and full-time employment, the overall composition of employment remains remarkably stable from 1988 to 2025, as shown in Figure 1.6. Table 1.2 confirms this pattern quantitatively: full-time employment accounts

for approximately 93 % of the relevant employment measure on average, while involuntary part-time employment accounts for the remaining 7 %.

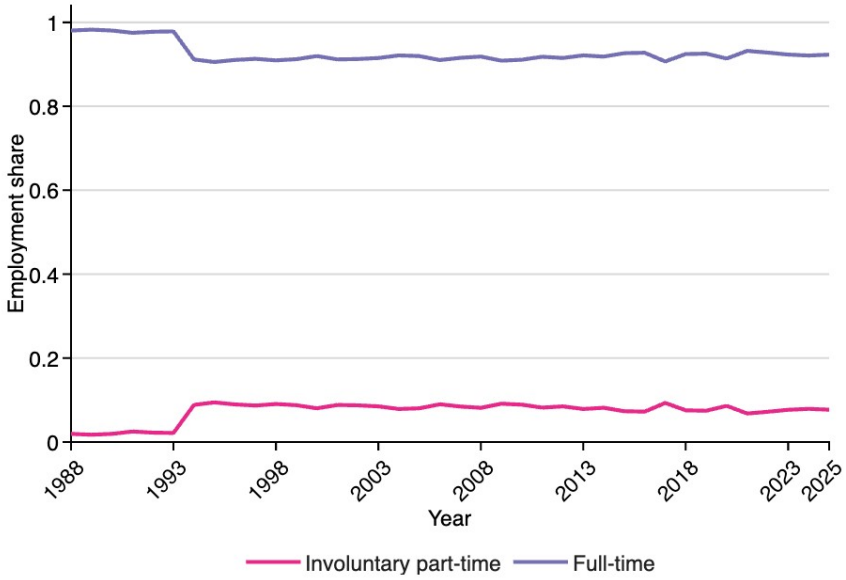


Figure 1.6: Share of full-time and involuntary part-time employment in the United States, 1988–2025.

Notes: The figure reports the share of full-time employment, α^{IP}, and the share of involuntary part-time employment, $1 - \alpha^{IP}$, among full-time and involuntary part-time workers. Full-time workers are defined as employed individuals with actual weekly hours worked of at least 35 (`ahrsworkt` ≥ 35). Involuntary part-time workers are defined as employed individuals with `WKSTAT` = 12 or 21, corresponding to part-time work for economic reasons. The sample includes employed workers (`EMPSTAT` = 10, 12) with positive person weights. Employment shares are computed using `ASECWT`, with the 2014 ASEC split-sample adjustment (`ASECWT` $\times 5/8$ if `HFLAG`=0; $\times 3/8$ if `HFLAG`=1). Population counts are $n_t^F = \sum_i w_{it} \mathbf{1}\{\text{FT}\}$ and $n_t^{IP} = \sum_i w_{it} \mathbf{1}\{\text{IPT}\}$; $\alpha_t^{IP} = n_t^F / (n_t^F + n_t^{IP})$. The series begins in 1988, when `WKSTAT` codes for involuntary part-time employment are available. *Source:* IPUMS CPS ASEC.

For the historical periods in which direct measures are unavailable, I use the sample-average employment shares as calibration values. For total part-time employment, the average full-time employment share is

$$\alpha^P = 0.75, \tag{1.13}$$

implying that approximately three-quarters of employed workers are full-time

workers and one-quarter are part-time workers. This value is used to calibrate the pre-1976 historical period. For the period in which CPS ASEC data are available, beginning in 1976, I incorporate the observed time-series values of α_t.

For involuntary part-time employment, the corresponding average full-time employment share is

$$\alpha^{IP} = 0.93, \tag{1.14}$$

implying that approximately 93 % of workers are employed full time, while 7 % are employed part time for economic reasons. Since the involuntary part-time series is available beginning in 1988, I use $\alpha^{IP} = 0.93$ to calibrate earlier historical periods and use the observed time-series values of α_t^{IP} after 1988.

Table 1.1: Summary Statistics of Full-Time and Total Part-Time Employment Shares

	Full-time (n^F)	Part-time (n^P)	α^P	$1-\alpha^P$
Mean	93.75 million	30.52 million	0.753	0.247
SD	16.70 million	4.55 million	0.011	0.011
Min	61.01 million	21.42 million	0.733	0.220
Max	120.82 million	37.25 million	0.780	0.267
N	50	50	50	50

Notes: Annual employment counts from the CPS Annual Social and Economic Supplement (ASEC), 1976–2025 ($N = 50$). The sample includes employed workers (`empstat` = 10, 12) with positive person weights (`asecwt`). In 2014, weights are adjusted for the ASEC split sample (`asecwt` × 5/8 if `hflag`=0; × 3/8 if `hflag`=1). Full-time is defined by actual weekly hours worked ≥ 35 (`ahrsworkt`); total part-time by actual hours < 35 (workers with valid actual hours only). Population counts are $n_t^F = \sum_i w_{it}\mathbf{1}\{\text{FT}\}$ and $n_t^P = \sum_i w_{it}\mathbf{1}\{\text{PT}\}$; $\alpha_t^P = n_t^F/(n_t^F + n_t^P)$ and $1-\alpha_t^P = n_t^P/(n_t^F + n_t^P)$. For years before 1976 in the quarterly figures, we set α to the sample mean reported above (0.75) and use the annual series α_t^P from 1976 onward.

Table 1.2: Summary Statistics of Full-Time and Involuntary Part-Time Employment Shares

	Full-time (n^F)	Invol. part-time (n^{IP})	α^{IP}	$1-\alpha^{IP}$
Mean	101.02 million	8.16 million	0.927	0.073
SD	11.56 million	2.78 million	0.023	0.023
Min	82.33 million	1.64 million	0.906	0.019
Max	120.82 million	11.24 million	0.981	0.094
N	38	38	38	38

Notes: Same CPS ASEC sample and weighting as Table 1.1, 1988–2025 ($N = 38$). Full-time: actual weekly hours worked ≥ 35 (`ahrsworkt`). Involuntary part-time (IPT): `wkstat` = 12 or 21 (at work part-time for economic reasons, or usually part-time for economic reasons); no additional hours restriction. Counts: n_t^F, n_t^{IP}; $\alpha_t^{IP} = n_t^F/(n_t^F + n_t^{IP})$. The IPT series is meaningful from 1988 when `wkstat` is available; before 1988 we calibrate α^{IP} to the sample mean above (0.93) in the quarterly figures and use observed α_t^{IP} thereafter.

Share of working hours

The second sufficient statistic is the relative-hours contribution of part-time workers. I construct two versions of this statistic. The first uses total part-time employment, while the second uses involuntary part-time employment. The distinction is useful because total part-time employment captures the aggregate-hours contribution of part-time work, whereas involuntary part-time employment isolates labor underutilization among workers who would like to work more hours. For total part-time employment, I define the relative-hours parameter as

$$\gamma_t^P = \frac{H_t^P}{H_t^F}, \tag{1.15}$$

where H_t^P and H_t^F denote the average weekly hours worked by part-time and full-time workers, respectively. Table 1.3 shows that part-time workers work 20.92 hours per week on average, while full-time workers work 44.26 hours. Figure 1.7 plots the corresponding time series of average weekly hours for full-time and part-time workers from 1976 to 2025. The implied relative-hours ratio averages 0.47

and is highly stable over time, with a standard deviation of only 0.012. Therefore, for the historical period in which annual hours data are not available, I use the sample average $\gamma^P = 0.47$ as the calibration value. For the period in which the data are available, I use the observed time-series value γ_t^P.

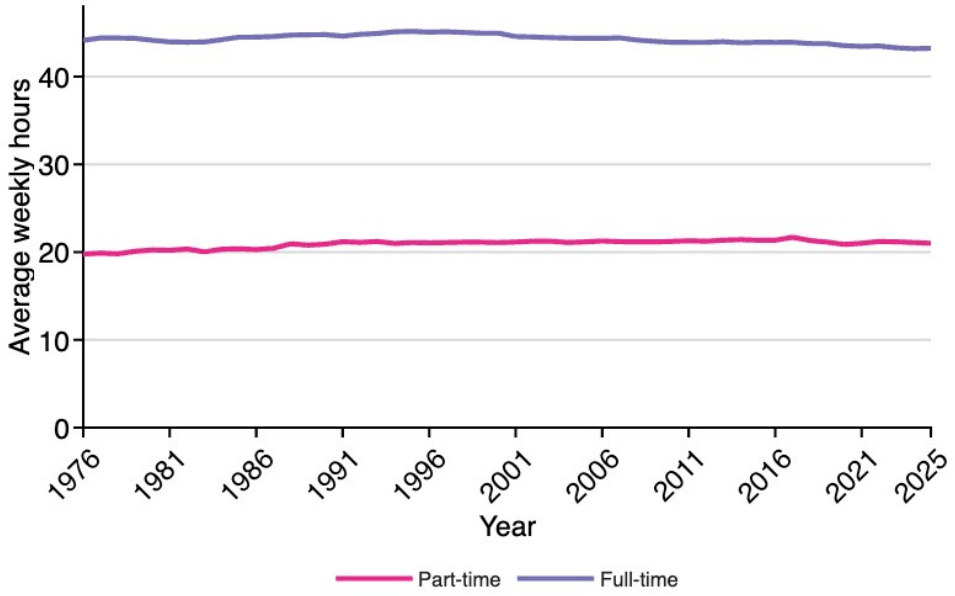


Figure 1.7: Average weekly hours worked by full-time and part-time workers in the United States, 1976–2025.

Notes: The figure reports annual average actual weekly hours worked by full-time and part-time workers. Workers are classified using actual hours worked last week (`ahrsworkt`): full-time if `ahrsworkt` ≥ 35, and part-time if `ahrsworkt` < 35, using valid actual hours only. The sample includes employed workers (`EMPSTAT` = 10, 12) with positive person weights. Averages are population-weighted means of `ahrsworkt` among full-time and part-time workers, using `ASECWT` with the 2014 ASEC split-sample adjustment (`ASECWT` $\times 5/8$ if `HFLAG`=0; $\times 3/8$ if `HFLAG`=1). *Source:* IPUMS CPS ASEC.

For involuntary part-time employment, I define

$$\gamma_t^{IP} = \frac{H_t^{IP}}{H_t^F}, \tag{1.16}$$

where H_t^{IP} denotes the average weekly hours worked by involuntary part-time workers. Figure 1.8 reports average weekly hours for full-time and involuntary

part-time workers in the United States from 1988 to 2025. Full-time workers work approximately 44 hours per week on average, while involuntary part-time workers work approximately 25 hours per week.

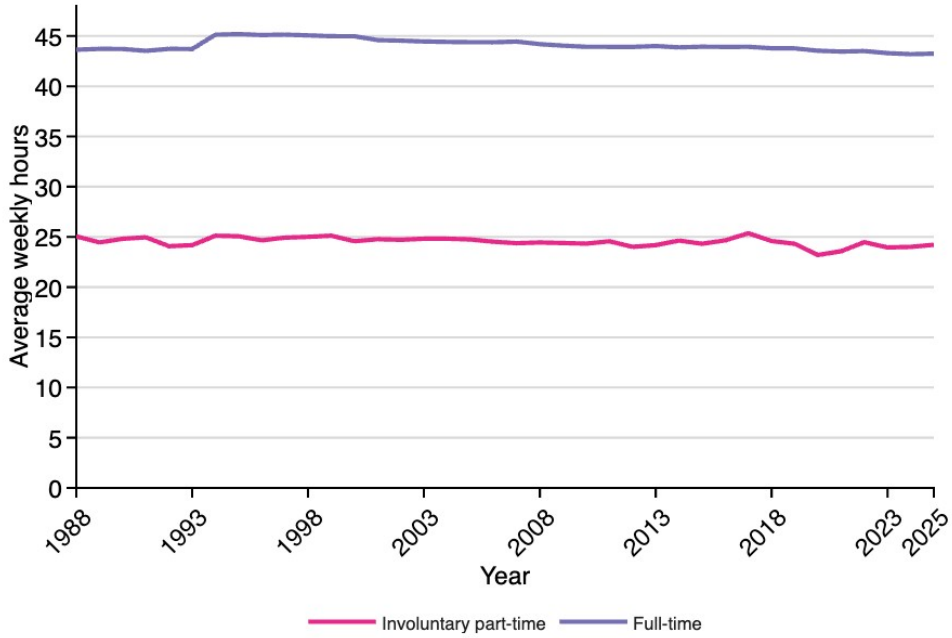


Figure 1.8: Average weekly hours worked by full-time and involuntary part-time workers in the United States, 1988–2025.

Notes: The figure reports annual average actual weekly hours worked by full-time and involuntary part-time workers. Full-time workers are defined as employed individuals with actual weekly hours worked of at least 35 (`ahrsworkt` ≥ 35). Involuntary part-time workers are defined as employed individuals with `WKSTAT` = 12 or 21, corresponding to part-time work for economic reasons. The sample includes employed workers (`EMPSTAT` = 10, 12) with valid actual hours and positive person weights. Averages are population-weighted means of `ahrsworkt` among full-time and involuntary part-time workers, using `ASECWT` with the 2014 ASEC split-sample adjustment (`ASECWT` $\times 5/8$ if `HFLAG`=0; $\times 3/8$ if `HFLAG`=1). The series begins in 1988, when `WKSTAT` codes for involuntary part-time employment are available.
Source: IPUMS CPS ASEC.

As shown in Table 1.4 and Figure 1.8, the relative-hours ratio for involuntary part-time workers averages 0.56, with a standard deviation of only 0.008. This implies that involuntary part-time workers supply approximately 56 % as many market hours as full-time workers. As with total part-time employment, I use the

sample average $\gamma^{IP} = 0.56$ for the historical period without available annual data and use the observed time-series value γ_t^{IP} for the period in which the data are available.

Table 1.3: Summary Statistics of Average Weekly Hours: Total Part-Time (γ^P)

	Full-time hours (H^F)	Part-time hours (H^P)	γ
Mean	44.26	20.92	0.473
SD	0.52	0.47	0.012
Min	43.18	19.77	0.446
Max	45.16	21.69	0.494
N	50	50	50

Notes: Annual averages from the CPS ASEC, 1976–2025 ($N = 50$), using the same sample and weights as Table 1.1. Workers are classified into full-time and total part-time by actual hours worked (`ahrsworkt` ≥ 35 vs. < 35). H_t^F and H_t^P are population-weighted means of actual weekly hours worked (`ahrsworkt`) among full-time and part-time workers, respectively (valid actual hours only). $\gamma_t = H_t^P / H_t^F$. For years before 1976 in the quarterly figures, we set γ^P to the sample mean (0.47) and use the annual series from 1976 onward.

Table 1.4: Summary Statistics of Average Weekly Hours: Involuntary Part-Time (γ^{IP})

	Full-time hours (H^F)	IPT hours (H^{IP})	γ^{IP}
Mean	44.27	24.52	0.554
SD	0.59	0.44	0.008
Min	43.18	23.19	0.533
Max	45.16	25.37	0.578
N	38	38	38

Notes: Same CPS ASEC sample and weighting as Table 1.2, 1988–2025 ($N = 38$). Full-time and IPT groups are defined as in Table 1.2. H_t^F and H_t^{IP} are population-weighted means of `ahrsworkt` among full-time and IPT workers (valid actual hours only). $\gamma_t^{IP} = H_t^{IP} / H_t^F$. For years before 1988 in the quarterly figures, we set γ^{IP} to the sample mean (0.56) and use the annual series from 1988 onward.

The average relative-hours ratio is 0.47 for total part-time workers and 0.56 for involuntary part-time workers. Although involuntary part-time workers supply somewhat more hours than the average part-time worker, the quantitative difference is relatively small. I nevertheless distinguish between the two measures

for consistency with the corresponding employment definitions. In the Japanese application, however, separate hours data for involuntary part-time workers are unavailable. Therefore, I assume that the relative-hours ratio for involuntary part-time workers is equal to that of total part-time workers.

Proposition 1: Consider a point on the Beveridge curve with the share of full-time workers α, recruiting cost c, Beveridge elasticity ϵ, share of time γ and the marginal rate of substitution between unemployment and employment z, tightness θ is inefficiently high if $\theta > \frac{\alpha+(1-\alpha)\cdot(\gamma+(1-\gamma)\cdot z)-z}{c\epsilon}$, inefficiently low if $\theta < \frac{\alpha+(1-\alpha)\cdot(\gamma+(1-\gamma)\cdot z)-z}{c\epsilon}$, efficient if

$$\theta^* = \frac{\alpha + (1-\alpha)\cdot[\gamma + (1-\gamma)\cdot z] - z}{c\epsilon} \tag{1.17}$$

The derivation treats the employment composition α and relative-hours parameter γ as fixed at the point where efficiency is evaluated. The tangency condition therefore characterizes efficient labor-market tightness for a given employment composition and hours structure. In the empirical implementation, α_t and γ_t are allowed to vary over time and enter the sufficient-statistics formula quarter by quarter.

Proof. Efficiency is achieved where the Beveridge curve is tangent to an isowelfare curve (Michaillat and Saez, 2021a). To locate the tangency point, we first need the slope of the isowelfare curve, which is minus the marginal rate of substitution between unemployment and vacancies in the welfare function. The slope of the

Beveridge curve is $v'(u)$. Therefore,

$$\begin{aligned}
-v'(u) &= -\frac{\partial W/\partial u}{\partial W/\partial v} \\
&= \frac{-\partial W/\partial u - \partial W/\partial(n_F + n_P)}{\partial W/\partial v} \\
&= \frac{-\partial W/\partial u + \alpha \cdot \partial W/\partial n_F + (1-\alpha) \cdot \partial W/\partial n_P}{\partial W/\partial v} \\
&= \frac{\frac{-\partial W/\partial u + \alpha \cdot \partial W/\partial n_F + (1-\alpha) \cdot \partial W/\partial n_P}{\partial W/\partial n_F}}{(\partial W/\partial v)/(\partial W/\partial n_F)} \\
&= \frac{\alpha + (1-\alpha) \cdot [\gamma + (1-\gamma)z] - z}{c}.
\end{aligned}$$

Thus,

$$-v'(u) = \frac{\alpha + (1-\alpha) \cdot [\gamma + (1-\gamma)z] - z}{c}. \tag{1.18}$$

The Beveridge elasticity satisfies $\epsilon\theta = -v'(u)$, where $\theta \equiv v/u$ denotes labor-market tightness. Thus, the efficient tightness is

$$\theta^* = \frac{\alpha + (1-\alpha) \cdot [\gamma + (1-\gamma)z] - z}{c\epsilon}. \tag{1.19}$$

The Beveridge curve is isoelastic. Therefore, the efficient unemployment rate can be expressed as

$$u^* = \left[\frac{c\epsilon}{\alpha + (1-\alpha) \cdot [\gamma + (1-\gamma)z] - z} \cdot \frac{v}{u^{-\epsilon}} \right]^{\frac{1}{1+\epsilon}}. \tag{1.20}$$

□

If $\alpha = 1$, all employed workers are full-time workers. The part-time margin is then shut down, and the model collapses to the baseline functional form in Michaillat and Saez (2021a).

Figure 1.9 illustrates the role of part-time employment in determining the

efficient unemployment rate. Once part-time employment is introduced, the isowelfare curve becomes flatter. The reason is that part-time jobs contribute fewer effective hours than full-time jobs, since $\gamma < 1$. Therefore, an additional employed worker generates less output when a larger share of employment is part-time. This reduces the social gain from pushing unemployment lower through additional vacancy creation.

As a result, the efficient allocation features lower labor-market tightness and a higher efficient unemployment rate. In the figure, the efficient point shifts slightly to the right along the Beveridge curve. Part-time employment can therefore make the observed unemployment rate appear lower by moving workers into employment, while still leaving the economy with lower effective labor input. This raises the efficient unemployment rate relative to the baseline model that treats all employment as full-time. This mechanism is central to the paper: accounting for part-time employment changes the benchmark against which labor-market slack should be measured.

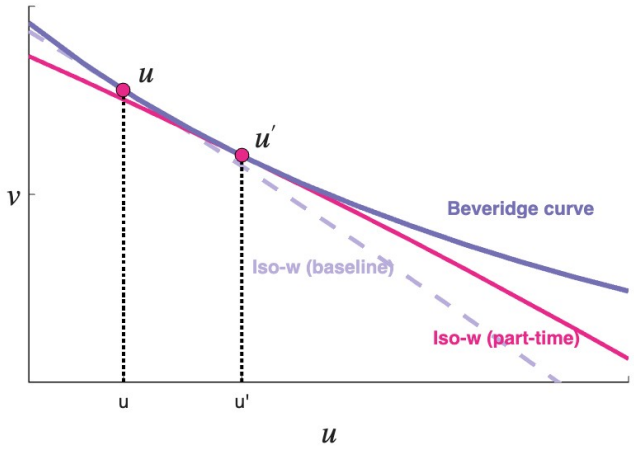


Figure 1.9: Efficient unemployment rate with and without part-time employment.

Notes: The dashed line shows the baseline isowelfare curve from Michaillat and Saez (2021a), while the solid line shows the isowelfare curve after incorporating part-time employment. Efficient unemployment is determined by the tangency between the Beveridge curve and the isowelfare curve. Accounting for part-time employment changes the slope of the isowelfare curve and shifts the efficient unemployment rate from u to u'.

1.4.3 Estimation of remaining sufficient statistics

Following Michaillat and Saez (2021a), I estimate the remaining sufficient statistics required to compute the efficient unemployment rate. These include the Beveridge elasticity, the social value of nonwork, and recruiting costs.

By regressing the log vacancy rate (using the vacancy proxy constructed by Barnichon (2010)) on the log unemployment rate (U.S. Bureau of Labor Statistics) from 1951Q1 to 2026Q1, I estimate the Beveridge elasticity in the United States. The statistical model has m breaks and $m+1$ regimes:

$$\ln v(t) = \ln \hat{a}_j - \epsilon_j \ln u(t) + z(t), \qquad t = T_{j-1}+1, \dots, T_j, \tag{1.21}$$

for $j = 1, \dots, m+1$. The observed independent variable is the logarithm of the unemployment rate, $\ln u(t)$, and the dependent variable is the logarithm of the

vacancy rate, $\ln v(t)$. The m break dates are denoted by $T_1, \ldots, T_m$; in this case, $T_0 = 0$ and $T_{m+1} = 300$. The parameter ϵ_j is the Beveridge elasticity in regime j, and $\hat{a}_j$ is the intercept. These parameters, together with the break dates, are jointly estimated using the Bai–Perron algorithm of Bai and Perron (1998, 2003), which first determines the number of structural breaks m and then estimates the parameters and break dates.

I follow the Bai–Perron setup of Michaillat and Saez (2021a), allowing for different variances of the errors across regimes and for autocorrelation in the errors. To obtain standard errors robust to autocorrelation and heteroskedasticity, the algorithm uses a quadratic kernel with automatic bandwidth selection based on an AR(1) approximation (Andrews, 1991). This setup also allows for different distributions of the independent and dependent variables across regimes. As required by the Bai–Perron procedure (Bai and Perron, 2003), I set the maximum number of breaks to three and the trimming parameter to 0.20, so that each regime contains at least 20% of the sample (60 quarters).

The next step is to determine the number of structural breaks. To examine whether structural breaks are present, the algorithm runs supF tests of no structural break versus m breaks for $m = 1, 2, 3$. The tests reject the null hypothesis of no break at the 1% significance level. I then use two information criteria—the Bayesian information criterion (Yao, 1988) and the modified Schwarz criterion (Liu et al., 1997)—to select the number of breaks. Both criteria select three breaks.

The estimated break dates occur in 1965Q4, 1990Q3, and 2011Q1, yielding four Beveridge-curve regimes: 1951Q1–1965Q4, 1966Q1–1990Q2, 1990Q3–2011Q1, and 2011Q2–2026Q1. Table 1.5 reports the resulting Beveridge elasticity estimates and corrected standard errors. The elasticity estimates range from 0.48

to 0.79 and average 0.68 over 1951–2026; the corrected standard errors range from 0.03 to 0.19. The overall fit is $R^2 = 0.77$. Figure 1.10 plots the estimated Beveridge curve within each regime.

The post-2020 period displays a distinct outward shift of the Beveridge curve, reflecting unusual labor-market disruptions associated with the COVID-19 pandemic. These disruptions include health concerns, changes in work preferences, and changes in the efficiency of job matching (Barlevy et al., 2023). Because the pandemic period represents a sharp departure from the historical Beveridge relationship, I present the main results separately for the pre-pandemic period (1951–2019) and the post-pandemic period (2020–2026). In the pre-pandemic sample, I estimate the Beveridge curve with the Bai–Perron procedure, setting the maximum number of breaks to five (with trimming parameter 0.15), as in Michaillat and Saez (2021a). In the post-pandemic subsample, the Bai–Perron break procedure is not reliable given the short sample, so I instead estimate the Beveridge elasticity using a simple OLS regression.

Table 1.5: Beveridge elasticity estimates by regime, 1951Q1–2026Q1

Regime	$\hat{\epsilon}_j$	SE
1951Q1–1965Q4	0.73	0.03
1966Q1–1990Q2	0.48	0.07
1990Q3–2011Q1	0.79	0.03
2011Q2–2026Q1	0.71	0.19

Notes: Estimates are obtained from the Bai–Perron procedure with a maximum of three breaks and trimming parameter 0.20. Standard errors are corrected for autocorrelation and heteroskedasticity.

Second, I set the social value of nonwork following Michaillat and Saez (2021a). They measure this statistic from revealed-preference estimates in two studies. Using administrative data from the U.S. Army, Borgschulte and Martorell (2018) use soldiers' reenlistment decisions as a natural experiment to estimate the cost of un-

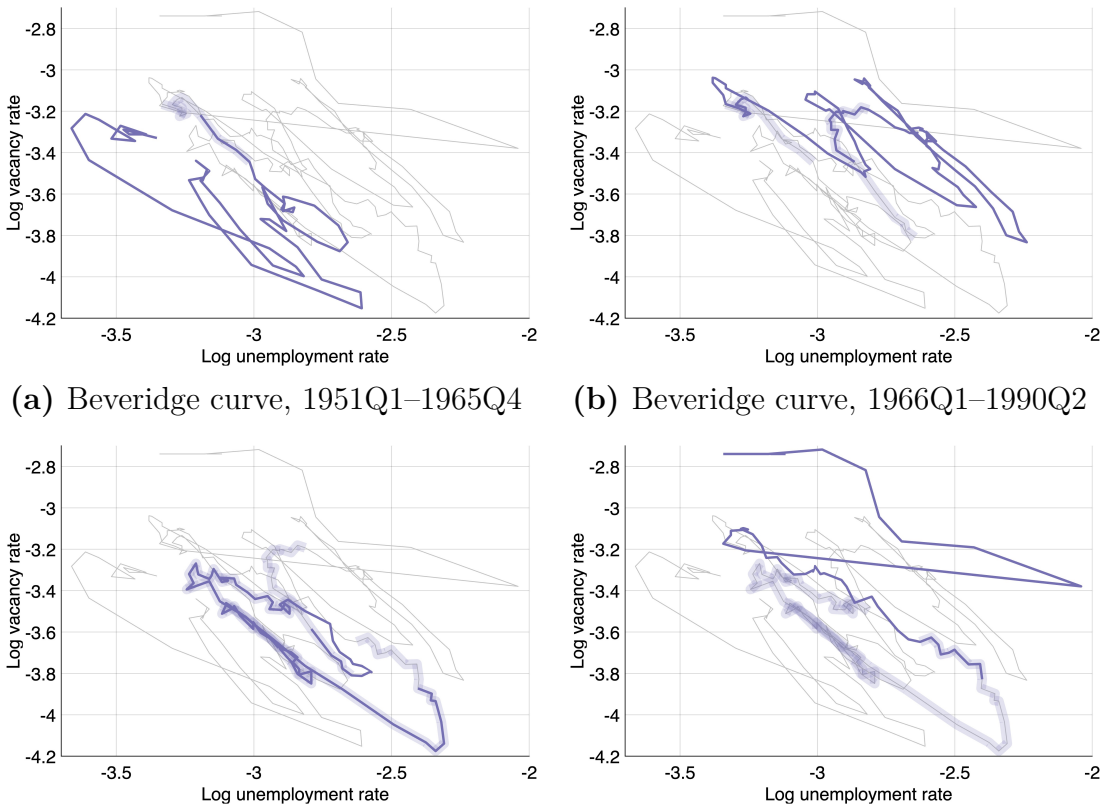
employment and find that home production and recreation during unemployment offset between 13% and 35% of earnings losses. In a separate field experiment, Mas and Pallais (2017) estimate that the value of nonwork time during unemployment is worth 58% of predicted earnings. Michaillat and Saez (2021a) translate these estimates into social values of nonwork by adjusting predicted earnings to the marginal product of labor and subtracting the value of public benefits received during unemployment (7% of the marginal product of labor). This yields a plausible range of 0.03–0.49. Michaillat and Saez (2021a) set the statistic to its midrange value, $z = 0.26$, and I adopt the same calibration.

Third, I set the recruiting cost following Michaillat and Saez (2021a), who measure it from the 1997 National Employer Survey conducted by the Bureau of the Census, as in Villena-Roldán (2010). The survey reports that establishments devote 3.2% of labor costs to recruiting. Adjusting for the unemployment rate in 1997, Michaillat and Saez (2021a) obtain a recruiting cost of $c = 0.92$. Because there is no comprehensive time series of recruiting effort in the United States, they assume that the recruiting cost remains constant at its 1997 value, as is standard in matching models. I adopt the same value and assumption.

Together with the part-time employment statistics estimated above, these sufficient statistics allow me to compute efficient labor-market tightness, the efficient unemployment rate, and the unemployment gap under the extended framework.

1.5 Application to the United States

I first extend the baseline Beveridgean unemployment-gap results of Michaillat and Saez (2021a) through 2026. Because the Beveridge curve displays unusual movements during and after the COVID-19 pandemic, I report the results separately for the pre-pandemic and post-pandemic periods.



(a) Beveridge curve, 1951Q1–1965Q4 **(b)** Beveridge curve, 1966Q1–1990Q2

(c) Beveridge curve, 1990Q3–2011Q1 **(d)** Beveridge curve, 2011Q2–2026Q1

Figure 1.10: Beveridge curve estimation in the United States, 1951–2026.

Notes: The figure plots the Beveridge curve separately for four labor-market regimes identified using the structural break procedure of Bai and Perron (1998, 2003). The estimated regimes are 1951Q1–1965Q4, 1966Q1–1990Q2, 1990Q3–2011Q1, and 2011Q2–2026Q1.

I then incorporate part-time employment using two calibrations. The first uses total part-time employment, which includes both voluntary and involuntary part-time workers. Under this calibration, the full-time employment share is $\alpha^P = 0.75$ and the relative-hours parameter is $\gamma^P = 0.47$. The second uses involuntary part-time employment, defined as workers employed part time for economic reasons. Under this calibration, the corresponding values are $\alpha^{IP} = 0.93$ and $\gamma^{IP} = 0.56$. Comparing the two calibrations allows me to distinguish the aggregate-hours effect of part-time employment from the underemployment effect.

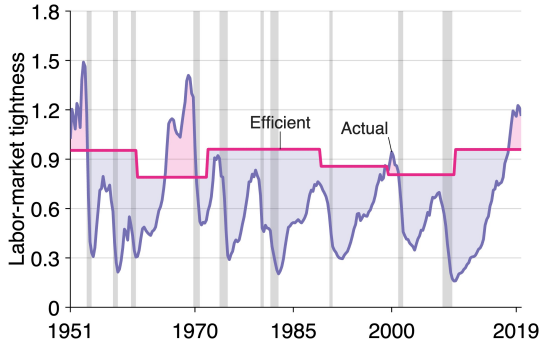
1.5.1 Baseline results

Figure 1.11 reports the baseline efficient-unemployment results for the United States through 2026. The figure separates the pre-pandemic and post-pandemic periods because the Beveridge curve displays unusual movements during and after the COVID-19 pandemic.

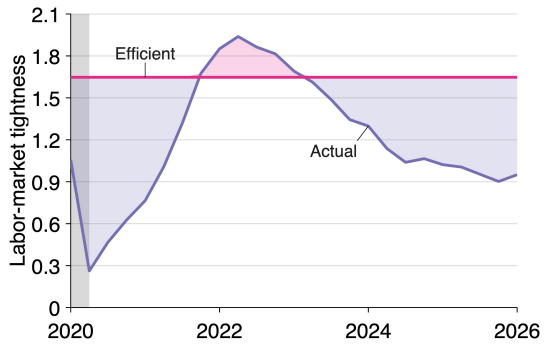
Panels A and B plot actual and efficient labor-market tightness. Before the pandemic, actual tightness rose during several well-known episodes, including the Korean War period, the late-1960s peak of the Vietnam War, and the dot-com boom of 1999–2000. After the pandemic, actual tightness increased sharply in 2021–2022 before easing in 2023. Because the post-pandemic sample is short, the Bai–Perron structural-break procedure is not applicable to this subsample. I therefore estimate the post-pandemic Beveridge elasticity using a simple OLS regression, which yields an elasticity of 1.65. This estimate serves as the post-pandemic benchmark.

Panels C and D report the efficient unemployment rate. Before the pandemic, the efficient unemployment rate averages 4.3% and ranges from 3.0% to 5.4%. After the pandemic, the efficient unemployment rate averages 3.5% and ranges

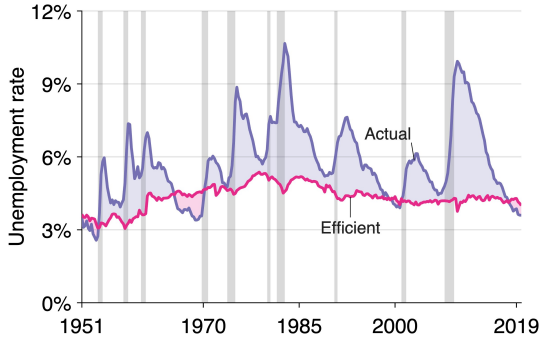
from 2.9% to 4.4%, while the actual unemployment rate averages 4.8%. Thus, under the baseline framework that treats employment as homogeneous and does not account for part-time employment, the United States generally exhibits a positive unemployment gap. This suggests that labor-market slack remains present even before incorporating part-time employment into the framework.



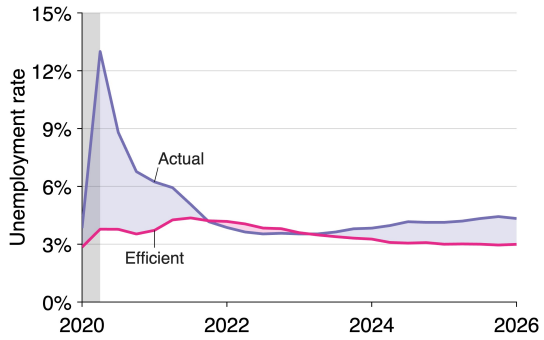
(a) Efficient labor-market tightness, 1951–2019



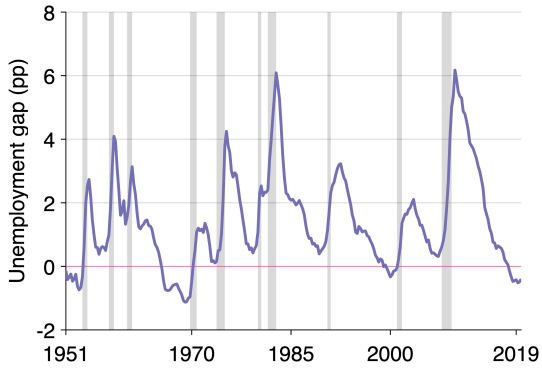
(b) Efficient labor-market tightness, 2020–2026



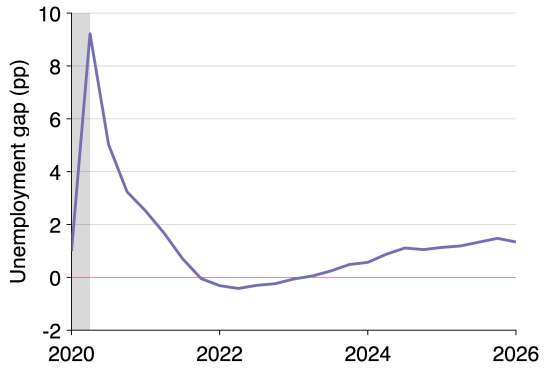
(c) Efficient unemployment rate, 1951–2019



(d) Efficient unemployment rate, 2020–2026



(e) Unemployment gap, 1951–2019



(f) Unemployment gap, 2020–2026

Figure 1.11: Efficient labor-market tightness, efficient unemployment, and unemployment gap in the United States under the full-time baseline, 1951–2026.

Notes: The figure reports the baseline results following Michaillat and Saez (2021a), in which employment is treated as homogeneous and all employed workers are effectively treated as full-time workers. The sample is split between the pre-pandemic period, 1951–2019, and the post-pandemic period, 2020–2026. These results provide the benchmark against which the part-time employment extensions are compared.

1.5.2 Incorporation of part-time employment

Total part-time employment

I now incorporate total part-time employment using the sufficient statistics estimated in Section 1.4. Under this extension, the calibration sets the full-time employment share to $\alpha^P = 0.75$ and the relative-hours parameter to $\gamma^P = 0.47$. The resulting measures remain broadly consistent with the full-time baseline, but several meaningful differences emerge. These differences are most visible in Panel A of Figure 1.12, where the circled episodes show actual labor-market tightness approaching or exceeding efficient labor-market tightness before several major downturns.

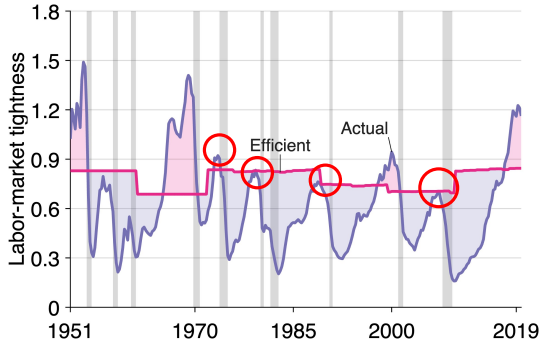
In particular, the extended framework identifies episodes in which actual labor-market tightness approaches or reaches the efficient allocation before several major recessions: the 1973–75 recession associated with the first oil shock, the early-1980s recessions associated with the Volcker disinflation period, the 1990–91 recession associated with the oil-price shock and credit tightening, and the 2007–09 Great Recession associated with the global financial crisis. In these episodes, accounting for part-time employment lowers the efficient-tightness benchmark, so that actual tightness converges toward—and in some cases effectively meets—the efficient allocation. Under the baseline framework that treats all employment as full-time, the same episodes instead appear inefficiently slack. The part-time extension therefore suggests that labor-market conditions were closer to the efficient allocation immediately before these downturns than implied by the full-time baseline.

Before COVID, actual labor-market tightness averages 0.62, while efficient tightness under the part-time calibration averages 0.78. After COVID, actual tightness averages 1.21, while efficient tightness averages 1.43. Thus, the extended

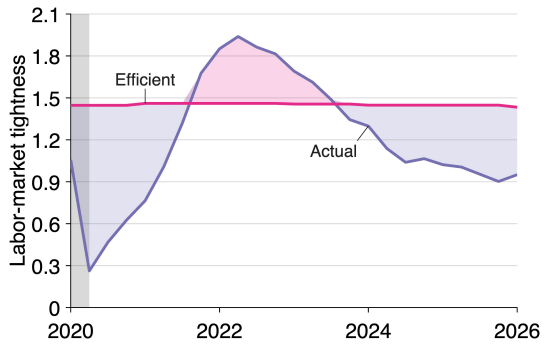
framework continues to indicate a generally slack labor market on average, but it also reveals several episodes in which actual tightness approaches the efficient level.

The corresponding efficient unemployment rate also shifts under the part-time calibration. Before COVID, actual unemployment averages 5.77%, while efficient unemployment averages 4.68%, with a range from 3.30% to 5.81%. After COVID, actual unemployment averages 4.82%, compared with an efficient unemployment rate of 3.86%, which ranges from 3.13% to 4.80%. These estimates imply a positive unemployment gap on average, but the gap narrows substantially during the highlighted pre-recession episodes.

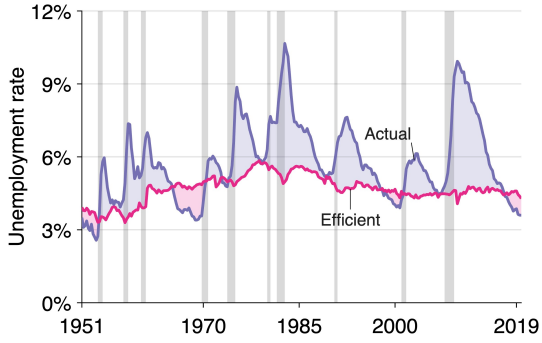
Figure 1.13 makes this comparison explicit by plotting the baseline results against the part-time-employment extension. Although the quantitative differences are modest, they are systematic. Incorporating part-time employment shifts the efficient benchmark by lowering efficient labor-market tightness and raising efficient unemployment. Consequently, the distance between actual and efficient labor-market conditions narrows during several pre-recession episodes, including those preceding the 1973–75, 1980, 1990–91, and 2007–09 recessions. This pattern suggests that labor-market data often contain information about the buildup of cyclical pressures before downturns. The finding is consistent with Michaillat and Saez (2025), who show that labor-market indicators provide useful information for assessing recession risk and identifying turning points in the business cycle.



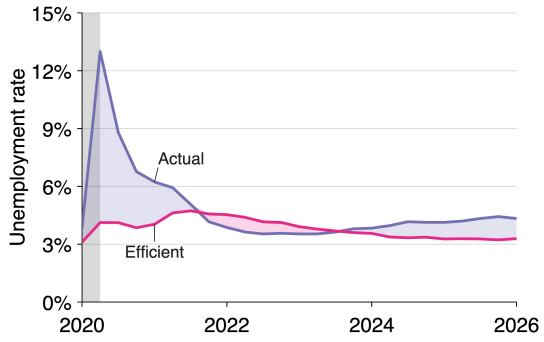
(a) Efficient labor-market tightness with part-time employment, 1951–2019



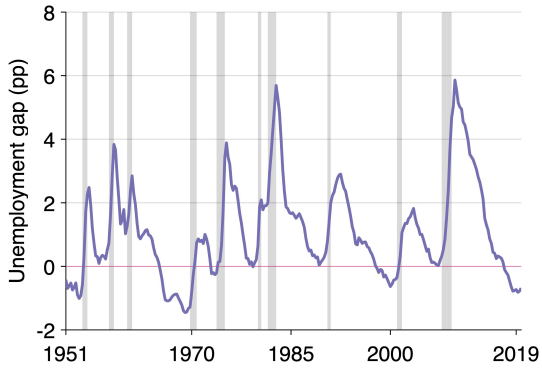
(b) Efficient labor-market tightness with part-time employment, 2020–2026



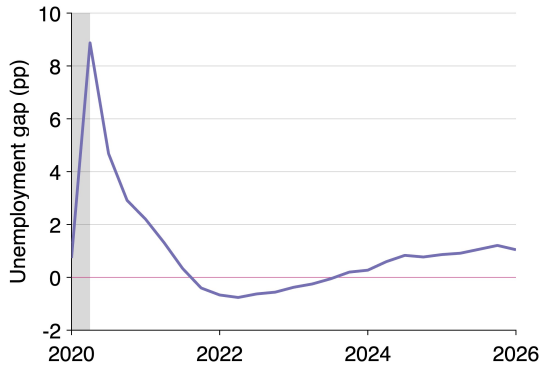
(c) Efficient unemployment rate with part-time employment, 1951–2019



(d) Efficient unemployment rate with part-time employment, 2020–2026



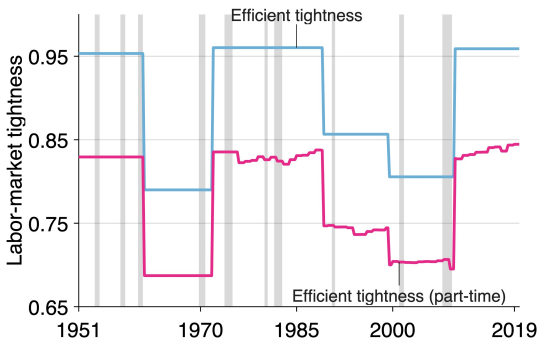
(e) Unemployment gap, 1951–2019



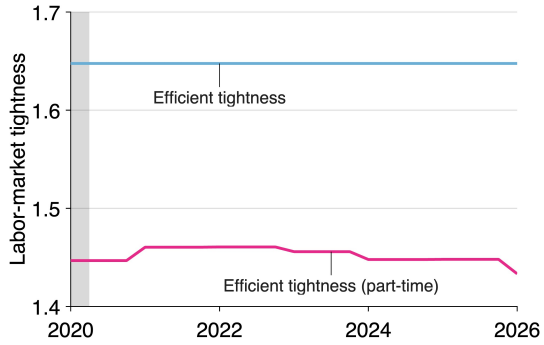
(f) Unemployment gap, 2020–2026

Figure 1.12: Efficient labor-market tightness, efficient unemployment, and unemployment gap with part-time employment in the United States, 1951–2026.

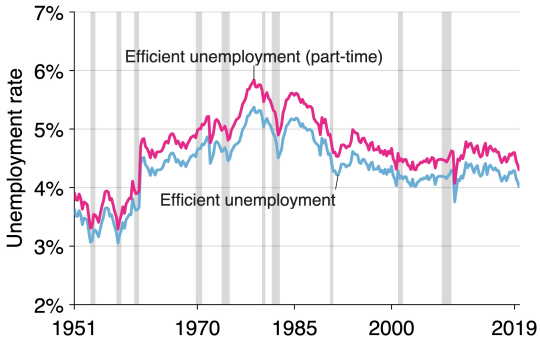
Notes: The figure reports efficient labor-market tightness, the efficient unemployment rate, and the unemployment gap under the extended framework with part-time employment. The calibration sets the full-time employment share to $\alpha^P = 0.75$ and the relative-hours parameter to $\gamma^P = 0.47$. The sample is split between the pre-pandemic period, 1951–2019, and the post-pandemic period, 2020–2026.



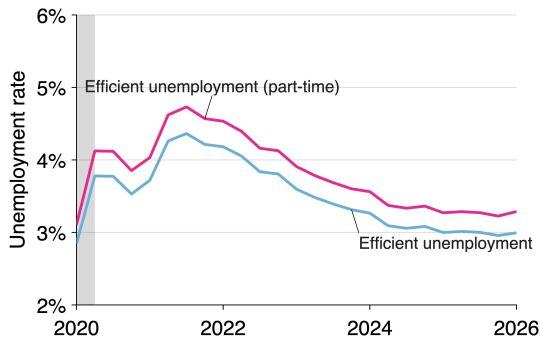
(a) Efficient labor-market tightness comparison, 1951–2019



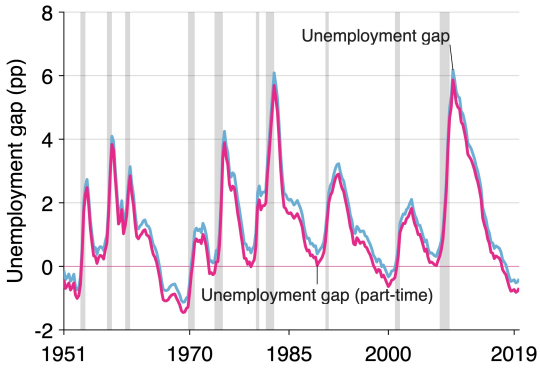
(b) Efficient labor-market tightness comparison, 2020–2026



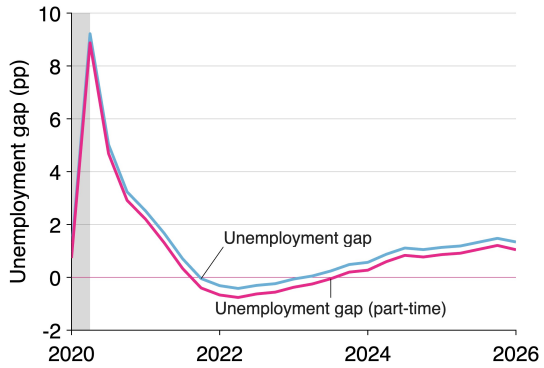
(c) Efficient unemployment rate comparison, 1951–2019



(d) Efficient unemployment rate comparison, 2020–2026



(e) Unemployment gap comparison, 1951–2019



(f) Unemployment gap comparison, 2020–2026

Figure 1.13: Comparison of efficient labor-market tightness, efficient unemployment, and unemployment gaps with and without total part-time employment in the United States.

Notes: The figure compares the baseline framework of Michaillat and Saez (2021a) and the extended framework that incorporates total part-time employment.

Incorporation of involuntary part-time

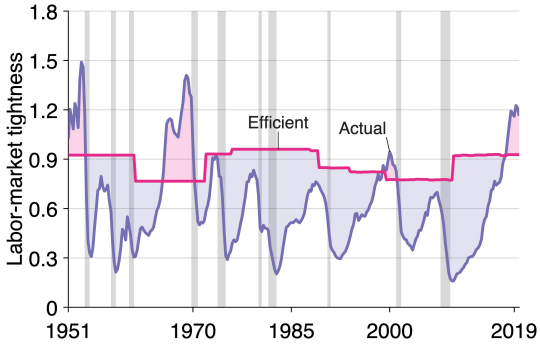
I next focus on involuntary part-time employment, defined as employment part time for economic reasons. This exercise is not a full decomposition of employment because voluntary part-time workers are excluded from the construction of α^{IP}. Instead, it isolates the underemployment margin by asking how the efficient-unemployment benchmark changes when the part-time margin is restricted to workers who would prefer additional hours.

Under this definition, involuntary part-time employment represents a much smaller share of employment than total part-time employment. The calibration therefore sets the full-time employment share to $\alpha^{IP} = 0.93$ and the relative-hours parameter to $\gamma^{IP} = 0.56$.

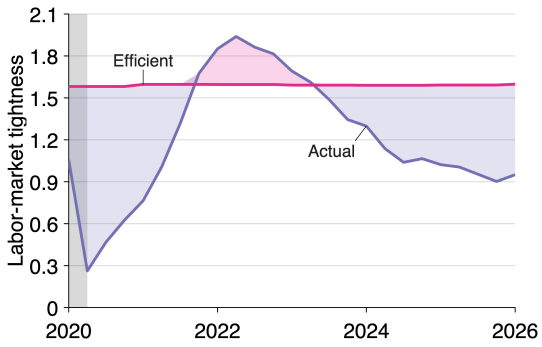
Figure 1.14 reports efficient labor-market tightness, efficient unemployment, and the unemployment gap under the involuntary part-time extension. Before COVID, actual labor-market tightness averages 0.620, while efficient tightness averages 0.867. After COVID, actual tightness averages 1.206, compared with an efficient level of 1.597. Efficient tightness ranges from 0.766 to 0.951 before COVID. In the post-pandemic sample, efficient tightness fluctuates only slightly around 1.597 because the Beveridge elasticity is estimated over a single regime and the remaining sufficient statistics exhibit limited variation.

The corresponding efficient unemployment rate averages 4.42 % before COVID and 3.58 % after COVID, compared with actual unemployment rates of 5.77 % and 4.82 %, respectively. Before COVID, efficient unemployment ranges from 3.10 % to 5.47 %, while after COVID it ranges from 2.91 % to 4.46 %. These results suggest that incorporating involuntary part-time employment shifts the efficient benchmark but does not substantially alter the qualitative behavior of the unemployment gap.

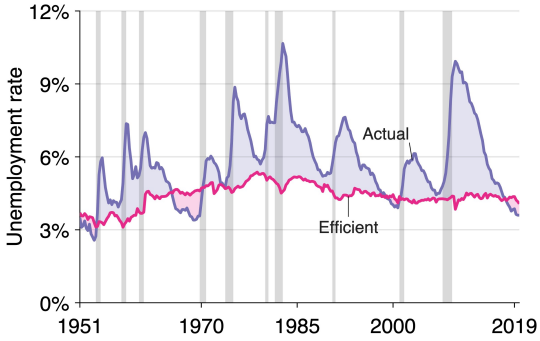
Figure 1.15 compares these results with the full-time baseline. The involuntary part-time extension produces estimates that are close to the baseline because workers employed part time for economic reasons account for only a small fraction of employment in the United States. Consequently, incorporating involuntary part-time employment has only a modest effect on measured labor-market efficiency. Nevertheless, this group remains economically important because it captures workers who would prefer to work more hours but are unable to do so, making it a useful indicator of labor-market slack.



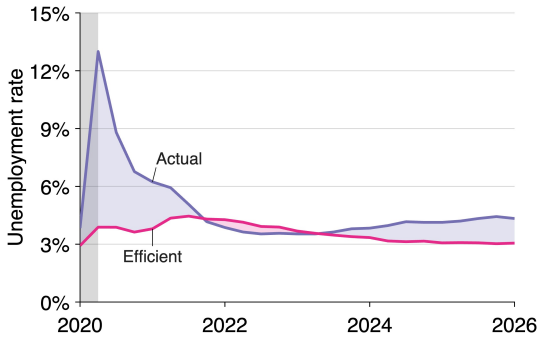
(a) Efficient labor-market tightness with involuntary part-time, 1951–2019



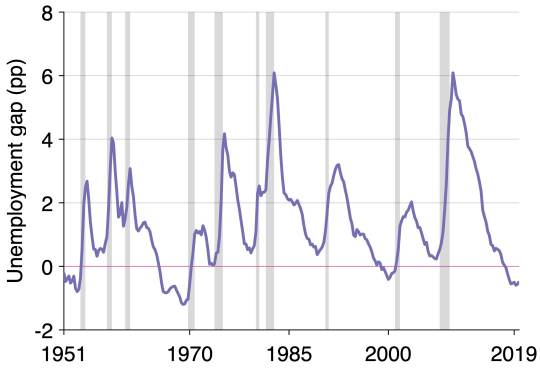
(b) Efficient labor-market tightness with involuntary part-time, 2020–2026



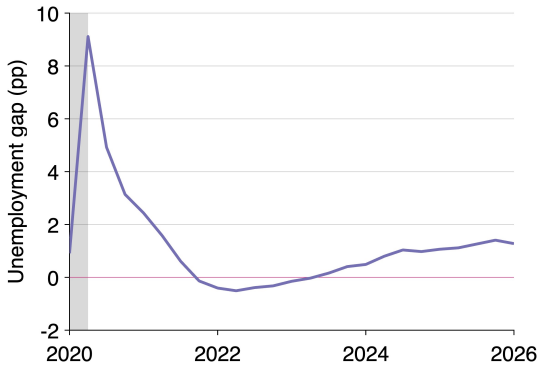
(c) Efficient unemployment rate with involuntary part-time, 1951–2019



(d) Efficient unemployment rate with involuntary part-time, 2020–2026



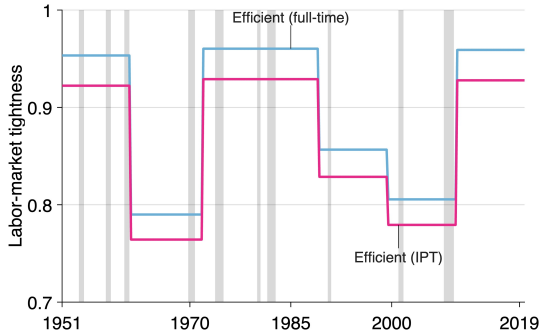
(e) Unemployment gap with involuntary part-time, 1951–2019



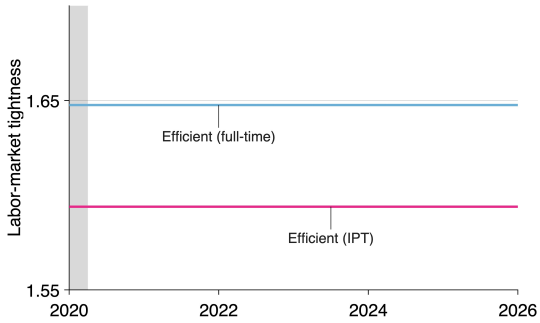
(f) Unemployment gap with involuntary part-time, 2020–2026

Figure 1.14: Efficient labor-market tightness, efficient unemployment, and unemployment gap in the United States.

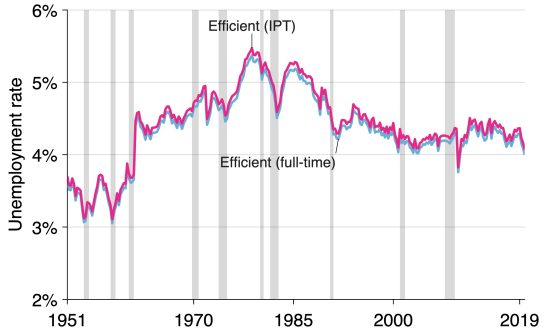
Notes: The figure reports efficient labor-market tightness, the efficient unemployment rate, and the unemployment gap under the extended framework with involuntary part-time employment. The sample is split between the pre-pandemic period, 1951–2019, and the post-pandemic period, 2020–2026.



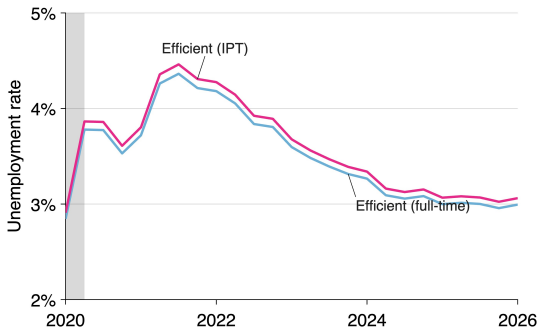
(a) Efficient labor-market tightness comparison, 1951–2019



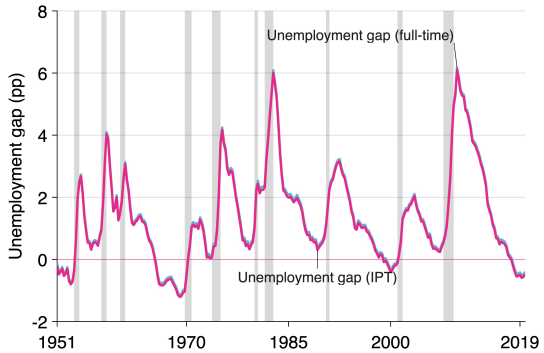
(b) Efficient labor-market tightness comparison, 2020–2026



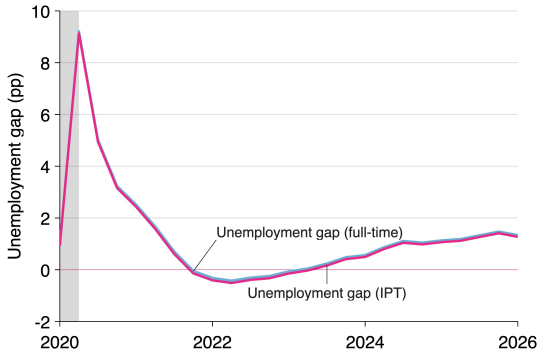
(c) Efficient unemployment rate comparison, 1951–2019



(d) Efficient unemployment rate comparison, 2020–2026



(e) Unemployment gap comparison, 1951–2019



(f) Unemployment gap comparison, 2020–2026

Figure 1.15: Comparison of efficient labor-market tightness, efficient unemployment, and unemployment gaps with and without involuntary part-time employment in the United States.

Notes: The figure compares the baseline framework of Michaillat and Saez (2021a) and the extended framework that incorporates involuntary part-time employment.

1.5.3 Beveridgean unemployment gap in the United States, 1951–2019, 2020–2026

Using the estimated Beveridge elasticity, social value of nonwork, recruiting cost, employment shares, and relative working hours, I compute efficient unemployment for the United States under both total part-time and involuntary part-time employment definitions. I also estimate the framework over the full 1951–2026 sample with five structural breaks. Because the Beveridge curve becomes unusually flat during the COVID-19 period, full-sample estimates provide limited information about underlying labor-market efficiency. I therefore focus on separate pre-COVID and post-COVID specifications as the main analysis, while reporting the full-sample results in Appendix A for completeness.

Under the total part-time calibration, efficient labor-market tightness averages 0.78 before COVID and 1.43 after COVID, compared with actual tightness of 0.62 and 1.21. The corresponding efficient unemployment rate averages 4.68% before COVID and 3.86% after COVID, compared with actual unemployment rates of 5.77% and 4.82%. The unemployment gap remains strongly countercyclical and widens sharply in recessions.

On average, the U.S. labor market remains slack relative to the efficient allocation, both before and after the pandemic. This average pattern, however, does not rule out episodic tightness. Actual tightness approaches or exceeds the efficient benchmark in several pre-recession episodes, suggesting that the part-time extension changes the interpretation of labor-market conditions at business-cycle peaks.

Accounting for part-time employment lowers efficient labor-market tightness and raises efficient unemployment because part-time workers contribute fewer effective hours than full-time workers. Thus, some convergence between actual and

efficient labor-market conditions is a direct consequence of incorporating differences in labor input. The relevant question is whether this adjustment is quantitatively important. The results suggest that it is. Relative to the full-time baseline results, the unemployment gap narrows substantially in several episodes once differences in hours worked are taken into account, indicating that headcount employment alone may overstate labor-market slack.

As a result, the extended framework identifies several episodes in which actual labor-market tightness approaches or exceeds the efficient benchmark, most notably before the 1973–75 recession associated with the First Oil Crisis and again around the early-1980s downturns, before the 1990–91 recession, and before the 2007–09 Global Financial Crisis. These episodes appear considerably closer to the efficient allocation than under the full-time benchmark, highlighting the importance of the hours margin when assessing labor-market conditions.

Under the involuntary part-time calibration, efficient tightness averages 0.87 before COVID and 1.60 after COVID, and efficient unemployment averages 4.42% and 3.58%. Because involuntary part-time employment accounts for only a small share of total employment in the United States, these estimates remain close to the full-time baseline of Michaillat and Saez (2021a).

Overall, accounting for part-time employment mainly affects the efficient benchmark through effective labor input: part-time workers supply fewer market hours than full-time workers, so identical measured unemployment can imply different welfare depending on employment composition. This adjustment matters more under the total part-time definition than under the involuntary part-time definition, but the cyclical behavior of the unemployment gap remains broadly similar to the baseline framework.

1.5.4 Robustness of U.S. results

To further assess the robustness of the findings, I follow the procedure in Michaillat and Saez (2021a). Figures 1.16 and 1.17 report alternative calibrations of the sufficient statistics for the 1951–2019 and 2020–2026 samples, respectively. Figures 1.18 and 1.19 report the corresponding inverse-optimum sufficient statistics. Summary statistics for all panels appear in Tables A.4–A.6 and Tables A.7–A.9 in Appendix A.

Alternative calibrations

Under the extended framework with part-time employment, Proposition 1 implies that labor-market tightness $\theta \equiv v/u$ is efficient when

$$\theta^* = \frac{\alpha + (1-\alpha)\left[\gamma + (1-\gamma)z\right] - z}{c\,\epsilon}. \tag{1.22}$$

Imposing the isoelastic Beveridge curve $v = \bar{\alpha}\, u^{-\epsilon}$ and solving for the unemployment rate yields the efficient unemployment rate

$$u^* = \left[\frac{c\,\epsilon}{\alpha + (1-\alpha)\left[\gamma + (1-\gamma)z\right] - z} \cdot \frac{v}{u^{-\epsilon}}\right]^{\frac{1}{1+\epsilon}}. \tag{1.23}$$

Equation (1.23) is evaluated quarter by quarter using observed unemployment u and vacancy v, the branch-specific Beveridge elasticity ϵ estimated with Bai–Perron break dates, baseline calibrations $z = 0.26$ and $c = 0.92$, and the part-time parameters $\alpha = 0.75$ (whole part-time series) and time-varying γ from U.S. working-hours data.

For each panel of Figures 1.16 and 1.17, I substitute alternative values of one sufficient statistic into (1.23) while holding the others fixed at baseline:

- **Panel A (Beveridge elasticity).** I set ϵ to its branch-specific point esti-

mate and to the lower and upper ends of its 95% confidence interval, computed from the corrected standard errors in the Bai–Perron output.

- **Panel B (social value of nonwork).** I set z to its baseline value 0.26 and to the endpoints of the calibration range [0.03, 0.49].

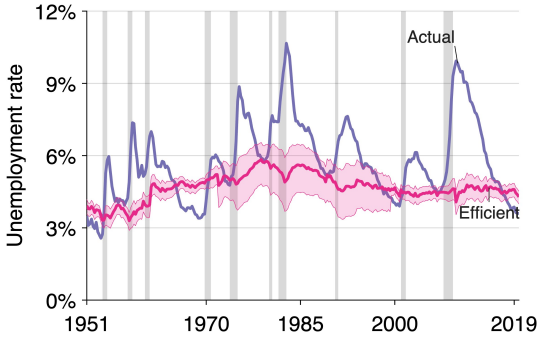
- **Panel C (recruiting cost).** I set c to its baseline value 0.92 and to the endpoints of the calibration range $[\frac{2}{3} \times 0.92, \frac{4}{3} \times 0.92]$.

In each panel, the thick pink line plots u^* under the baseline calibration of the displayed statistic; the thin pink lines and shaded pink area trace u^* at the lower and upper bounds; and the purple line plots the actual unemployment rate.

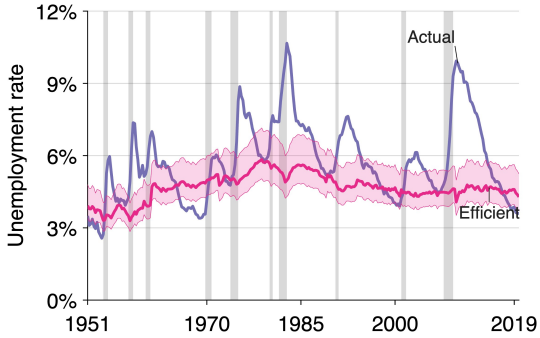
The efficient unemployment rate is stable across alternative calibrations of the Beveridge elasticity, the social value of nonwork, and recruiting costs. Consistent with Michaillat and Saez (2021a), calibration uncertainty has limited effects before COVID. Over 1951–2019, varying the Beveridge elasticity within its 95% confidence interval shifts the efficient unemployment rate by about ± 0.5 percentage point on average, with a maximum deviation of 1.3 percentage points (Table A.4). Similar patterns hold for the social value of nonwork and recruiting costs. Alternative calibrations typically move the efficient rate by less than 1 percentage point on average, and the maximum deviations remain below 1.3 percentage points (Tables A.5 and A.6). Overall, the main unemployment-gap patterns are not driven by a particular choice of Beveridge elasticity, social value of nonwork, or recruiting cost.

The 2020–2026 sample exhibits somewhat greater sensitivity, reflecting the unusually flat Beveridge curve during and immediately after the pandemic. Under alternative calibrations, the efficient unemployment rate differs from baseline by about 0.7–1.0 percentage point on average in panel A, with a maximum deviation

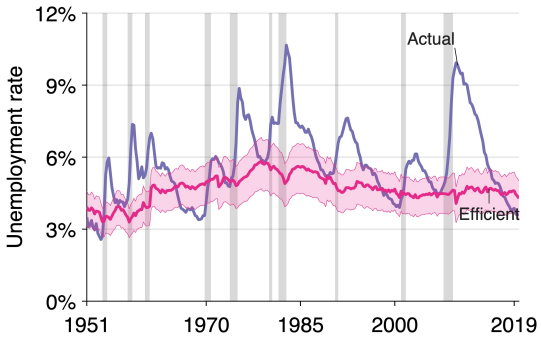
of 1.4 percentage points in 2020Q2. Panels B and C show similarly modest post-COVID changes, with maximum deviations between 1.1 and 1.3 percentage points. Overall, the efficient unemployment rate remains quantitatively stable across a wide range of parameter values.



(a) Beveridge elasticity: ϵ in the 95% confidence interval

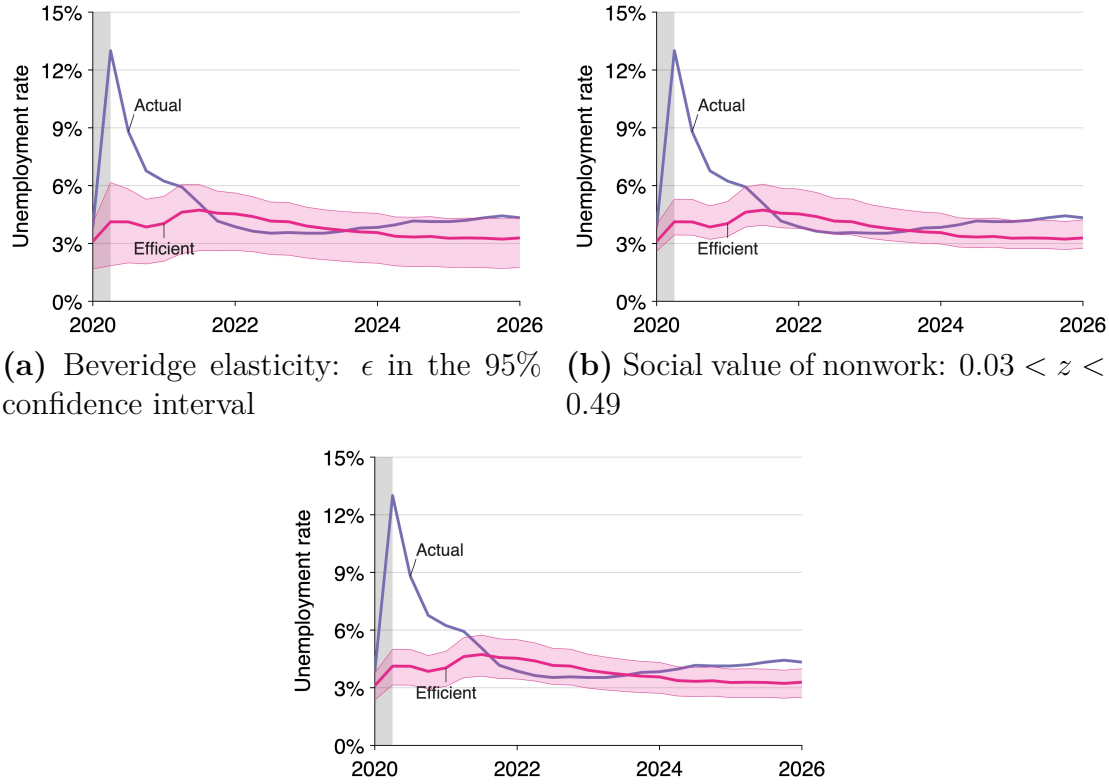


(b) Social value of nonwork: $0.03 < z < 0.49$



(c) Recruiting cost: $0.61 < c < 1.23$

Figure 1.16: Efficient unemployment rates for alternative sufficient statistics in the United States, 1951–2019. *Notes:* The figure reports the efficient unemployment rate implied by the model under alternative values of the Beveridge elasticity, the social value of nonwork, and recruiting costs. Each panel varies one sufficient statistic while holding the remaining sufficient statistics at their baseline values.



(a) Beveridge elasticity: ϵ in the 95% confidence interval

(b) Social value of nonwork: $0.03 < z < 0.49$

(c) Recruiting cost: $0.61 < c < 1.23$

Figure 1.17: Efficient unemployment rates for alternative sufficient statistics in the United States, 2020–2026. *Notes:* The figure reports the efficient unemployment rate implied by the model under alternative values of the Beveridge elasticity, the social value of nonwork, and recruiting costs. Each panel varies one sufficient statistic while holding the remaining sufficient statistics at their baseline values.

Inverse-optimum sufficient statistics

The inverse-optimum exercise asks which value of each sufficient statistic would make *observed* labor-market tightness efficient, holding the other sufficient statistics at their baseline calibrations. From (1.22), actual tightness is efficient when the Beveridge elasticity satisfies

$$\epsilon^* = \frac{\alpha + (1-\alpha)\left[\gamma + (1-\gamma)z\right] - z}{c\,\theta}. \tag{1.24}$$

Similarly, actual tightness is efficient when the social value of nonwork satisfies

$$z^* = \frac{\alpha + (1-\alpha)\gamma - c\,\epsilon\,\theta}{1-(1-\alpha)(1-\gamma)}, \tag{1.25}$$

and when the recruiting cost satisfies

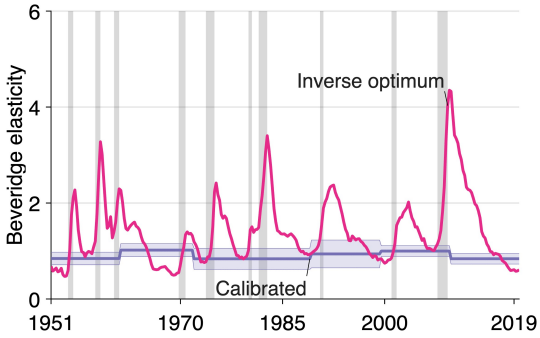
$$c^* = \frac{\alpha + (1-\alpha)\left[\gamma + (1-\gamma)z\right] - z}{\epsilon\,\theta}. \tag{1.26}$$

In each panel of Figures 1.18 and 1.19, the pink line plots the corresponding inverse-optimum series (ϵ^*, z^*, or c^*); the thick purple line plots the calibrated value; and the thin purple lines and shaded purple area mark the 95% confidence interval (panel A) or calibration range (panels B and C).

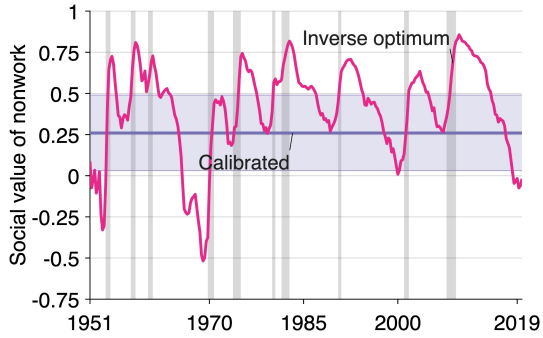
Over 1951–2019, the inverse-optimum sufficient statistics frequently depart from their calibrations, especially in recessions, reproducing the patterns in Michaillat and Saez (2021a). The inverse-optimum Beveridge elasticity averages 1.42, compared with a calibrated value of about 0.91 and lies above its 95% confidence interval in 61% of quarters (Table A.7). The inverse-optimum social value of nonwork averages 0.41, exceeds its calibration range in 46% of quarters, and turns negative in 12% of quarters (Table A.8). The inverse-optimum recruiting cost averages 1.46, well above the calibrated $c = 0.92$, and exceeds its calibration range in 51% of quarters (Table A.9). These departures indicate that, under the calibrated sufficient statistics, observed labor-market tightness was often lower than efficient—consistent with inefficient slack, especially in downturns.

Over 2020–2026, the inverse-optimum values align more closely with their calibrations. The inverse-optimum Beveridge elasticity averages 0.64 and lies within its 95% confidence interval in 76% of quarters; the inverse-optimum social value of nonwork averages 0.28 and lies within [0.03, 0.49] in 84% of quarters; and the

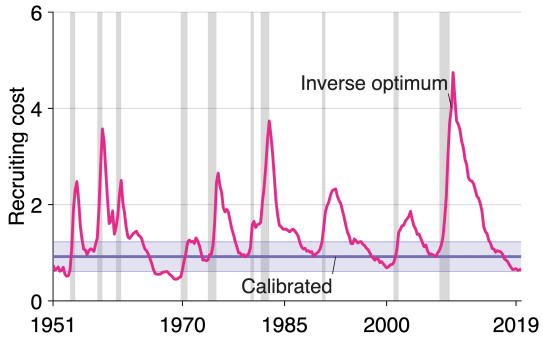
inverse-optimum recruiting cost averages 1.20 and lies within its calibration range in 76% of quarters. Large deviations are concentrated in 2020Q2, when unemployment spiked and the Beveridge curve was unusually flat.



(a) Inverse-optimum Beveridge elasticity

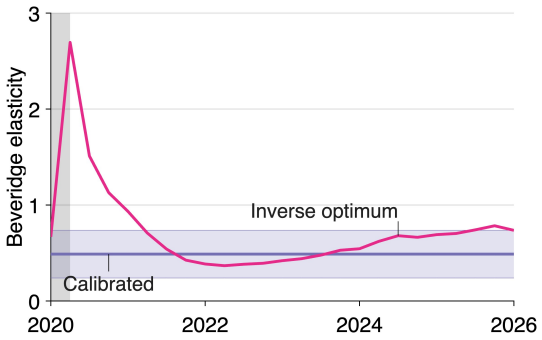


(b) Inverse-optimum social value of nonwork

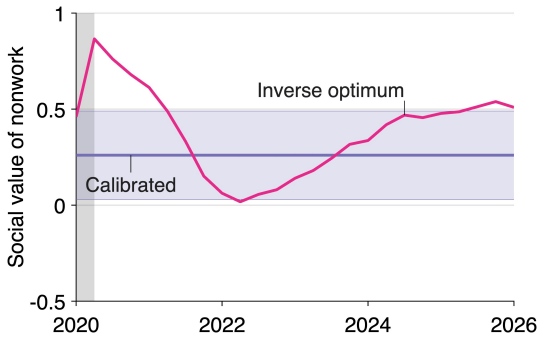


(c) Inverse-optimum recruiting cost

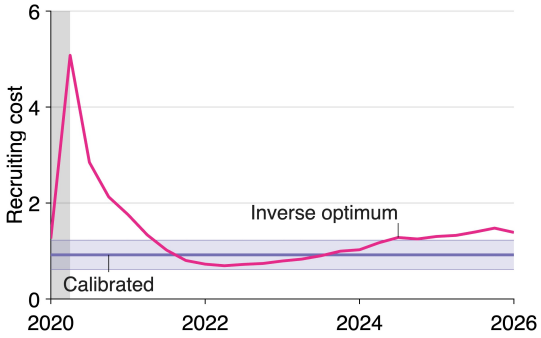
Figure 1.18: Inverse-optimum sufficient statistics in the United States, 1951–2019. *Notes:* The figure reports the inverse-optimum Beveridge elasticity, social value of nonwork, and recruiting cost implied by the efficient-unemployment framework. Each panel recovers the value of a sufficient statistic consistent with the observed unemployment rate and labor-market tightness.



(a) Inverse-optimum Beveridge elasticity



(b) Inverse-optimum social value of nonwork



(c) Inverse-optimum recruiting cost

Figure 1.19: Inverse-optimum sufficient statistics in the United States, 2020–2026. *Notes:* The figure reports the inverse-optimum Beveridge elasticity, social value of nonwork, and recruiting cost implied by the efficient-unemployment framework. Each panel recovers the value of a sufficient statistic consistent with the observed unemployment rate and labor-market tightness.

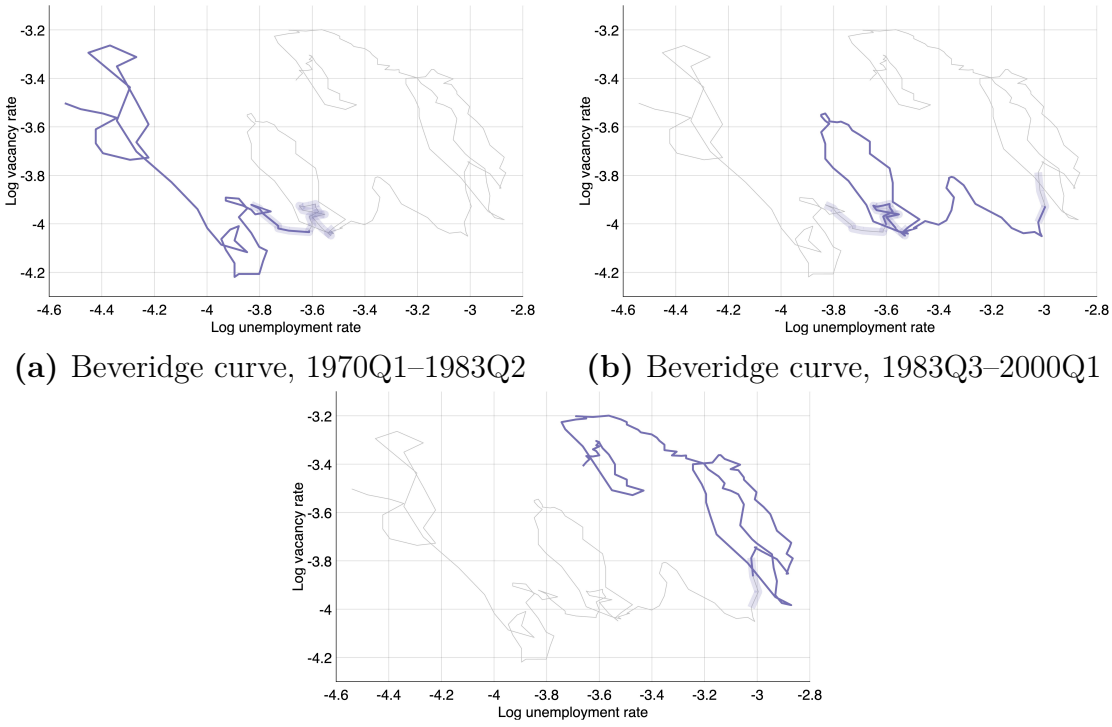
The efficient unemployment rate is stable across a wide range of plausible calibrations, consistent with Michaillat and Saez (2021a). The inverse-optimum exercise further shows that pre-COVID labor-market outcomes can be reconciled with efficiency only under often-implausible parameter values. These robustness exercises support the main U.S. finding: accounting for part-time employment does not overturn the conclusion that the labor market was generally inefficiently slack, especially during recessions. The post-COVID evidence is weaker because the unusually flat Beveridge curve reduces the informativeness of the sufficient-statistics approach, but it remains broadly consistent with this conclusion.

1.6 Application to Japan

Part-time employment is a central and growing feature of the Japanese labor market, as documented in Figures 1.3 and 1.21. From 2002 to 2025, full-time employment remained relatively stable while part-time employment increased substantially. Table 1.6 indicates that Japan had, on average, 48.96 million full-time workers and 13.75 million part-time workers during this period. Part-time workers accounted for 21.9 % of total employment among full-time and part-time workers.

However, Figure 1.23 and Tables 1.6–1.8 show that the rise in part-time employment is not primarily driven by involuntary part-time work. Involuntary part-time employment accounts for only 5.9 % of full-time and involuntary part-time employment on average, and only 22.9 % of total part-time employment. Moreover, the involuntary part-time share has declined over time. This distinction is important because total part-time employment captures the broader hours-margin mechanism, while involuntary part-time employment captures hidden underemployment.

I therefore apply the extended Beveridgean unemployment framework to Japan using total part-time employment. This specification captures the mechanism emphasized in the model. Part-time workers count as employed in headcount terms, but they supply fewer effective hours than full-time workers. I then consider a more restrictive specification based on involuntary part-time employment. This second exercise asks whether the results are driven by hidden underemployment rather than by the broader expansion of lower-hours employment.



(a) Beveridge curve, 1970Q1–1983Q2

(b) Beveridge curve, 1983Q3–2000Q1

(c) Beveridge curve, 2000Q2–2025Q3

Figure 1.20: Beveridge curve estimation in Japan, 1970–2025.

Notes: The number of structural breaks and their dates are estimated using the Bai–Perron (1998, 2003) algorithm.

To implement the measurement for Japan, I first estimate the Beveridge elasticity by regressing the log vacancy rate on the log unemployment rate. The unemployment rate is obtained from the OECD database, and the vacancy rate is obtained from Statistics Bureau of Japan. The estimation sample runs from 1970Q1 to 2025Q3 (223 quarterly observations). As in the U.S. application, I allow for structural breaks in the Beveridge relationship using the Bai–Perron algorithm of Bai and Perron (1998, 2003), following the setup in Michaillat and Saez (2021a). I allow for different variances of the errors across regimes, for autocorrelation in the errors, and for different distributions of the independent and dependent variables across regimes. Standard errors are corrected for autocorre-

lation and heteroskedasticity using a quadratic kernel with automatic bandwidth selection based on an AR(1) approximation (Andrews, 1991). I set the maximum number of breaks to two and the trimming parameter to 0.20, so that each regime contains at least 20% of the sample.

The supF tests reject the null hypothesis of no structural break at the 1% significance level. Both the Bayesian information criterion (Yao, 1988) and the modified Schwarz criterion (Liu et al., 1997) select two breaks. The estimated break dates occur in 1983Q2 and 2000Q1, yielding three Beveridge-curve regimes: 1970Q1–1983Q2, 1983Q3–2000Q1, and 2000Q2–2025Q3. The resulting Beveridge elasticity estimates are 0.86, 0.46, and 0.63, with corrected standard errors of 0.15, 0.10, and 0.08, respectively. The overall fit is $R^2 = 0.80$. Figure 1.20 plots the estimated Beveridge curve within each regime. The figure shows a clear outward shift in the Japanese Beveridge curve over time, indicating that the same unemployment rate has been associated with higher vacancy rates in later subperiods.

For the remaining sufficient statistics, I set $z = 0.26$ and $c = 0.92$ following Michaillat and Saez (2021a). I construct α from the OECD part-time employment share and set $\gamma = 0.56$ using hours data from the Statistics Bureau of Japan, Labour Force Survey Detailed Tabulation. The Japanese hours data imply average weekly hours of 24.19 for part-time workers and 43.59 for full-time workers, so $\gamma = H^P/H^F \approx 0.56$. Using the U.S. values of z and c keeps the cross-country comparison transparent and isolates the role of Japan's employment composition and relative hours. However, these parameters may differ across countries because recruiting institutions, employment protection, job-search behavior, and the value of nonmarket time are not identical in Japan and the United States. The Japan results should therefore be interpreted as measuring how incorporating part-time employment changes the efficient-unemployment benchmark while holding the

baseline preference and recruiting-cost parameters fixed. Accordingly, the levels of efficient unemployment should not be interpreted as directly comparable across the United States and Japan. The relevant comparison is within each country, between the full-time benchmark and the part-time calibrations. To assess the sensitivity of this choice, I later vary z and c over plausible ranges and compute inverse-optimum values of the sufficient statistics.

The Japanese calibration combines OECD employment counts with hours data from the Labour Force Survey Detailed Tabulation. This introduces a potential definitional difference because part-time employment is not always measured using the same 35-hour threshold used in the U.S. CPS classification. I therefore interpret the Japan results as using the OECD/LFS classification of part-time employment rather than imposing a U.S.-specific hours cutoff. The relative-hours parameter γ partially addresses this issue by measuring the actual hours gap between part-time and full-time workers in the Japanese data. Nevertheless, differences in national definitions of part-time employment should be kept in mind when comparing the magnitude of the part-time adjustment across countries.

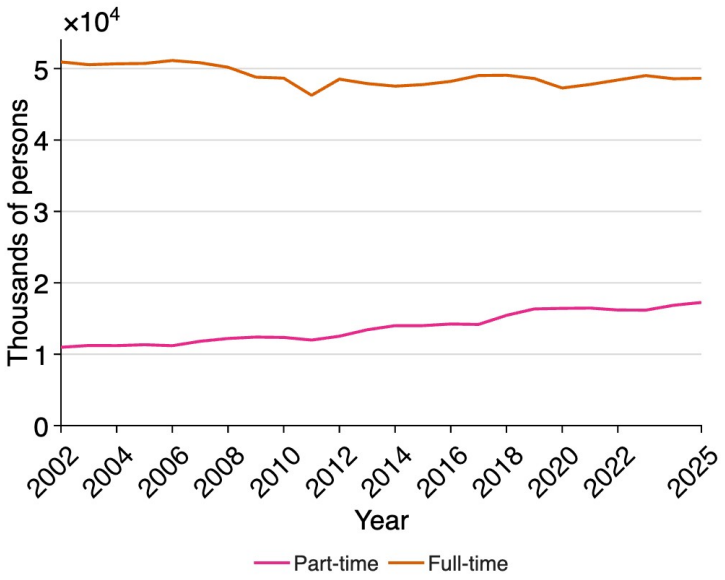


Figure 1.21: Full-time and part-time employment in Japan, 2002–2025.

Notes: The figure reports annual full-time and part-time employment in Japan. Full-time employment remained relatively stable over the sample period, fluctuating around 49 million workers. In contrast, part-time employment increased steadily from approximately 11 million workers in 2002 to more than 17 million workers in 2025, reflecting the growing importance of part-time work in the Japanese labor market. *Source:* OECD Employment Database.

As shown in Figure 1.21, full-time employment remained relatively stable over 2002–2025, whereas part-time employment increased steadily. At the same time, the nature of part-time employment also changed. Figure 1.22 shows that although total part-time employment expanded substantially, the share of involuntary part-time workers declined from roughly 30 % to less than 20 % over the same period.

Table 1.6: Summary Statistics of Full-Time and Part-Time Employment

Statistic	Full-time (n_F)	Part-time (n_P)	α^P	$1-\alpha^P$
Mean	48.96 million	13.75 million	0.781	0.219
SD	1.31 million	2.15 million	0.971	0.029
Min	46.26 million	10.97 million	0.823	0.177
Max	51.13 million	17.25 million	0.738	0.262
N	24	24	24	24

Notes: n_F denotes full-time employment, n_P denotes total part-time employment, and $n_P/(n_F+n_P)$ is the share of part-time employment in total employment. The sample contains 24 annual observations from 2002 to 2025. *Source*: OECD Database.

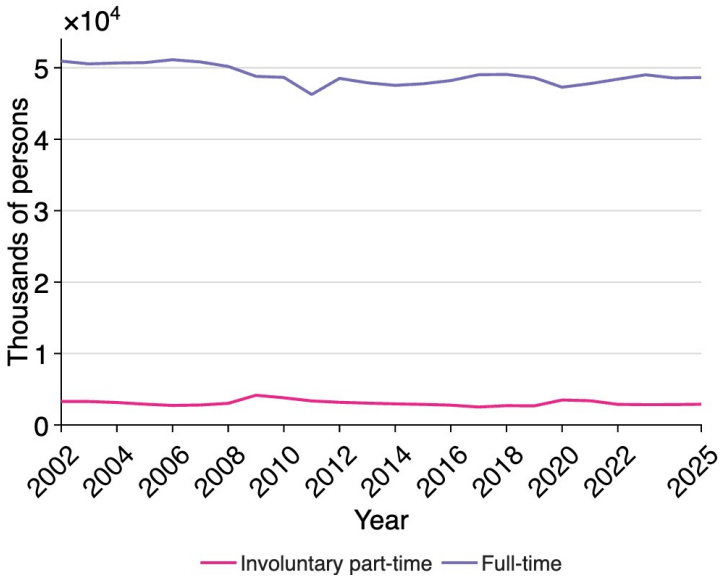


Figure 1.22: Involuntary part-time and full-time employment in Japan, 2002–2025.

Notes: The figure reports total part-time employment, involuntary part-time employment, and the involuntary part-time share in Japan. Although total part-time employment increased steadily over the sample period, involuntary part-time employment remained much smaller and the involuntary share declined over time. On average, involuntary part-time workers account for about 20–25 % of total part-time employment. *Source:* OECD Employment Database.

Table 1.7: Summary Statistics of Full-Time and Involuntary Part-Time Employment

Statistic	Full-time (n_F)	Involuntary part-time (n_{IP})	α^{IP}	$1-\alpha^{IP}$
Mean	48.96 million	3.06 million	0.941	0.059
SD	1.32 million	0.38 million	0.007	0.007
Min	46.26 million	2.51 million	0.921	0.049
Max	51.13 million	4.16 million	0.951	0.079
N	24	24	24	24

Notes: n_F denotes full-time employment, n_{IP} denotes involuntary part-time employment, and $\alpha = n_F/(n_F + n_{IP})$. The sample contains 24 annual observations. *Source*: OECD database.

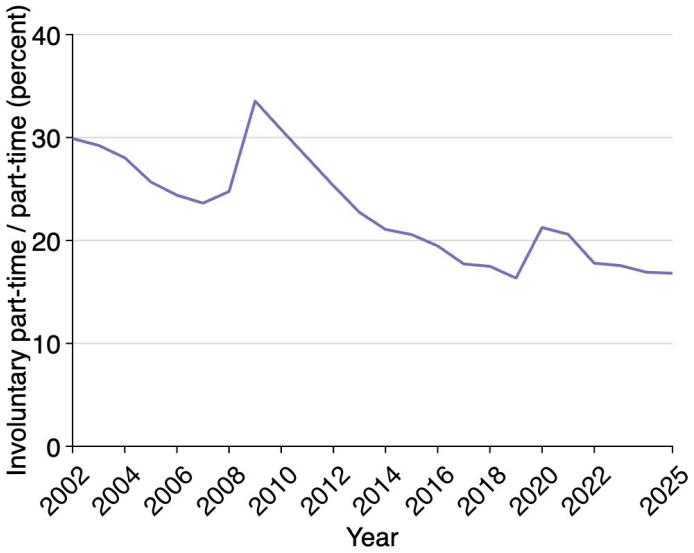


Figure 1.23: Involuntary part-time employment as a share of total part-time employment in Japan, 2002–2025.

Notes: The figure reports the ratio of involuntary part-time employment to total part-time employment. Although part-time employment expanded substantially over the sample period, the involuntary share declined from roughly 30 % to less than 20 %. *Source:* OECD Employment Database.

Table 1.8: Summary Statistics of Part-Time and Involuntary Part-Time Employment

Statistic	Part-time (n_P)	Involuntary part-time (n_{IP})	n_{IP}/n_P
Mean	13.76 million	3.06 million	0.229
SD	2.15 million	0.38 million	0.051
Min	10.97 million	2.51 million	0.163
Max	17.25 million	4.16 million	0.335
N	24	24	24

Notes: n_P denotes total part-time employment, n_{IP} denotes involuntary part-time employment, and n_{IP}/n_P is the share of part-time workers who report working part-time for economic reasons. The sample contains 24 annual observations. *Source*: OECD database.

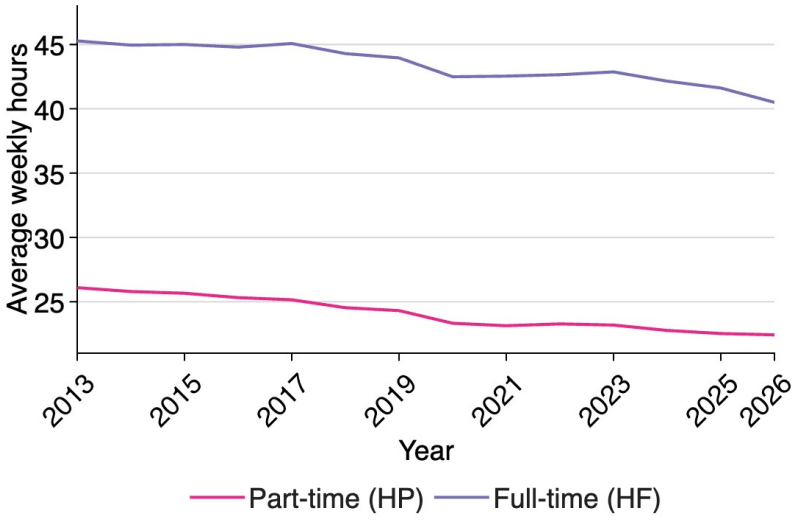


Figure 1.24: Average weekly hours for part-time and full-time workers in Japan, 2013–2026.

Notes: The figure reports average weekly hours for part-time workers (H_P) and full-time workers (H_F). Part-time workers consistently work far fewer hours than full-time workers. Over the sample period, average weekly hours decline for both groups: full-time hours fall from about 45.5 to 40.5 hours per week, while part-time hours fall from about 26 to 22.5 hours per week, with a sharper decline for full-time workers around 2020. On average, $H_P = 24.2$ hours and $H_F = 43.6$ hours, implying $\gamma = H_P/H_F \approx 0.56$. *Source:* Statistics Bureau of Japan, *Labour Force Survey Detailed Tabulation*.

Table 1.9: Summary Statistics of Weekly Hours and Relative Hours in Japan

Statistic	H_P	H_F	γ
Mean	24.19	43.59	0.555
SD	1.30	1.90	0.017
Min	21.80	37.40	0.522
Max	26.70	46.30	0.600
N	160	160	160

Notes: H_P denotes average weekly hours worked by part-time employees, H_F denotes average weekly hours worked by full-time employees, and $\gamma = H_P/H_F$ is the relative-hours parameter used in the calibration. The sample consists of monthly observations from January 2013 to April 2026. *Source*: Statistics Bureau of Japan, *Labour Force Survey Detailed Tabulation*.

The calibration of the part-time specification uses two additional inputs. First, I calibrate the relative-hours parameter, γ, using average weekly hours worked by full-time and part-time workers. Figure 1.24 shows that part-time workers consis-

tently work substantially fewer hours than full-time workers in Japan. Table 1.9 reports the corresponding summary statistics. From January 2013 to April 2026, part-time workers worked 24.19 hours per week on average, while full-time workers worked 43.59 hours. I therefore set

$$\gamma = \frac{H_P}{H_F} = 0.56,$$

so that a part-time worker supplies about 56 % as many weekly hours as a full-time worker.

Second, I calibrate the employment-composition parameter using the full-time and part-time employment shares. Table 1.6 shows that, from 2002 to 2025, Japan had on average 48.96 million full-time workers and 13.75 million part-time workers. Part-time workers accounted for 21.9 % of full-time and part-time employment. I therefore set the pre-2002 full-time share to

$$\alpha = 1 - 0.219 \approx 0.78.$$

For 2002–2025, I use the annual OECD time series of the full-time and part-time employment shares directly. This allows the part-time share to vary over time in the post-2002 period, while using the average post-2002 composition as the calibration for the earlier period. Although OECD total part-time employment data are available before 2002, I start the main Japan employment-share calibration in 2002 to align the total part-time and involuntary part-time series. Since the involuntary part-time series is not available before 2002, using 2002 as the common starting point keeps the two calibrations comparable. For the pre-2002 period, I therefore use the 2002–2025 average full-time share, $\alpha = 0.78$, in the total part-time calibration.

This choice may overstate the historical importance of part-time employment if Japan's part-time share was lower before 2002. In that case, the pre-2002 part-time calibration would overstate the upward shift in efficient unemployment and the downward shift in efficient labor-market tightness before 2002.

As a robustness check, I recompute the pre-2002 results using a higher full-time share, $\alpha = 0.90$, which corresponds to a smaller historical part-time share. Appendix Figure A.7 reports the robustness check using $\alpha = 0.90$ before 2002. The alternative calibration produces a smaller pre-2002 part-time adjustment, as expected, but the post-2002 paths are unchanged because both calibrations use observed OECD part-time employment shares from 2002 onward. Importantly, the main episodes emphasized below—the period around the Global Financial Crisis, the aftermath of the 2011 earthquake, and the post-pandemic period—occur after 2002 and therefore use observed OECD part-time employment shares rather than the backward extrapolation.

The extended Beveridgean unemployment framework changes the interpretation of employment growth in Japan. A rise in employment mechanically lowers the unemployment rate, but it does not necessarily imply a proportional increase in effective labor input when a growing share of employment consists of part-time jobs. Since part-time workers supply fewer hours than full-time workers, part-time employment can reduce measured unemployment while still leaving effective labor utilization lower than suggested by headcount employment.

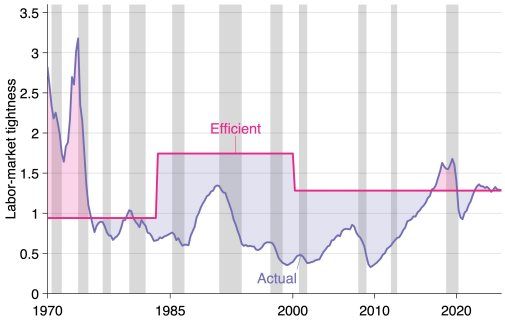
Figure 1.25 compares the efficient-unemployment results under the full-time benchmark and the part-time calibration. Under the full-time benchmark, average efficient labor-market tightness is 1.34 and the average efficient unemployment rate is 2.28 %. After incorporating part-time employment, efficient tightness falls to 1.07 and the efficient unemployment rate rises to 2.71 %. This shift reflects

the lower effective labor input supplied by part-time workers. Because reducing unemployment increasingly adds lower-hours employment, the welfare gain from additional vacancy creation is smaller, which lowers efficient tightness and raises the efficient unemployment benchmark.

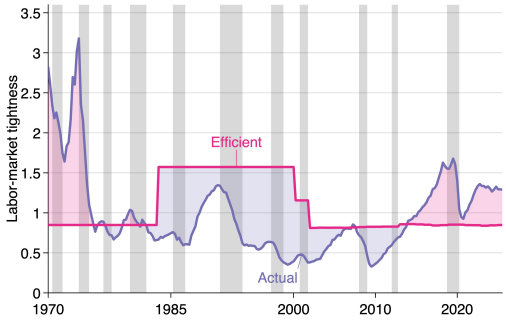
The unemployment gap also changes. Under the full-time benchmark, the average unemployment gap is 0.88 %age points. After incorporating part-time employment, the average gap falls to 0.45 %age points. These results suggest that ignoring part-time employment overstates efficient labor-market tightness and understates the efficient unemployment rate. In headcount terms, Japan may appear more slack than it is relative to the part-time-adjusted benchmark. Once the hours margin is taken into account, the efficient unemployment rate shifts upward and the measured unemployment gap becomes smaller.

The efficient unemployment series in Figure 1.25 inherits several discrete movements from the Bai–Perron estimation because the Beveridge elasticity is allowed to change across regimes. These movements are mechanical consequences of the piecewise elasticity estimates rather than abrupt changes in labor-market fundamentals. To verify that the main conclusion is not driven by these discontinuities, Figure 1.26 recomputes the efficient unemployment rate after smoothing the Beveridge elasticity around the estimated break dates. The smoothed series removes the sharp jumps while leaving the substantive comparison across the full-time, total part-time, and involuntary part-time calibrations largely unchanged. This confirms that the differences across calibrations are not driven by the discontinuous treatment of structural breaks.

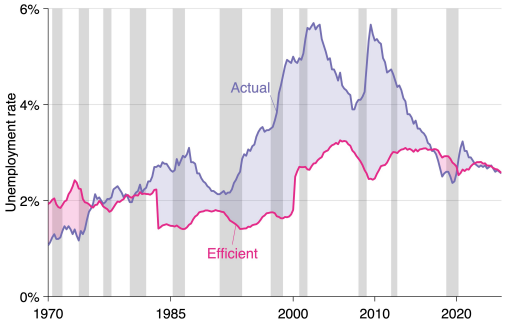
I next examine whether this result reflects hidden underemployment or the broader hours margin associated with part-time work.



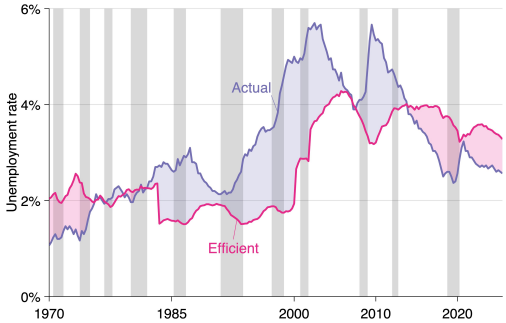
(a) Labor-market tightness without part-time (full-time only), Japan



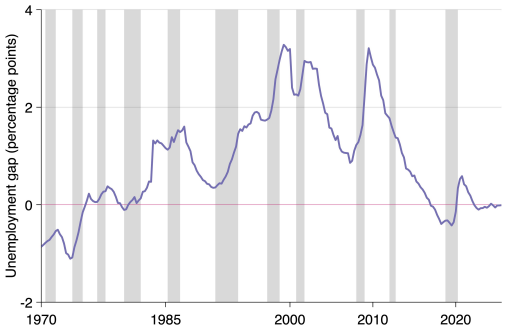
(b) Labor-market tightness with total part-time, Japan



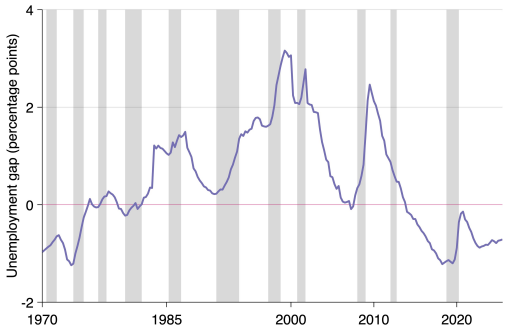
(c) Efficient unemployment rate without part-time (full-time only), Japan



(d) Efficient unemployment rate with total part-time, Japan



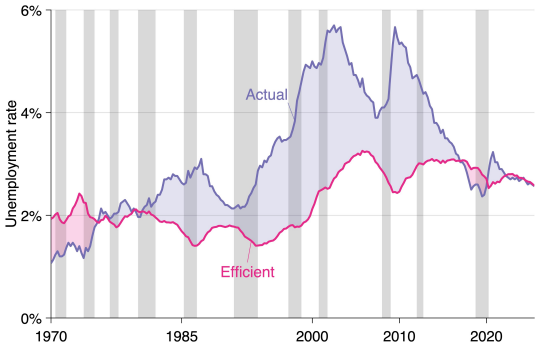
(e) Unemployment gap without part-time (full-time only), Japan



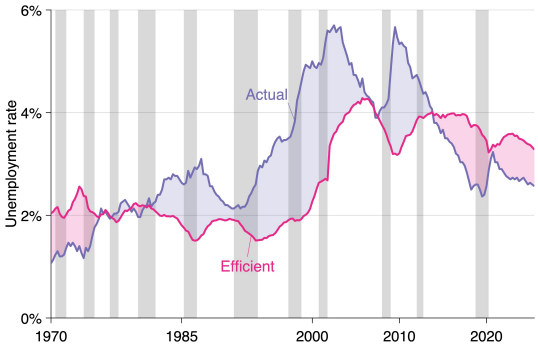
(f) Unemployment gap with total part-time, Japan

Figure 1.25: Efficient labor-market tightness, efficient unemployment rate, and unemployment gap in Japan, 1970–2025.

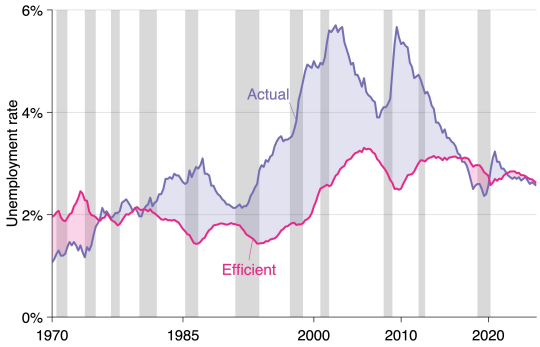
Notes: Left column: full-time benchmark ($\alpha = 1$). Right column: total part-time employment incorporation ($\alpha = 0.78$ through 2001, part-time share from OECD thereafter). Shaded areas are recessions. Panels A–D compare efficient and actual series; panel E–F show the unemployment gap.



(a) Efficient unemployment rate, $\alpha = 1$



(b) Efficient unemployment rate, $\alpha = 0.78$



(c) Efficient unemployment rate, $\alpha = 0.94$

Figure 1.26: Efficient unemployment in Japan using smoothed Beveridge elasticities estimated over Bai–Perron regimes.

Notes: The figure compares the full-time benchmark, with $\alpha = 1$, to part-time calibrations. For total part-time employment, $\alpha = 0.78$ through 2001 and follows the OECD part-time employment share from 2002 onward. The $\alpha = 0.94$ case uses the corresponding involuntary part-time employment path.

1.6.1 2008 Financial Crisis

Before the global financial crisis, the actual unemployment rate and the efficient unemployment rate move close together under the part-time calibration. This suggests that the Japanese labor market was near its efficient allocation during this period once differences in hours worked are taken into account. Without accounting for part-time employment, the same episode appears more slack be-

cause the full-time baseline treats all employed workers as supplying the same effective labor input. The part-time calibration raises the efficient unemployment benchmark by recognizing that a larger share of employment consists of lower-hours jobs. As a result, the observed unemployment rate lies closer to the efficient rate than under the full-time benchmark.

In this sense, the results suggest that labor-market conditions contain information about cyclical turning points. The part-time extension does not merely alter the level of efficient unemployment; it also changes the interpretation of several pre-recession episodes, indicating that employment composition may be relevant for recession monitoring and business-cycle assessment.

1.6.2 2011 Earthquake Aftermath

After the 2011 earthquake, the part-time calibration indicates a clear tightening of the Japanese labor market. Actual unemployment falls below efficient unemployment, implying that the labor market became tight relative to the efficient benchmark. This tightening is less visible under the full-time benchmark, which treats all employed workers as supplying the same effective labor input. Once part-time employment is incorporated, the lower hours supplied by part-time workers raise the efficient-unemployment benchmark, making the post-2011 tightening more apparent.

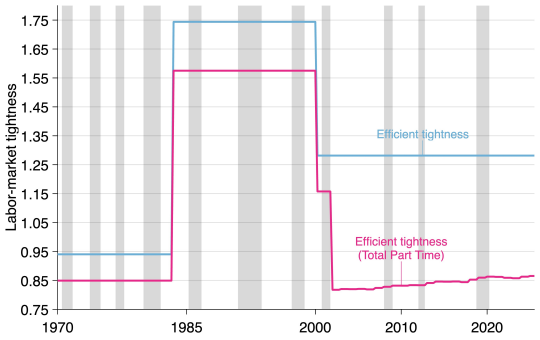
This pattern is consistent with the labor-market disruption following the earthquake. The shock generated job losses and regional mismatch, while reconstruction-related vacancies also increased. The part-time calibration therefore captures a feature of the period that is much less visible under the full-time benchmark. More broadly, the post-2011 period marks the beginning of a longer phase in which Japan's labor market appears increasingly tight, a pattern that continues

into the post-pandemic period.

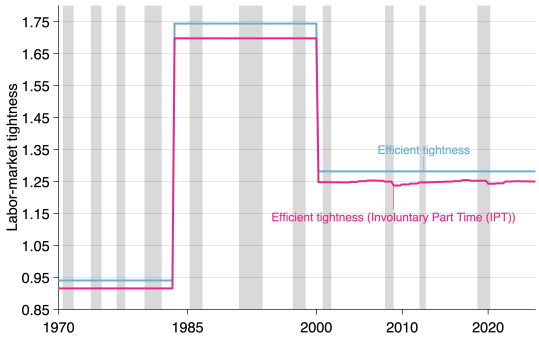
1.6.3 Pandemic Period

Around the COVID-19 pandemic, the labor market initially moved closer to the efficient benchmark. Part of this convergence reflects a contraction in labor supply and labor-force participation. In the post-pandemic period, however, the Japanese labor market tightened again. Actual unemployment remains below the efficient unemployment rate, indicating that labor-market tightness increased relative to the efficient benchmark.

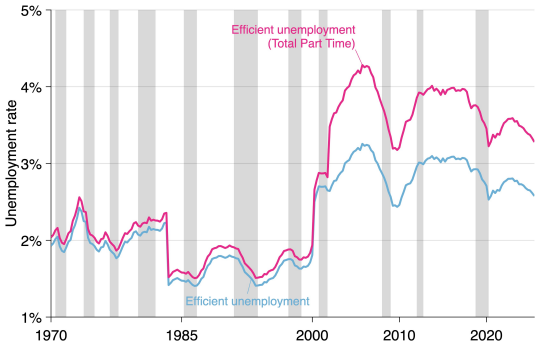
This tightening becomes much more apparent once part-time employment is incorporated into the analysis. Even though Japan appears inefficiently slack on average over the full sample, the periods surrounding the global financial crisis, the aftermath of the 2011 earthquake, and the post-pandemic recovery all exhibit episodes of labor-market tightness that persist into recent years. In several of these episodes, the full-time benchmark would continue to classify the labor market as relatively slack. Accounting for part-time employment therefore provides a more informative assessment of labor-market conditions.



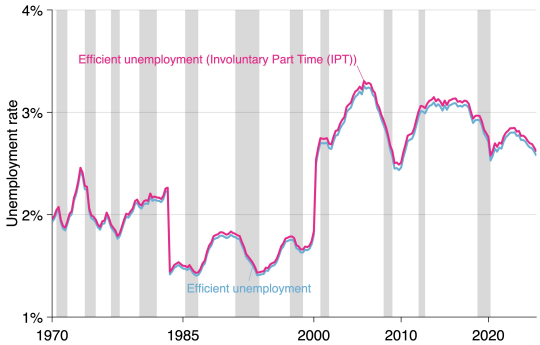
(a) Labor-market tightness: part-time ($\alpha = 0.78$) vs full-time ($\alpha = 1$)



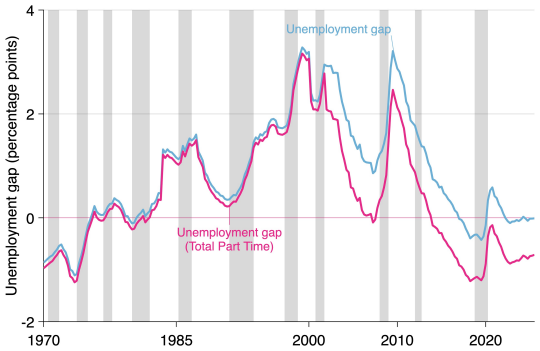
(b) Labor-market tightness: IPT ($\alpha = 0.94$) vs full-time ($\alpha = 1$)



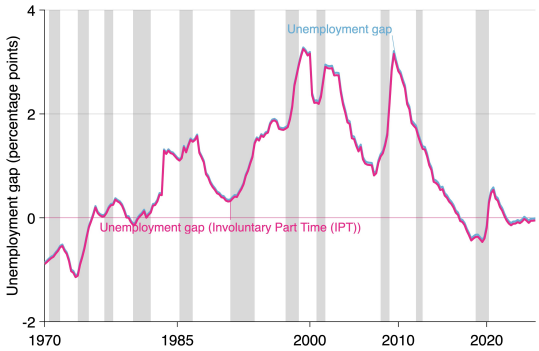
(c) Efficient unemployment rate: part-time ($\alpha = 0.78$) vs full-time ($\alpha = 1$)



(d) Efficient unemployment rate: IPT($\alpha = 0.94$) vs full-time ($\alpha = 1$)



(e) Unemployment gap: part-time ($\alpha = 0.78$) vs full-time ($\alpha = 1$)



(f) Unemployment gap: IPT ($\alpha = 0.94$) vs full-time ($\alpha = 1$)

Figure 1.27: Labor-market tightness, efficient unemployment rate, and unemployment gap in Japan, 1970–2025.

Notes: Each panel compares a part-time specification to the full-time benchmark ($\alpha = 1$): Left column: $\alpha = 0.78$ through 2001 (OECD part-time share from 2002 thereafter); right column: $\alpha = 0.94$ through 2001 (involuntary part-time share from OECD thereafter).

1.6.4 Involuntary part-time

Because hours for involuntary part-time workers are not available for Japan, I set $\gamma^{IP} = \gamma$ in this exercise. The comparison between total part-time employment and involuntary part-time employment therefore mainly reflects differences in employment composition, captured by α, rather than differences in relative hours. The involuntary part-time results should be interpreted as isolating the underemployment share of part-time employment under the same relative-hours assumption used for total part-time workers. To distinguish between the labor-utilization effect of total part-time employment and the underemployment effect of involuntary part-time employment, Figure 1.27 compares two alternative specifications. The left column incorporates total part-time employment, using $\alpha^P = 0.78$ before 2002 and the OECD part-time employment share from 2002 onward. The right column incorporates involuntary part-time employment, using $\alpha^{IP} = 0.94$ before 2002 and the OECD involuntary part-time employment share from 2002 onward.

The results differ substantially across the two specifications. When total part-time employment is incorporated, efficient labor-market tightness falls noticeably, the efficient unemployment rate rises, and the unemployment gap becomes considerably smaller relative to the full-time benchmark. In contrast, incorporating IPT produces only minor changes in all three measures. The estimated paths of efficient labor-market tightness, efficient unemployment, and the unemployment gap remain very close to those obtained under the full-time specification.

The difference in magnitude reflects the relative size of the two groups. As reported in Tables 1.6, 1.7, and 1.8, part-time workers account for 21.9 % of full-time and part-time employment on average, whereas involuntary part-time workers account for only 5.9 %. Moreover, involuntary part-time workers represent only 22.9 % of total part-time employment. Consequently, most part-time employment in

Japan is not classified as involuntary.

These findings suggest that the quantitative importance of part-time employment in Japan stems primarily from the hours margin rather than from hidden underemployment. The large effect of total part-time employment indicates that employment growth increasingly occurs through lower-hours jobs. By contrast, the relatively small effect of IPT implies that hidden labor-market slack plays a much more limited role. Therefore, the rise of part-time employment in Japan appears to reflect a structural shift in labor utilization rather than a large increase in workers who are unable to obtain full-time employment.

This pattern also helps explain why Japan can have a large and growing part-time sector without a correspondingly large increase in hidden slack. Much of the rise in part-time employment appears to reflect demographic and labor-supply changes, including greater participation by women and older workers in shorter-hour jobs, rather than workers being unable to obtain full-time employment.

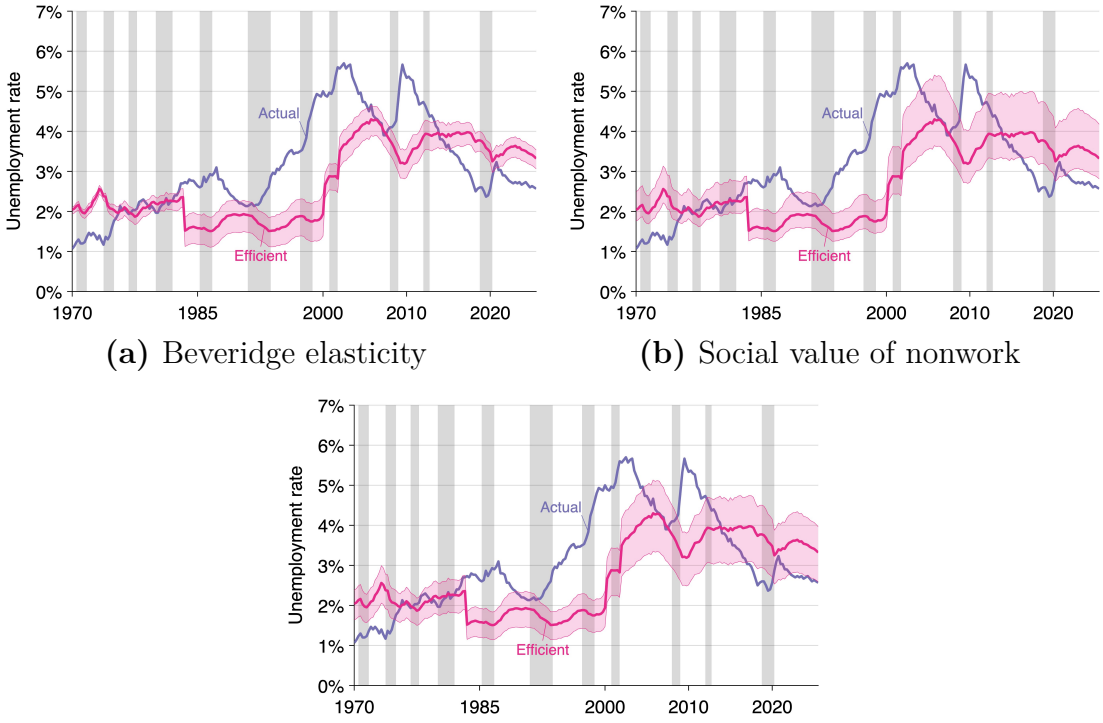
1.6.5 Inverse-optimum sufficient statistics

There remains uncertainty about the precise values of the sufficient statistics. Yet, Figure 1.28 shows that the efficient unemployment rate is remarkably stable across plausible calibrations of the Beveridge elasticity, the social value of nonwork, and recruiting costs. When the Beveridge elasticity is at the bottom (top) end of its 95% confidence interval, the efficient unemployment rate tracks the baseline series but is on average 0.33 %age point lower (0.28 %age point higher); at no point does it deviate by more than 0.55 percentage point from the baseline calibration, and the full envelope spans only 1.1% to 4.6% (Table A.4). The corresponding average shifts for alternative calibrations of the social value of nonwork are 0.41 and 0.69 percentage point, with a maximum deviation of 1.1 percent-

age points and an envelope of 1.3%–5.4% (Table A.5); for recruiting costs, they are 0.59 and 0.52 percentage point, with a maximum deviation of 0.9 percentage point and an envelope of 1.1%–5.1% (Table A.6). Although alternative parameter values shift the level of efficient unemployment, they do not materially alter the qualitative assessment of labor-market conditions.

Figure 1.29 provides further evidence of robustness. The inverse-optimum Beveridge elasticity, social value of nonwork, and recruiting cost frequently depart from their calibrated values, particularly during the Lost Decades and other periods of prolonged labor-market weakness. On average, the inverse-optimum Beveridge elasticity is 0.84, compared with a calibrated value of 0.63; it falls within the 95% confidence interval in only 24% of quarters and lies above the upper bound in 50% (Table A.7). The inverse-optimum social value of nonwork averages 0.27 and falls within the calibration range [0.03, 0.49] in 42% of quarters; it turns negative in tight-labor episodes, reaching -1.77 in 1973Q4, and exceeds the upper bound in 36% of quarters (Table A.8). The inverse-optimum recruiting cost averages 1.34, well above the calibrated $c = 0.92$, and lies within its calibration range in 46% of quarters while exceeding the upper bound in 40% (Table A.9). In many episodes, reconciling observed labor-market outcomes with efficiency would require a Beveridge elasticity substantially above its estimated value, an implausibly high social value of nonwork, or recruiting costs far exceeding their calibrated level.

Taken together, the results indicate that Japan's labor market was frequently inefficiently slack during the Lost Decades, but moved closer to efficiency after 2011 and appears increasingly tight in the post-pandemic period.

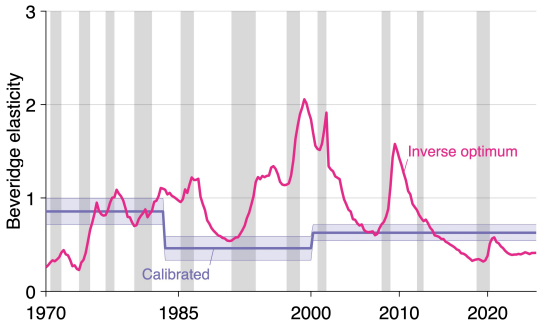


(a) Beveridge elasticity

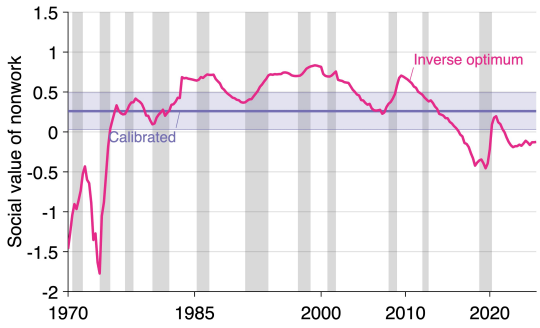
(b) Social value of nonwork

(c) Recruiting cost

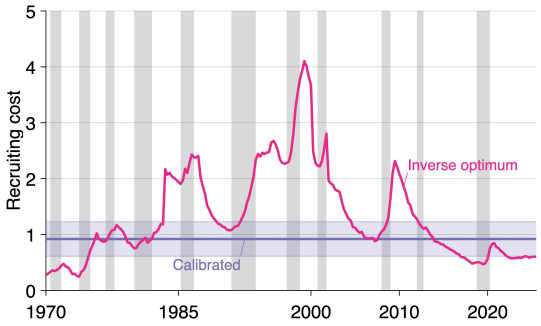
Figure 1.28: Efficient unemployment rates for alternative sufficient statistics in Japan, 1970–2025.



(a) Inverse-optimum Beveridge elasticity



(b) Inverse-optimum social value of nonwork



(c) Inverse-optimum recruiting cost

Figure 1.29: Inverse-optimum sufficient statistics in Japan, 1970–2025.

1.7 Conclusion

This paper extends the efficient-unemployment framework of Michaillat and Saez (2021a) by allowing employed workers to differ in hours supplied. The original framework treats all employed workers as homogeneous. I extend the sufficient-statistics formula through two additional objects, the employment-composition parameter α and the relative-hours parameter γ. When all workers are full-time, so that $\alpha = 1$, the extended formula collapses to the original Michaillat–Saez expression, providing a consistency check on the extension. The same welfare function is used throughout; the total and involuntary part-time calibrations differ only in how employment is partitioned when measuring α and γ.

Applying the extended framework to the United States from 1951 to 2026 shows that incorporating involuntary part-time employment produces only limited changes in efficient unemployment and the unemployment gap. The resulting series remain close to the baseline estimates, reflecting the relatively small role of involuntary part-time employment in the U.S. labor market. Incorporating total part-time employment has more important implications. On average, the unemployment gap remains positive and strongly countercyclical, so the broad assessment of labor-market slack is unchanged. At the same time, the total part-time calibration reveals a different cyclical pattern around major downturns. Before the 1973–75, early-1980s, 1990–91, and 2007–09 recessions, actual labor-market tightness approaches or reaches the efficient allocation once differences in hours worked are taken into account. Under the full-time benchmark, these same episodes instead appear inefficiently slack. Accounting for part-time employment therefore changes the interpretation of labor-market conditions around several recession turning points, highlighting the importance of the hours margin when assessing cyclical conditions.

The Japanese evidence points to the same mechanism but with quantitatively larger effects. As in the United States, incorporating involuntary part-time employment produces only limited changes in efficient unemployment and unemployment gaps. By contrast, incorporating total part-time employment has substantial effects on the efficient benchmark. The average efficient unemployment rate rises, efficient labor-market tightness falls, and the unemployment gap becomes substantially smaller than under the full-time benchmark. Although Japan remains inefficiently slack on average over the full sample, accounting for part-time employment changes the interpretation of several important episodes. Before the global financial crisis, the Japanese labor market appears close to efficient under

the part-time calibration. After the 2011 earthquake, the part-time calibration reveals a tighter labor market, and during the post-pandemic period it again indicates increasing tightness. Thus, while the U.S. results mainly alter the interpretation of conditions around recession turning points, the Japanese results reveal extended periods of labor-market tightness that are much less visible under the full-time benchmark.

Additional exercises using involuntary part-time employment indicate that these results are not primarily driven by underemployment. Involuntary part-time workers account for only a small share of total part-time employment in both countries, and a particularly small and declining share in Japan. Restricting the part-time margin to this group therefore yields estimates that remain close to the full-time benchmark.

This exercise should be interpreted cautiously. Because hours worked by involuntary part-time workers are unavailable for Japan, I set $\gamma^{IP} = \gamma$. The difference between the total and involuntary part-time results therefore reflects mainly the employment-share statistic α, rather than a separately measured hours margin.

Overall, the results suggest that accounting for employment heterogeneity can materially affect the measurement of labor-market efficiency. In both countries, incorporating total part-time employment raises efficient unemployment and lowers efficient labor-market tightness because part-time workers supply fewer effective hours than full-time workers. The effect is modest for involuntary part-time employment, indicating that underemployment alone is not the primary mechanism. Instead, the main channel operates through differences in labor input across workers. More generally, the findings highlight the importance of measuring labor input not only by the number of workers employed but also by the amount of work they perform.

Chapter 2

Recession Detection in Japan Using Labor Market Data

2.1 Introduction

As emphasized by Hamilton (2011), policymakers, firms, and households are often more concerned with identifying when the economy is turning than with measuring the precise magnitude of GDP growth. Economic turning points mark regime changes in the overall state of the economy, signaling transitions between expansion and recession. These transitions are associated with broad declines in output, weakening labor market conditions, reduced demand faced by firms, and rising economic uncertainty. Because of this importance, the study of recession detection has been extensively developed in the United States.

This paper aims to extend these labor-market-based recession detection methods to a non-U.S. context by asking the question: can labor-market-based recession detection rules be effectively applied to Japan? While many recession detection methods are developed and evaluated primarily in the U.S. context, it

remains unclear whether such methods generalize well to economies with different labor market structures, institutional settings, and business cycle dynamics. Japan provides an especially informative setting because it combines a large advanced economy with detailed and consistently measured labor market data spanning several decades. At the same time, Japan exhibits distinct labor market characteristics and unusually long delays in official recession dating by the Economic and Social Research Institute (ESRI), making timely recession detection particularly informative.

The timely identification of recessions remains a central challenge in Japan. Using the business cycle chronology from the ESRI, we compare recession start dates (business cycle peaks) to their official announcement dates and document large and systematic delays. On average, recessions are announced 26.5 months after their onset, with a standard deviation of about 10 months (Table 2.1). As a result, policymakers and markets often operate for extended periods without reliable confirmation that the economy has entered a recession, limiting the scope for timely policy responses. This issue is particularly important in Japan, where three decades of secular stagnation and relatively muted fluctuations in unemployment make recession detection both more difficult and more consequential. At the same time, Japan remains exposed to recession risk. In 2025, Japan narrowly avoided a technical recession, commonly defined as two consecutive quarters of negative GDP growth, highlighting both the difficulty of real-time assessment and the economy's continuing vulnerability to downturns.

To illustrate the economic meaning of the Japanese recession dates, Figure 2.1 plots HP-filtered real GDP together with the official recession dates announced by the ESRI. The figure shows that the recession dates generally coincide with sharp declines in detrended economic activity. This is important because some audiences

may be unfamiliar with the Japanese recession dating system or uncertain about what these recession dates are intended to measure. The HP-filtered GDP series demonstrates that the official recession dates correspond to meaningful turning points in aggregate economic activity rather than arbitrary classifications. In other words, the recession dates capture periods in which the Japanese economy experiences broad and significant slowdowns relative to trend.

A natural benchmark for real-time recession detection is the Sahm Rule, which signals a recession when the three-month moving average of the unemployment rate rises sufficiently above its recent minimum. While the rule has performed well in the United States, it relies solely on unemployment as an indicator of labor market conditions. This may present limitations in Japan, where unemployment fluctuations are substantially more muted than in more volatile economies such as the United States. As shown in Figures 2.2 and 2.3, U.S. recessions are characterized by sharp increases in unemployment and declines in vacancies, whereas Japanese recessions exhibit much smaller movements in unemployment. In contrast, vacancy rates in Japan display stronger cyclical variation and may better capture fluctuations in labor demand that are not fully reflected in unemployment data.

Recent work by Pascal Michaillat and Emmanuel Saez (2025) proposes incorporating both unemployment and vacancy rates into recession detection through a combined unemployment–vacancy indicator, referred to as the Michez Rule. By integrating information from both labor supply and labor demand conditions, the approach aims to improve the timeliness and accuracy of recession detection.

Building on Michaillat (2025), this paper shows that labor-market recession rules can be successfully adapted to Japan. We utilize monthly Japanese unemployment data from FRED, monthly job vacancy rate data from Japan's Ministry

of Health, Labour and Welfare, and monthly official recession dates from the ESRI. We construct 95,832 recession indicators using combinations of unemployment and vacancy data with alternative smoothing methods, transformations, and thresholds. Among these, 193 classifiers are perfect classifiers that lie on the anticipation–precision frontier. They produce no false negatives (actual recessions missed by the classifiers) and no false positives (non-recessions mistakenly detected by the classifiers). The selected classifier ensemble signals recessions on average 0.06 months after their true onset with a standard deviation of detection errors of 2.72 months. The classifier ensemble correctly identifies all 11 Japanese recessions between 1970 and 2021 without any false positives. These findings are consistent with the evidence for the United States reported by Michaillat (2025), where the classifier ensemble detected recession starts on average 1.2 months after onset between 1979 and 2021, compared to 6.3 months for the National Bureau of Economic Research (NBER) Business Cycle Dating Committee. Overall, the results suggest that slack-based labor-market indicators provide a robust framework for real-time recession detection outside the United States.

2.2 Literature Review

This paper contributes to several strands of the recession-detection literature. First, it contributes to the literature on real-time recession dating and business-cycle turning points. Hamilton (2011) studies the difficulty of identifying recession turning points in real-time, emphasizing challenges such as data revisions, structural change, and noisy macroeconomic signals. The paper also highlights the trade-off between using simple and robust recession indicators versus incorporating large information sets to improve turning-point detection. This trade-off motivates our approach, which focuses on extracting recession signals from a small set

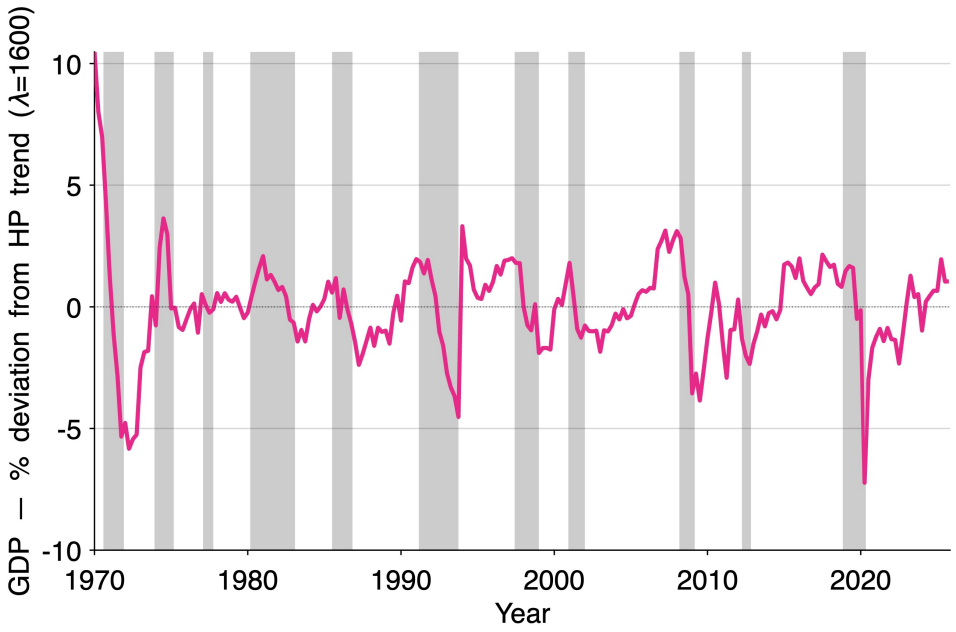


Figure 2.1: Japanese real GDP deviations from trend, 1970–2025

Notes: The figure plots the percentage deviation of Japanese real GDP from its Hodrick–Prescott (HP) trend with smoothing parameter $\lambda = 1600$. Grey shaded regions indicate recessions dated by the Economic and Social Research Institute (ESRI). Peaks and troughs in GDP relative to trend closely align with ESRI business-cycle turning points, highlighting the cyclical fluctuations associated with Japanese recessions.

Table 2.1: ESRI Recession Start Announcements

Cycle	Peak (Start Date)	Announcement Date	Delay (months)
11th	Feb 1991	Oct 1993	∼32
12th	May 1997	Jul 2000	∼38
13th	Nov 2000	May 2002	∼18
14th	Feb 2008	Jul 2009	∼17
15th	Mar 2012	Aug 2013	∼17
16th	Oct 2018	Nov 9, 2021	∼37
Mean			**26.5**
Std. Dev.			**10.25**

Notes: The table reports the delay between the official start dates of Japanese recessions, defined by ESRI business-cycle peaks, and the dates on which ESRI publicly announced these turning points. Delays are measured in months between the recession start date and the corresponding announcement date. The results highlight the substantial lag in official recession dating, motivating the need for timely recession detection methods.

of labor-market indicators rather than relying on large macroeconomic datasets.

A large part of the literature attempts to improve recession detection by incorporating increasingly large sets of macroeconomic indicators. These approaches typically rely on dynamic factor models or large collections of macroeconomic indicators. For example, Stock and Watson (2014) use large information sets to estimate business-cycle turning points, while Chen et al. (2011) and Fornaro (2016) similarly rely on many macroeconomic series to improve recession prediction. In contrast, another strand of the literature emphasizes that relatively simple recession rules can perform remarkably well in practice. Piger (2019) shows that narrow sets of carefully selected variables can provide strong classification performance, while Sahm (2019) proposes the Sahm Rule, a simple labor-market-based recession indicator using unemployment dynamics.

More recent work focuses on improving recession detection by combining multiple measures of labor-market slack and systematically searching across large model spaces to identify robust recession indicators. **?** introduce the Michez Rule, which combines unemployment and vacancy data to reduce noise in recession signals and improve real-time recession detection in U.S. data. Relatedly, Michaillat (2025) shows that recession detection can be improved by systematically searching across alternative smoothing methods, transformations, and thresholds to construct optimized recession classifiers from large model spaces.

This paper contributes to this literature by examining whether labor-market-based recession rules generalize outside the United States. Japan provides a useful setting because unemployment fluctuations are relatively muted while vacancy rates display stronger cyclical movements around business-cycle turning points. In addition, Japanese vacancy data may provide a more representative measure of labor-market conditions than vacancy data available for the United States, making

Japan a particularly informative environment for evaluating vacancy-based recession indicators. We therefore assess whether vacancy-augmented recession indicators retain predictive power in a labor market where unemployment alone may provide weaker recession signals. More broadly, the paper studies whether labor-market-based recession rules provide a general framework for real-time turning-point detection across countries.

2.3 Data

This section presents the data on Japanese recessions, unemployment, and job vacancies used in the paper. The data coverage is from January 1970 to December 2025.

2.3.1 Japan Time Series Data

Our analysis uses monthly unemployment, vacancy, employment, and business-cycle dating data for Japan spanning 1970 to 2025. The unemployment data are obtained from the Labour Force Survey conducted by the Statistics Bureau of Japan, while vacancy and employment data are taken from Labour Market Indicators (General, including Part-time). Business-cycle reference dates and GDP data are provided by the Economic and Social Research Institute (ESRI). Using the vacancy and employment data, we construct a monthly vacancy rate series.

2.3.2 Business Cycle Dating

Table 2.2 reports the recession start dates based on the chronology provided by the Economic and Social Research Institute. ESRI recession dates are widely treated as the official business-cycle chronology in Japan and are routinely refer-

enced in financial press coverage discussing recession risk (Economic and Social Research Institute (2014); Japan Times (2025)). Over our sample period, this chronology identifies 11 recession episodes. These dates serve as the ground truth for evaluating our recession indicators. Our objective is to assess how early and how accurately these indicators detect recession onsets in real-time.

To fix ideas, ESRI's role in Japan is broadly analogous to that of the **?** (NBER) in the United States. Similar to the NBER, ESRI identifies business-cycle turning points using a broad range of economic indicators rather than relying on a single variable. ESRI constructs a monthly coincident composite index designed to track overall economic activity using indicators related to production, employment, consumption, investment, retail sales, exports, and other measures of economic conditions. ESRI then determines business-cycle peaks and troughs by examining sustained movements in this composite index together with broader economic conditions (Economic and Social Research Institute (2014)).

Although ESRI's approach is conceptually similar to the NBER's business-cycle dating methodology, the two institutions differ in implementation. ESRI places greater emphasis on a formally constructed coincident composite index, whereas the NBER evaluates several macroeconomic indicators individually when determining turning points. Recent reporting has similarly emphasized the central role of the coincident index in ESRI's recession dating process (Japan Times (2025)).

2.3.3 GDP Data

Japan's official GDP data are produced by the ESRI through the Quarterly Estimates of GDP. The GDP series are reported at the quarterly frequency and include nominal and real measures under the expenditure approach to national

Table 2.2: Start Dates of Japanese Recessions (1970-2025)

No.	Episode	Start Month
1	Recession 1	Jul. 1970
2	Recession 2	Nov. 1973
3	Recession 3	Jan. 1977
4	Recession 4	Feb. 1980
5	Recession 5	Jun. 1985
6	Recession 6	Feb. 1991
7	Recession 7	May 1997
8	Recession 8	Nov. 2000
9	Recession 9	Feb. 2008
10	Recession 10	Mar. 2012
11	Recession 11	Oct. 2018

Notes: The table reports the start dates of Japanese recessions identified by the ESRI. ESRI determines business-cycle turning points using a broad range of macroeconomic indicators and a coincident composite index designed to track overall economic activity. By convention, recession start dates correspond to business-cycle peaks.

accounts. In this paper, GDP data are used as a broad measure of aggregate economic activity and for comparison with officially dated business-cycle turning points.

2.3.4 Unemployment and Vacancy Data

The unemployment rate data are obtained from the OECD Main Economic Indicators database through the Federal Reserve Economic Data (FRED) system and are based on Japan's Labour Force Survey conducted by the Statistics Bureau of Japan. The Labour Force Survey is a monthly household survey designed to measure employment and unemployment conditions in Japan. Following OECD and International Labour Organization (ILO) definitions, individuals are classified as unemployed if they are without work, available for work, and actively seeking

employment. The unemployment rate is measured as the number of unemployed persons divided by the labor force and is reported as a monthly seasonally adjusted series for individuals aged 15–64.

In Japan, the Ministry of Health, Labour and Welfare compiles monthly statistics on job openings, job seekers, and employment placements through the Public Employment Security Offices (Hello Work) system, and publishes indicators such as the job openings-to-applicants ratio as part of the General Employment Placement Situation statistics. In this paper, we use seasonally adjusted monthly data on the total number of active job openings (monthly active vacancies) obtained from the official e-Stat database. We also obtain seasonally adjusted monthly labor force data from e-Stat. The vacancy rate is then constructed as the total number of active job openings divided by the total labor force.

2.3.5 Labor Market Dynamics

To provide a reference point for the Japanese analysis, we first examine labor market dynamics in the United States, where recession detection rules such as the Sahm and Michez rules were originally developed and evaluated. Figure 2.2 plots the unemployment rate and the vacancy rate over time, with shaded regions indicating recessions dated by the National Bureau of Economic Research. A clear pattern emerges: unemployment rises during recessions, while vacancies decline, implying that the two series move in opposite directions over the business cycle.

We next examine labor market dynamics in Japan. Figure 2.3 plots the unemployment rate and the vacancy rate over time, with shaded regions indicating recessions dated by the Economic and Social Research Institute. As in the United States, unemployment tends to rise during recessions, while vacancies decline, implying that the two series move in opposite directions over the business cycle.

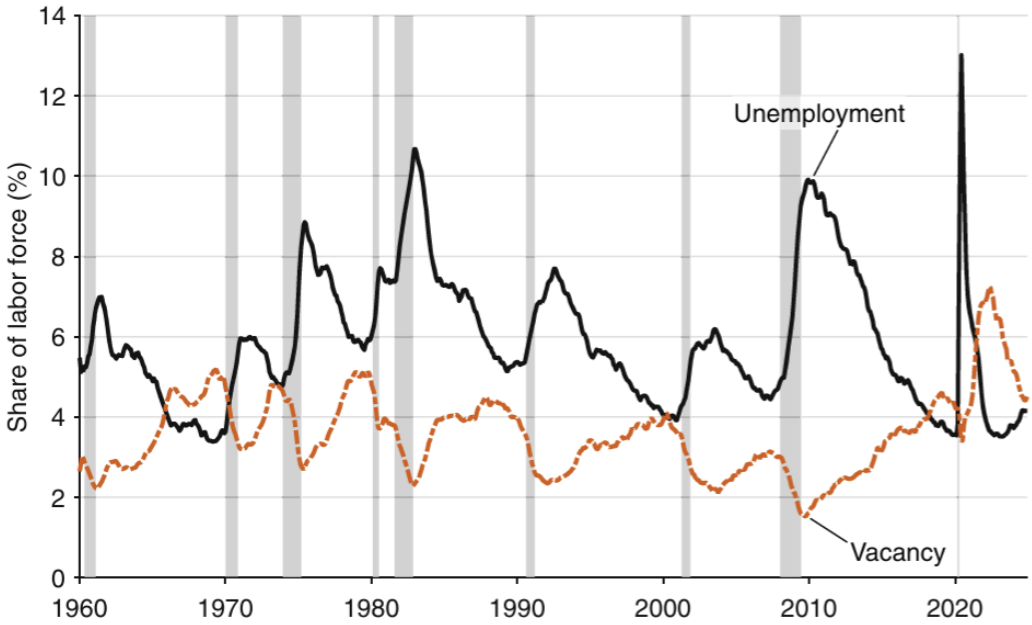


Figure 2.2: United States: Unemployment and Vacancy Rate

Notes: U.S. unemployment and vacancy rates, January 1960–December 2024. This figure is reproduced from Michaillat (2025). The unemployment rate (solid black line) is based on BLS data, and the vacancy rate (orange dashed line) combines estimates from **?** and the BLS. Both series are shown as three-month moving averages of monthly data. Shaded areas indicate recessions as dated by the NBER. We use this figure as a benchmark for comparison with our analysis.

However, these patterns are less pronounced in Japan. In particular, fluctuations in unemployment are more gradual, and signals appear less sharp around some recession episodes. This attenuation suggests that recession detection using labor market data alone may be more challenging in the Japanese context. At the same time, the underlying economic intuition remains: vacancies reflect labor demand, while unemployment captures labor market slack. Accordingly, combining these two indicators may yield a more informative signal for recession detection in Japan.

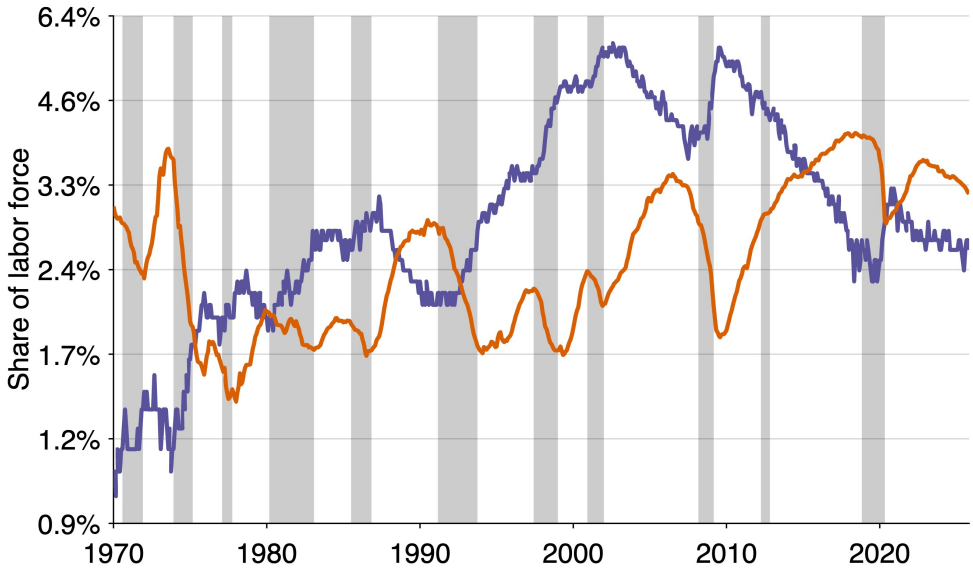


Figure 2.3: Japan: Unemployment Rate and Vacancy Rate

Notes: Monthly Japanese unemployment and vacancy rates from January 1970 to December 2025. The unemployment rate (solid blue line) is constructed using data from FRED, while the vacancy rate (orange line) is computed using Japan's labor force survey. Both series are expressed as three-month moving averages of monthly observations. Shaded areas indicate recessions dated by the Economic and Social Research Institute (ESRI).

2.4 Background: Sahm Rule and Michez Rule in Japan

This section presents the empirical results for recession detection in Japan using labor-market-based indicators. We first implement the Sahm and Michez rules in the Japanese context. While the Sahm rule tracks increases in unemployment relative to recent lows, the Michez rule combines unemployment and vacancy data to construct a less noisy recession indicator. We then evaluate the ability of these indicators to identify historical recession episodes in real-time.

2.4.1 Sahm Rule in Japan

First, we smooth the unemployment rate $u(t)$ by taking a three-month moving average:

$$\bar{u}(t) = \frac{u(t) + u(t-1) + u(t-2)}{3}. \tag{2.1}$$

This step reduces short-term fluctuations and isolates persistent movements in unemployment.

Second, we measure the increase in the smoothed unemployment rate relative to its lowest value over the preceding 12 months:

$$\tilde{u}(t) = \bar{u}(t) - \min_{0 \le s \le 12} \{\bar{u}(t-s)\}. \tag{2.2}$$

This transformation captures how much unemployment has risen from its recent trough. By construction, the indicator remains near zero when unemployment is stable or declining, and becomes positive when unemployment begins to increase.

A recession is identified when the indicator exceeds 0.5 percentage points:

$$\tilde{u}(t-1) < 0.5 \quad \text{and} \quad \tilde{u}(t) \ge 0.5 \tag{2.3}$$

Figure 2.4 shows that the Sahm indicator performs well in the United States. The horizontal line marks the 0.5 percentage point threshold used to signal a recession. The indicator rises sharply during recessions and typically crosses the threshold in these periods, successfully identifying downturns. However, because the rule relies entirely on unemployment, the signal can be noisy and occasionally misleading. In particular, the indicator generates false positives in several historical episodes, including 1934, when it peaked at 3.98 percentage points despite

no recession, as well as in 1947 and 1959, when it crossed the threshold without an associated downturn. Even in more recent data, smaller fluctuations, such as those around 2003, do not correspond to recessions. These patterns suggest that, while the Sahm Rule is simple and informative, reliance on a single labor market indicator limits precision around turning points.

The Sahm indicator performs poorly in Japan (Figure 2.5). It misses several recessions prior to 1990, detects the 1990s downturn with a delay, and fails to capture episodes in the late 1970s and around 2012. Overall, it identifies only 4 of 11 recessions, less than half, indicating that a direct application of the Sahm rule does not translate well to the Japanese context.

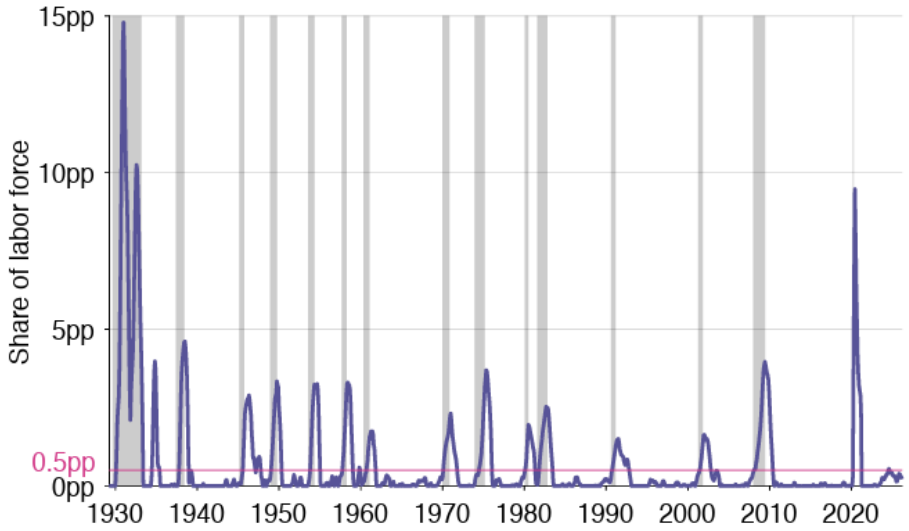


Figure 2.4: Sahm Rule Results for the United States

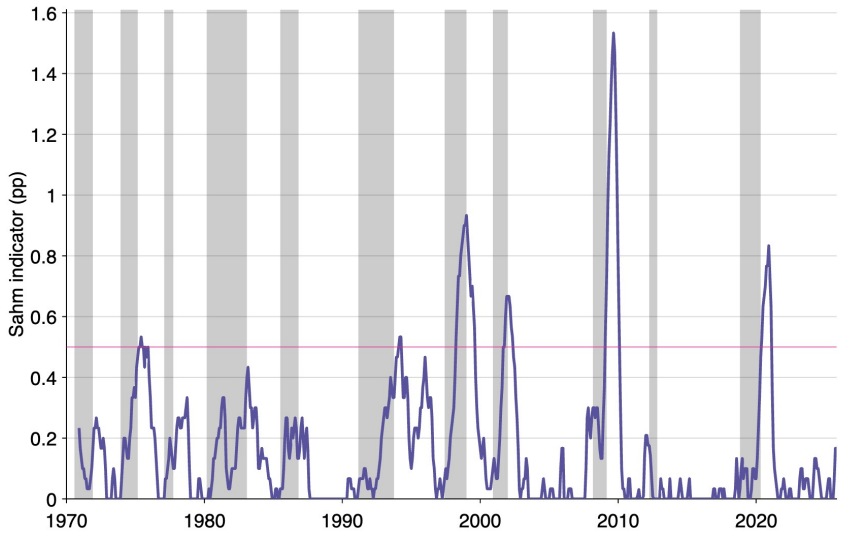


Figure 2.5: Sahm Rule Results for Japan

Notes: The top panel plots the Sahm indicator for the United States, defined as the increase in the three-month moving average of the unemployment rate relative to its twelve-month minimum, using data from the Bureau of Labor Statistics from 1930 to the present. Shaded areas denote recessions dated by the NBER. The recession threshold is 0.5 percentage points.

The bottom panel plots the Sahm indicator for Japan defined as the increase in the three-month moving average of the unemployment rate relative to its twelve-month minimum, using unemployment data from FRED spanning 1970–2025. Shaded areas denote recessions dated by the ESRI. The recession threshold is 0.5 percentage points.

2.4.2 Michez Rule in Japan

We implement the Michez rule, which combines unemployment and vacancy data to produce a less noisy recession indicator.

First, we construct the vacancy-based component. Let $v(t)$ denote the vacancy rate and define its three-month moving average:

$$\bar{v}(t) = \frac{v(t) + v(t-1) + v(t-2)}{3}. \tag{2.4}$$

We then measure the decline in the smoothed vacancy rate relative to its maximum over the preceding 12 months:

$$\tilde{v}(t) = \max_{0 \le s \le 12} \bar{v}(t-s) - \bar{v}(t). \tag{2.5}$$

Let $\tilde{u}(t)$ denote the unemployment-based indicator defined in Section 2.4.1.

The Michez rule combines these two indicators by taking their minimum as shown in Figure 2.6:

$$m(t) = \min\{\tilde{u}(t), \tilde{v}(t)\}. \tag{2.6}$$

The Michez rule further improves recession detection by incorporating vacancy data alongside unemployment. Following the original U.S. application, we implement the standard Michez rule using a fixed threshold of 0.29, such that recessions are detected whenever the indicator crosses the threshold. As shown in Figure 2.7, the minimum indicator closely tracks recession periods, with threshold crossings occurring near the onset of most downturns. Relative to the Sahm rule, the Michez rule generates fewer false positives, with the indicator remaining near zero during expansions and exhibiting fewer spurious spikes outside recession periods.

This improvement reflects the ability of the Michez rule to filter out transient

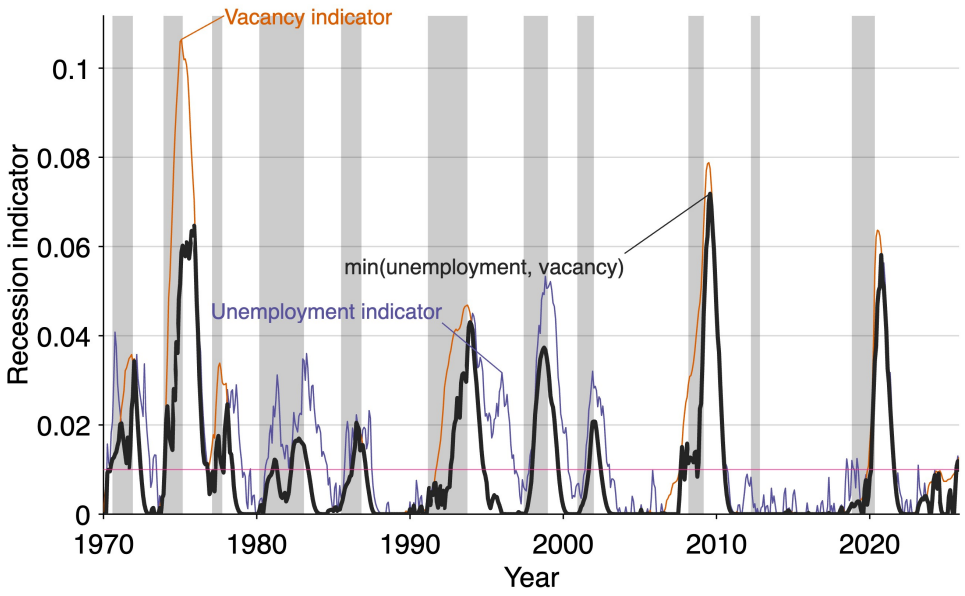


Figure 2.6: Michez Rule: Minimum of Two Indicators

Notes: The figure plots three recession indicators for Japan from 1970 onward. The blue line shows the unemployment-based indicator, measured as the increase in the unemployment rate relative to its recent minimum. The orange dashed line shows the vacancy-based indicator, measured as the decline in the vacancy rate relative to its recent maximum. The black line represents the Michez rule, defined as the minimum of the two indicators. Gray shaded regions indicate recessions dated by the Economic and Social Research Institute (ESRI).

labor market fluctuations by requiring both rising unemployment and declining vacancies. Incorporating vacancy data yields a more accurate and less noisy recession signal, improving the tradeoff between false positives and false negatives.

In our baseline specification, this fixed-threshold rule serves as a benchmark. We then extend the analysis by replacing the ad hoc threshold with the probabilistic classifier selection and aggregation procedure described above.

2.5 Construction of the Recession Indicators

Using unemployment and vacancy data, we construct 95,832 recession indicators in this section. The construction process consists of smoothing the data,

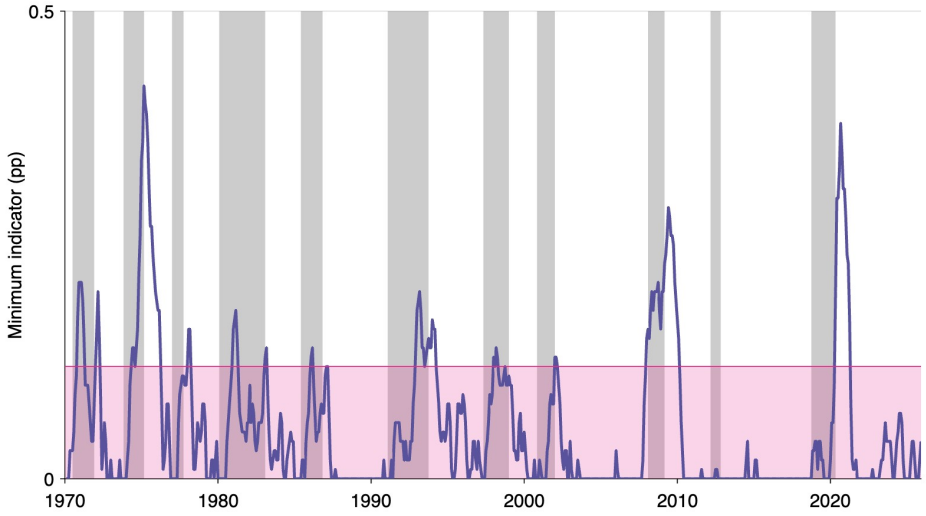


Figure 2.7: Michez Rule Results

Notes: The figure plots the Michez minimum indicator over time using Japanese labor-market data. Gray shaded regions indicate recessions dated by the Economic and Social Research Institute (ESRI). The horizontal line represents the recession threshold of 0.29. The Michez rule combines unemployment and vacancy indicators by taking their minimum, following Michaillat and Saez (2025). Labor-market data are obtained from the Labor Force Survey of the Statistics Bureau of Japan.

identifying turning points, scaling the resulting movements, and combining the single-variable measures into composite indicators. In the next section, we pair these indicators with thresholds to detect recession episodes.

2.5.1 Smoothing the Data

To construct the recession indicators, we first smooth the unemployment and vacancy series using both simple moving averages and exponentially weighted moving averages, applying the approach in Michaillat (2025). We begin by defining trailing moving averages for the unemployment and vacancy rate. In both equations $\alpha = 0, 1, \ldots, 11$ represents the smoothing window. For example, if $\alpha = 0$, the series is unchanged; if $\alpha = 11$, the formula produces a 12-month trailing average. The standard Sahm and Michez rule specifications correspond to $\alpha = 2$, which yields a 3-month moving average.

$$\bar{u}(t) = \frac{\sum_{k=0}^{\alpha} u(t-k)}{\alpha + 1}, \tag{2.7}$$

$$\bar{v}(t) = \frac{\sum_{k=0}^{\alpha} v(t-k)}{\alpha + 1}. \tag{2.8}$$

In addition, we examine exponentially weighted moving averages for the unemployment and vacancy rate where $\alpha \in [0.1, 1]$ determines the degree of persistence in the series. A value of $\alpha = 1$ implies no smoothing, while lower values of α generate smoother series.

$$\bar{u}(t) = \alpha u(t) + (1 - \alpha)\bar{u}(t-1), \tag{2.9}$$

$$\bar{v}(t) = \alpha v(t) + (1 - \alpha)\bar{v}(t-1). \tag{2.10}$$

2.5.2 Detecting Turning Points

To identify turning points in labor market conditions, we compare current unemployment and vacancy rates to their recent extrema over a trailing window, following Michaillat (2025). We define the trailing minimum unemployment rate and trailing maximum vacancy rate as:

$$u^{\min}(t) = \min_{0 \le k \le \beta} \bar{u}(t-k), \tag{2.11}$$

$$v^{\max}(t) = \max_{0 \le k \le \beta} \bar{v}(t-k), \tag{2.12}$$

where $\beta = 1, 2, \ldots, 18$ specifies the length of the trailing window. We then construct the unemployment and vacancy indicators as

$$\tilde{u}(t) = \bar{u}(t) - u^{\min}(t), \tag{2.13}$$

$$\tilde{v}(t) = v^{\max}(t) - \bar{v}(t). \tag{2.14}$$

The unemployment indicator measures how much unemployment has increased relative to recent lows, while the vacancy indicator measures how much vacancies have fallen relative to recent highs. These transformations are motivated by the fact that recessions are typically associated with rising unemployment and declining vacancies. Larger values of $\tilde{u}(t)$ and $\tilde{v}(t)$ therefore correspond to greater labor market deterioration.

2.5.3 Scaling Variations

Following Michaillat (2025), we consider alternative ways of measuring changes in labor market conditions. Since recessions may depend on either level or proportional changes in unemployment and vacancy rates, we apply a Box–Cox transformation to construct alternative indicators.

For unemployment, we define

$$\hat{u}(t) = \frac{[u(t)]^\gamma - [u^{\min}(t)]^\gamma}{\gamma} \simeq \frac{\tilde{u}(t)}{u^{\min}(t)^{1-\gamma}}, \tag{2.15}$$

where $\gamma \in \{0, 0.1, \dots, 1\}$ controls the scaling of the transformation. The parameter γ determines whether the indicator emphasizes absolute or proportional changes in unemployment. When $\gamma = 1$, the transformation reduces to level changes in unemployment, while $\gamma \to 0$ approximates percentage or log changes. Intermediate values of γ generate indicators that lie between these two cases.

As $\gamma \to 0$, the transformation becomes

$$\hat{u}(t) = \log\left(\frac{u(t)}{u^{\min}(t)}\right) \simeq \frac{\tilde{u}(t)}{u^{\min}(t)}, \tag{2.16}$$

which captures proportional changes in unemployment. When $\gamma = 1$, the transformation simplifies to

$$\hat{u}(t) = \tilde{u}(t), \tag{2.17}$$

which measures level changes in unemployment.

Similarly, for vacancies, we define

$$\hat{v}(t) = \frac{[v^{\max}(t)]^\gamma - [\bar{v}(t)]^\gamma}{\gamma} \simeq \frac{\tilde{v}(t)}{[v^{\max}(t)]^{1-\gamma}} \tag{2.18}$$

where γ has the same interpretation as in the unemployment transformation. The vacancy specification differs slightly from the unemployment specification because recessions are associated with declines in vacancies relative to a recent maximum, rather than increases in unemployment relative to a recent minimum. Lower values of γ place greater emphasis on proportional declines in vacancies, while $\gamma = 1$ corresponds to level declines.

2.5.4 Combining Single-Variable Indicators

Finally, we consider recession indicators based on unemployment alone, vacancies alone, and weighted combinations of the two measures constructed using alternative weighting schemes. $\delta = 0, 0.1, \ldots, 1$ determines the weight placed on the unemployment indicator. The Sahm Rule uses the unemployment indicator, corresponding to $\delta = 1$.

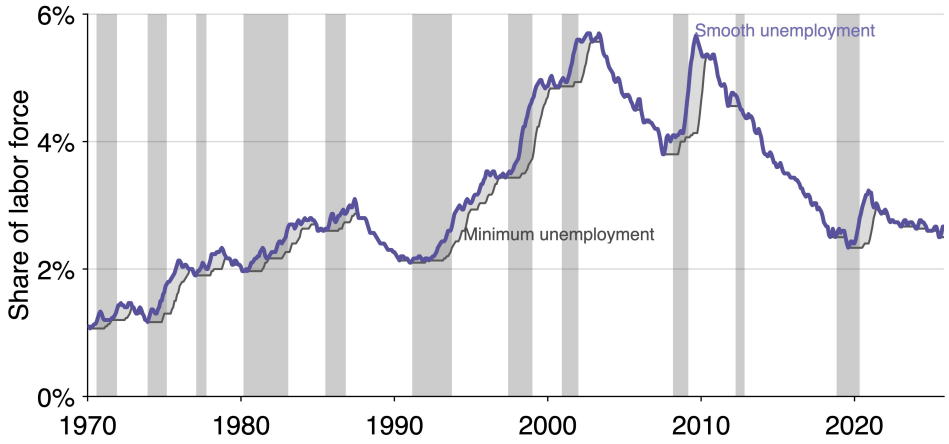
$$i(t) = \delta \hat{u}(t) + (1 - \delta)\hat{v}(t), \tag{2.19}$$

Following **?**, we also construct indicators based on the minimum and maximum of the unemployment and vacancy signals:

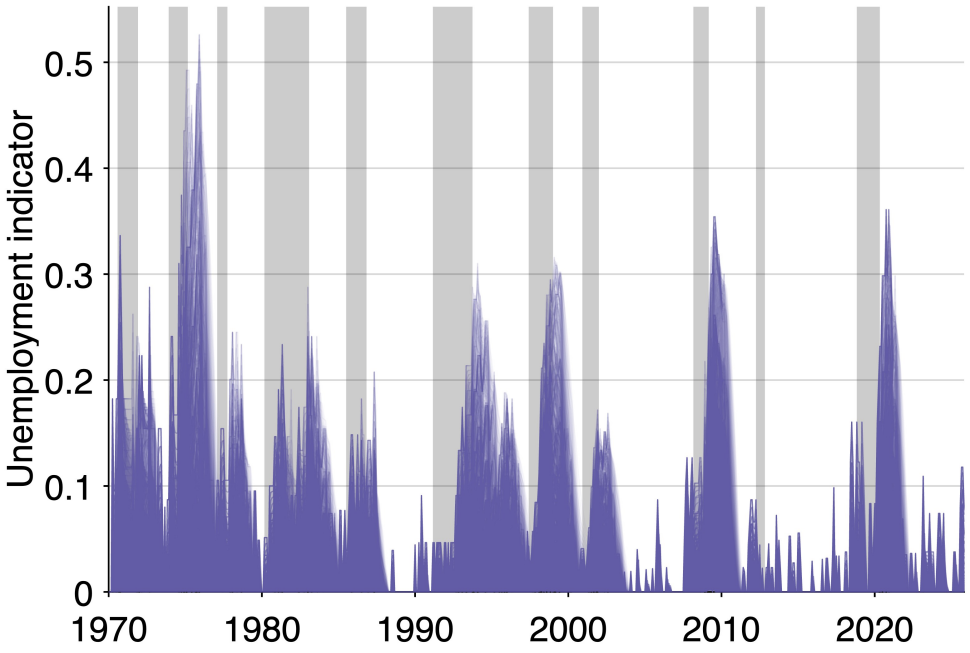
$$i(t) = \delta \min(\hat{u}(t), \hat{v}(t)) + (1 - \delta) \max(\hat{u}(t), \hat{v}(t)), \tag{2.20}$$

where $\delta = 0, 0.1, \ldots, 1$ determines the weight on the minimum indicator. The Michez rule uses the minimum indicator, corresponding to $\delta = 1$.

Rather than imposing a specific functional form, we consider a broad class of recession indicators constructed from unemployment and vacancy data. These indicators allow for different ways of measuring changes in labor market conditions, ranging from level changes to percentage changes through Box–Cox transforma-



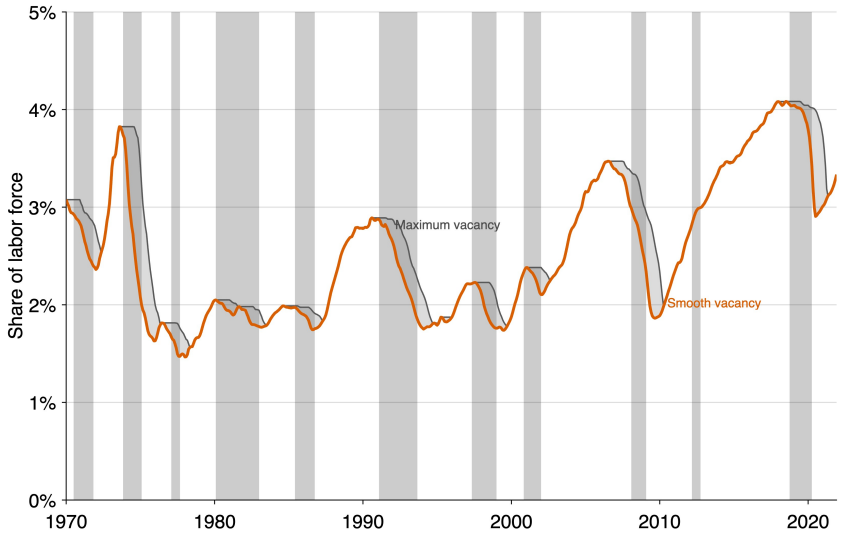
(a) Trailing Minimum of the Unemployment Rate



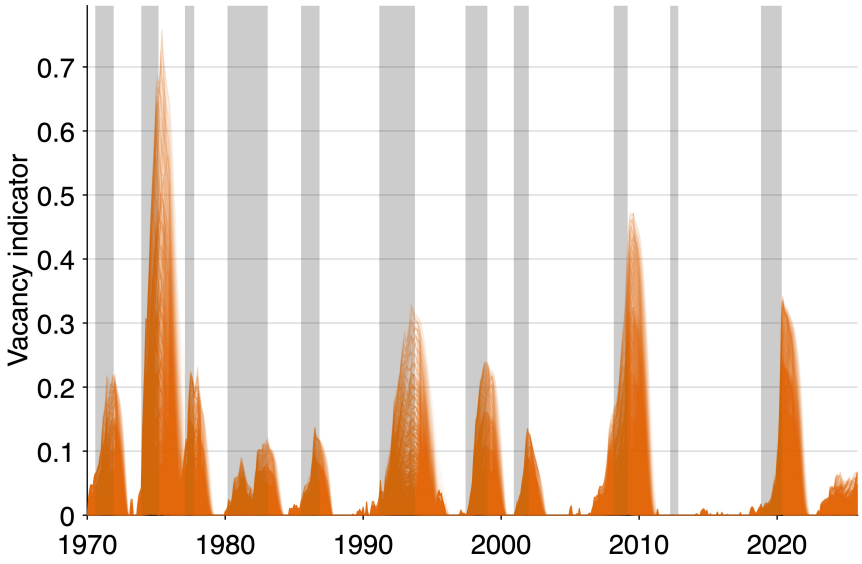
(b) Increases in the Unemployment Rate

Figure 2.8: Increase in the Japanese Unemployment Rate (1970–2025)

Notes: Panel (a) plots the trailing minimum of the smoothed unemployment rate computed using Equation (2.11). Panel (b) plots the unemployment increase defined in Equation (2.13) as the difference between the smoothed unemployment rate and its recent minimum. Gray shaded regions indicate recessions dated by the ESRI.



(a) Trailing Maximum of the Vacancy Rate



(b) Decreases in the Vacancy Rate

Figure 2.9: Decrease in the Japanese Vacancy Rate (1970–2025)

Notes: Panel (a) plots the trailing maximum of the smoothed vacancy rate computed using Equation (2.12). Panel (b) plots the vacancy decrease defined in Equation (2.14) as the difference between the smoothed vacancy rate and its recent maximum. Gray shaded regions indicate recessions dated by the ESRI.

tions. We further combine unemployment and vacancy signals using both linear combinations and nonlinear transformations, such as the minimum and maximum of the two series. This approach generates a large set of candidate indicators that differ in their sensitivity and robustness, allowing the data to determine which specifications best capture recession dynamics.

Figures 2.8b and 2.9b plot the unemployment- and vacancy-based indicators for Japan. The unemployment-based indicators capture increases relative to their recent minimum, while the vacancy-based indicators capture declines relative to their recent maximum. Both increase during recession periods, but they differ in their dynamics. The unemployment-based indicators evolve smoothly and adjust gradually, whereas the vacancy-based indicators exhibit sharper movements and align more closely with the onset of recessions. These patterns suggest that vacancy-based measures provide a more responsive signal of labor market conditions, while unemployment-based measures reflect slower adjustment. Taken together, the two series provide complementary information, motivating their combination to improve recession detection.

2.5.5 Summary

Using the unemployment rate, we generate a total of $(12+10)\times 18\times 11 = 4{,}356$ recession indicators by considering alternative smoothing methods, trailing windows, and scaling transformations. Applying the same procedure to the vacancy rate yields an additional 4,356 indicators.

Next, we combine the unemployment and vacancy indicators using linear combinations, producing $4{,}356\times 11 = 47{,}916$ composite indicators. We then construct an additional 47,916 indicators by taking the minimum and maximum across these combined measures. Altogether, this process generates $47{,}916 + 47{,}916 = 95{,}832$

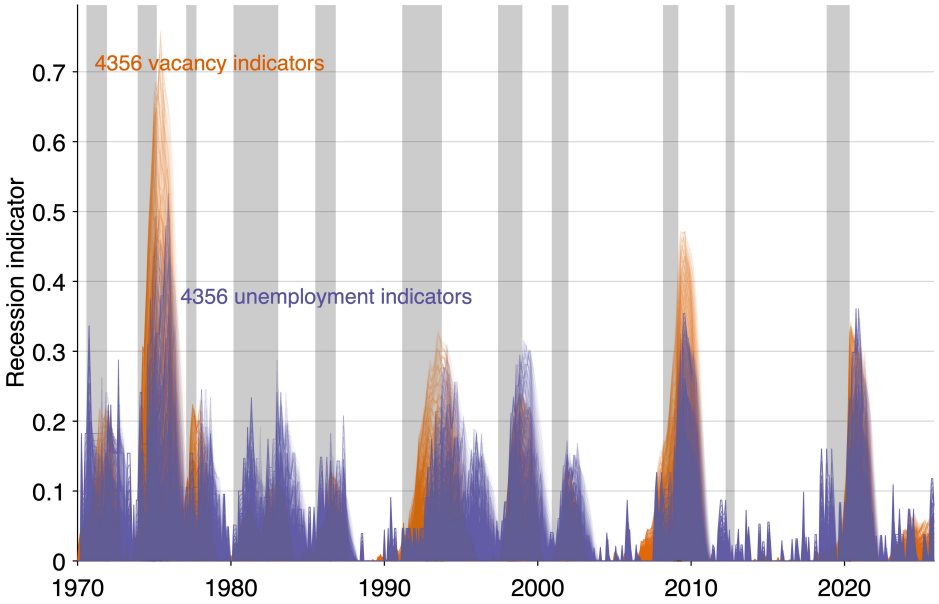


Figure 2.10: 4,356 Unemployment Indicators and 4,356 Vacancy Indicators

Notes: The unemployment indicators are constructed by applying the scaling transformation in Equation (2.15) to the unemployment increases shown in Figure 2.8b. Similarly, the vacancy indicators are constructed by applying the scaling transformation in Equation (2.18) to the vacancy decreases shown in Figure 2.9b. Gray shaded regions indicate recessions dated by the Economic and Social Research Institute (ESRI).

recession indicators, which serve as the basis for constructing recession classifiers.

2.6 Results

Following the framework proposed by Michaillat (2025), described in Section 2.5, we construct recession classifiers as combinations of labor-market indicators and threshold values. We then evaluate their performance using the anticipation–precision frontier. Among 193 frontier classifiers, we select six classifiers whose standard deviations of detection errors are below three months, which form the final high-precision set used in the analysis.

2.6.1 Detection Methodology

After obtaining 95,832 recession indicators, we generate hundreds of millions of recession classifiers by applying alternative thresholds to each indicator. As shown in Figure 2.7, the indicator can cross the threshold multiple times during recoveries, generating spurious recession signals. This pattern motivates the reset criterion requiring the indicator to return to zero before a new recession can be identified.

Therefore, a recession begins when the indicator crosses the threshold from below while the economy is currently in expansion. To avoid false recession signals caused by temporary movements around the threshold, the methodology keeps track of whether the economy is already in a recession or expansion state. In particular, the economy can only enter a new recession if it was previously in expansion, and it only returns to expansion once the indicator falls back to zero. This prevents short-lived rebounds in the indicator during recoveries from being incorrectly treated as new recessions. Although more computationally intensive than simple threshold rules, the procedure produces a more stable and economically intuitive classification of business-cycle turning points.

2.6.2 Selecting Perfect Classifiers

Using 95,832 recession indicators together with 2,500 alternative thresholds ranging from 0.0001 to 0.25, we generate 1,844,564 perfect recession classifiers. For each classifier, we evaluate its ability to identify the official Japanese recession episodes over the sample period. We retain only classifiers that correctly identify all recessions without generating false positives or false negatives. A false negative occurs when a classifier fails to identify a true recession, while a false positive occurs when a classifier incorrectly signals a recession during an expansion. After

restricting attention to the high-precision segment, we compare the six classifiers' mean and standard deviation of detection errors in order to identify recession rules that provide the strongest real-time turning-point performance.

2.6.3 Evaluating Perfect Classifiers

To select the classifiers included in the ensemble, we again follow Michaillat (2025) exactly. In principle, a policymaker with mean–variance preferences over detection error would choose the frontier classifier k that minimizes

$$\mu(k) + \lambda\sigma(k),$$

where $\mu(k)$ and $\sigma(k)$ denote the mean and standard deviation of the detection error, and $\lambda > 0$ captures the weight placed on precision relative to anticipation. Because λ is not observed, we do not select a single classifier. Instead, as in Michaillat (2025), we retain all classifiers on the anticipation–precision frontier whose standard deviation of detection errors satisfies

$$\sigma(k) < 3 \text{ months}.$$

Under the normal approximation for detection errors, this restriction implies that the associated 95% confidence interval for the recession start date has width

$$4\sigma(k) < 12 \text{ months}.$$

We therefore define the ensemble as

$$\mathcal{K}^* = \{k \in \mathcal{F} : \sigma(k) < 3\},$$

where $\mathcal{F}$ denotes the set of classifiers on the anticipation–precision frontier. This is the same selection rule used in Michaillat (2025): it focuses attention on the high-precision segment of the frontier and excludes classifiers whose implied timing uncertainty exceeds one year.

2.6.4 Finding the Anticipation-Precision Frontier

We then select classifiers along the anticipation–precision frontier, defined as those that minimize both the mean detection delay and its variability. This frontier highlights classifiers that achieve the best tradeoff between timely and reliable recession detection.

Figure 2.11a plots the mean and standard deviation of detection errors across a large set of recession classifiers. The anticipation–precision frontier identifies the set of optimal classifiers that minimize detection delay for a given level of dispersion, or equivalently, minimize dispersion for a given level of delay. The downward-sloping frontier reveals a clear tradeoff: classifiers that achieve greater anticipation (more negative mean errors) exhibit substantially higher variability in detection timing, while classifiers with low dispersion detect recessions more consistently but with a modest delay.

The frontier contains 193 classifiers out of the 1,844,564 perfect classifiers. The left-most classifier, which achieves the highest precision, has a mean detection error of approximately 0.06 months and has a standard deviation of detection errors of 2.7 months, indicating very accurate timing with minimal delay. In contrast, the right-most classifier exhibits substantial anticipation, with a mean detection error of -208.5 months and a standard deviation detection error of 158 months, reflecting a large loss in precision. Michaillat (2025) conducts this analysis for the United States from 1979–2021. The classifier ensemble in the United States

detects recessions only 1.2 months after the recession has started. In Japan, we are able to detect recessions in 0.06 months after onset, while on average the Japanese ESRI committee detects recessions 26.5 months later, as shown in Table 2.1. This shows that the algorithm is applicable to different labor markets worldwide and may perform even better in the Japanese context. Despite major differences between the United States and Japan in labor market structure and business cycle dynamics, the classifier ensemble is still able to detect recessions quickly and accurately. These findings suggest that combining unemployment and vacancy information provides a robust and internationally applicable recession indicator.

Table 2.3 reports episode-level recession detection results for the selected ensemble under alternative thresholds. For each threshold, the table lists the official start date, the detected start date, and the corresponding detection error for all 11 recession episodes, allowing us to examine how detection timing varies with the choice of threshold. Several patterns emerge.

First, lower thresholds tend to detect recessions earlier, often generating smaller or negative detection errors. In many episodes, detection occurs either at the official start date or slightly in advance, although some variation across recessions remains.

Second, higher thresholds generally produce later detection relative to lower thresholds, with more frequent delays across episodes. However, detection timing is not uniformly later in all cases, and variation across recessions persists.

Third, the intermediate threshold does not consistently improve performance across episodes. It produces a mix of early and delayed detections and does not uniformly reduce variation in detection timing relative to other thresholds.

Finally, the table reveals substantial heterogeneity across recession episodes.

Table 2.3: Detected Recession Starts by Threshold for the Main Ensemble

Ensemble Member	Threshold	Recession	Official Start	Official Trough	Detection Start	Error (Months)
			Threshold = 0.0005			
1	0.0005	1	1970.58	1971.92	1970.25	-4
1	0.0005	2	1973.92	1975.17	1974.08	2
1	0.0005	3	1977.08	1977.75	1977.33	3
1	0.0005	4	1980.17	1983.08	1980.58	5
1	0.0005	5	1985.50	1986.83	1985.75	3
1	0.0005	6	1991.17	1993.75	1991.42	3
1	0.0005	7	1997.42	1999.00	1997.50	1
1	0.0005	8	2000.92	2002.00	2001.17	3
1	0.0005	9	2008.17	2009.17	2008.42	3
1	0.0005	10	2012.25	2012.83	2012.50	3
1	0.0005	11	2018.83	2020.33	2019.08	3
			Threshold = 0.0056			
2	0.0056	1	1970.58	1971.92	1970.25	-4
2	0.0056	2	1973.92	1975.17	1973.67	-3
2	0.0056	3	1977.08	1977.75	1976.83	-3
2	0.0056	4	1980.17	1983.08	1979.92	-3
2	0.0056	5	1985.50	1986.83	1985.75	3
2	0.0056	6	1991.17	1993.75	1991.17	0
2	0.0056	7	1997.42	1999.00	1997.42	0
2	0.0056	8	2000.92	2002.00	2000.25	-8
2	0.0056	9	2008.17	2009.17	2008.00	-2
2	0.0056	10	2012.25	2012.83	2012.25	0
2	0.0056	11	2018.83	2020.33	2018.92	1
			Threshold = 0.0081			
3	0.0081	1	1970.58	1971.92	1970.25	-4
3	0.0081	2	1973.92	1975.17	1973.58	-4
3	0.0081	3	1977.08	1977.75	1976.75	-4
3	0.0081	4	1980.17	1983.08	1979.83	-4
3	0.0081	5	1985.50	1986.83	1985.75	3
3	0.0081	6	1991.17	1993.75	1990.75	-5
3	0.0081	7	1997.42	1999.00	1997.42	0
3	0.0081	8	2000.92	2002.00	2000.17	-9
3	0.0081	9	2008.17	2009.17	2007.92	-3
3	0.0081	10	2012.25	2012.83	2012.25	0
3	0.0081	11	2018.83	2020.33	2018.83	0

Notes: The table reports recession detection dates for Japan under alternative ensemble thresholds. Official start dates correspond to the benchmark recession chronology dated by the Economic and Social Research Institute (ESRI). Detection start denotes the first month in which the recession indicator crosses the specified threshold. Detection errors are measured as the difference, in months, between the detected and official recession start dates. Negative values indicate early detection, while positive values indicate delayed detection.

Some downturns are detected at similar times across thresholds, while others exhibit larger differences in detection timing depending on the threshold. Overall, the episode-level results demonstrate that the anticipation–precision tradeoff is not an artifact of aggregation, but arises consistently across recessions and varies with the choice of threshold.

2.6.5 Selecting the Frontier Classifiers with Highest Precision

The classifier ensemble is composed of the specifications on the anticipation–precision frontier with a standard deviation of detection errors below 3 months (Table 2.4). All selected classifiers use simple moving average (SMA) smoothing, with smoothing parameters α between 6 and 9 months and turning horizons β between 5 and 7 months. The Box–Cox transformation parameter γ varies across specifications, while the indicators are combined linearly with high weights ($\delta \in [0.8, 0.9]$). The associated thresholds are small, ranging from 0.0005 to 0.0073.

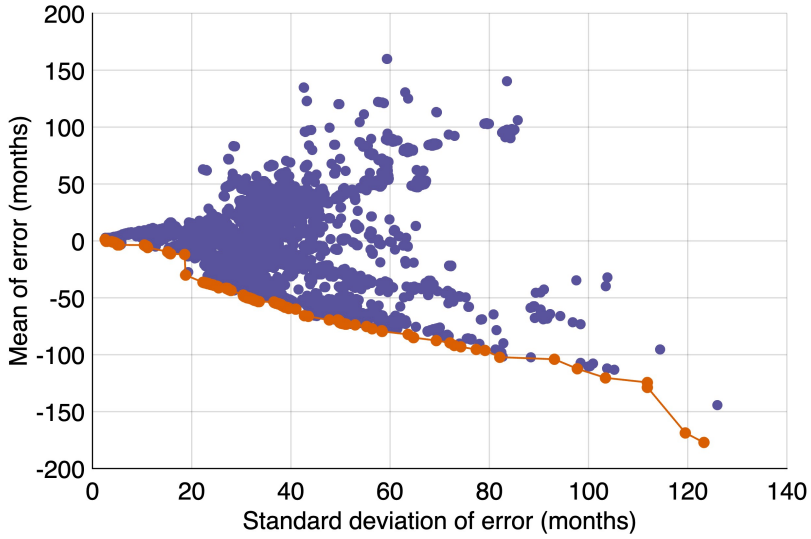
The selected classifiers detect all recession episodes in the sample without generating false positives. The mean detection error ranges from -0.45 to 1.18 months, with an average of 0.06 months, indicating that, on average, the classifiers detect recessions slightly after onset. The standard deviation of detection errors ranges from 2.5 to 2.9 months, with an average of 2.72 months, implying that detection timing remains tightly clustered across episodes.

Overall, the ensemble delivers recession signals that are timely and consistent, reflecting the restriction to the high-precision segment of the anticipation–precision frontier.

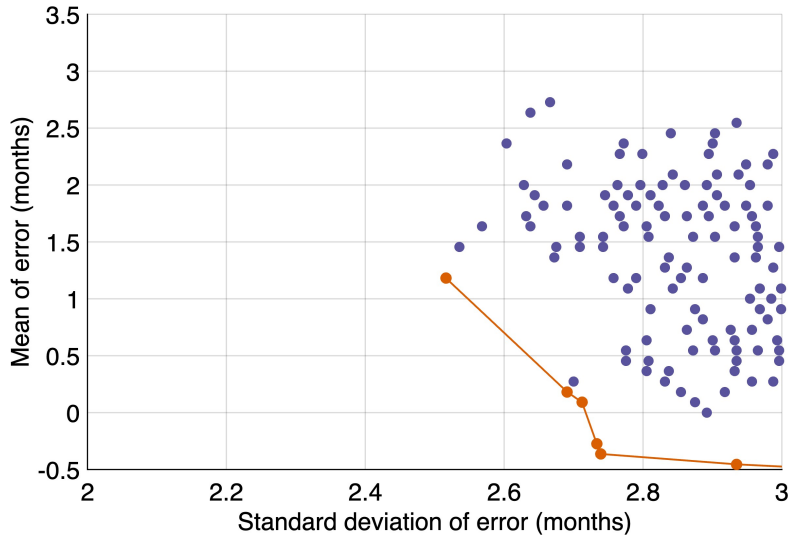
Table 2.4: Classifier Ensemble Selected from the Anticipation–Precision Frontier

Smoothing method	Smoothing parameter	Curving parameter	Turning parameter	Mixing method	Mixing parameter	Threshold	Standard error	Mean error
SMA	6	0.8	8	linear	0.9	0.0005	2.52	1.18
SMA	5	0.1	9	linear	0.8	0.0056	2.69	0.18
SMA	5	0.0	9	linear	0.9	0.0081	2.71	0.09
SMA	5	0.1	9	linear	0.8	0.0054	2.73	-0.27
SMA	5	0.0	9	linear	0.8	0.0073	2.74	-0.36
SMA	5	0.0	9	linear	0.8	0.0072	2.93	-0.45
Average	**5.17**	**0.17**	**8.83**	–	**0.83**	**0.0057**	**2.72**	**0.06**

Notes: Detection-error statistics are reported in months. `SMA` denotes simple moving average smoothing. The table reports classifiers selected from the anticipation–precision frontier subject to the restriction $\sigma(k) < 3$ months. The Box–Cox parameter γ governs the transformation of the recession indicator, while the turning parameter β determines the horizon used to compute turning points. Indicators are combined linearly using weight δ. A recession is identified when the resulting indicator crosses the threshold ζ from below. Detection errors are defined as the difference, in months, between estimated and official recession start dates dated by the Economic and Social Research Institute (ESRI).



(a) 193 Perfect Classifiers on the Anticipation–Precision Frontier



(b) 6 Perfect Classifiers satisfying $\sigma(k) < 3$ on the Anticipation–Precision Frontier

Figure 2.11: Anticipation and Precision of 1,844,564 Perfect Recession Classifiers in Japan, January 1970–December 2021

Notes: The figures plot the mean detection error against the standard deviation of detection errors for recession classifiers constructed using alternative indicators and thresholds. Each point represents a classifier. The orange frontier traces classifiers that maximize anticipation for a given level of dispersion, or equivalently maximize precision for a given level of anticipation. Panel (a) shows the full anticipation–precision frontier across all classifiers, while panel (b) focuses on the high-precision region satisfying $\sigma(k) < 3$. Detection errors are measured relative to recession start dates dated by the Economic and Social Research Institute (ESRI).

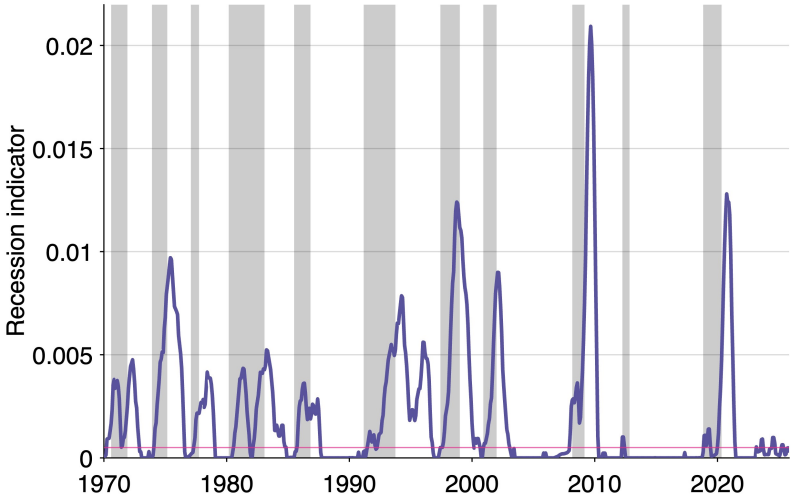


Figure 2.12: Classifier 1: Threshold 0.0005

Notes: The classifier specification corresponds to the first classifier reported in Table 2.4. The purple line represents the recession indicator underlying the classifier, while the pink line represents the corresponding recession threshold. Gray shaded regions indicate recessions dated by the Economic and Social Research Institute (ESRI).

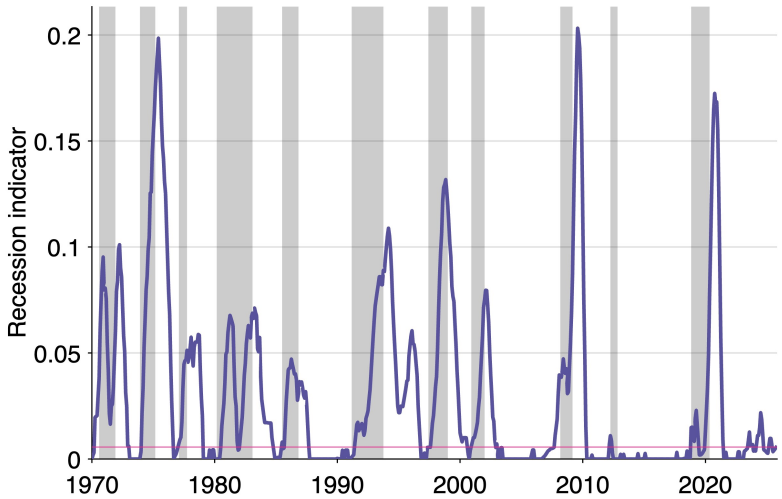


Figure 2.13: Classifier 2: Threshold 0.0056

Notes: The classifier specification corresponds to the second classifier reported in Table 2.4. The purple line represents the recession indicator underlying the classifier, while the pink line represents the corresponding recession threshold. Gray shaded regions indicate recessions dated by the Economic and Social Research Institute (ESRI).

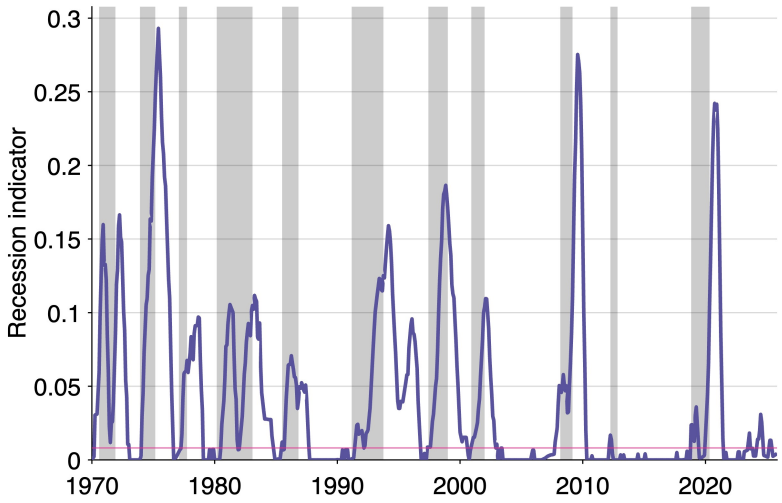


Figure 2.14: Classifier 3: Threshold 0.0081

Notes: The classifier specification corresponds to the third classifier reported in Table 2.4. The purple line represents the recession indicator underlying the classifier, while the pink line represents the corresponding recession threshold. Gray shaded regions indicate recessions dated by the Economic and Social Research Institute (ESRI).

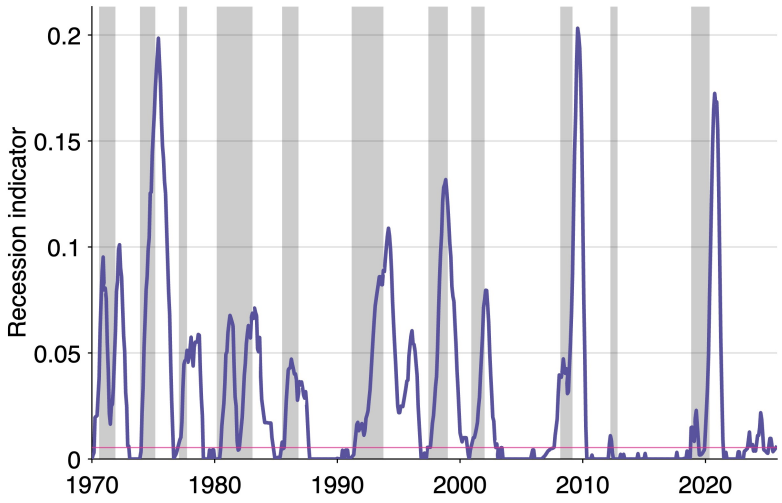


Figure 2.15: Classifier 4: Threshold 0.0054

Notes: The classifier specification corresponds to the fourth classifier reported in Table 2.4. The purple line represents the recession indicator underlying the classifier, while the pink line represents the corresponding recession threshold. Gray shaded regions indicate recessions dated by the Economic and Social Research Institute (ESRI).

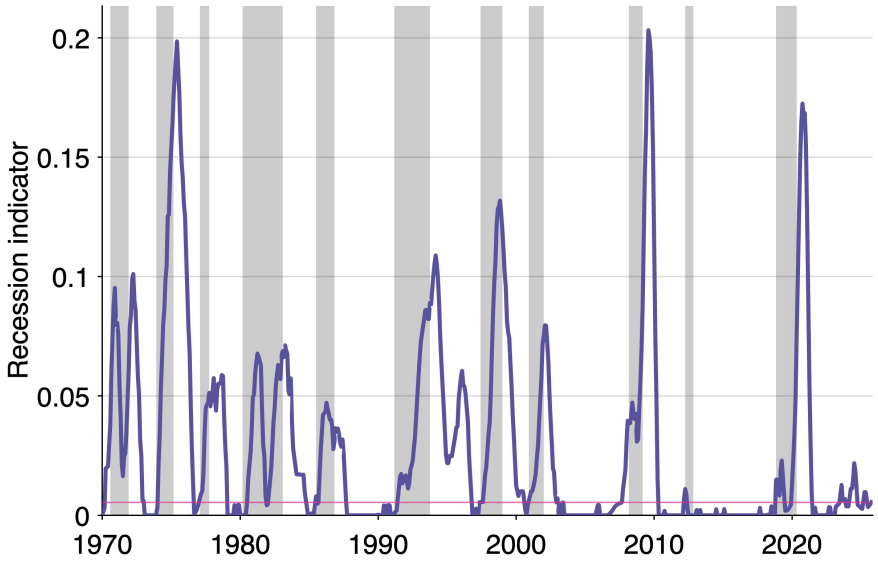


Figure 2.16: Classifier 5: Threshold 0.0073

Notes: The classifier specification corresponds to the fifth classifier reported in Table 2.4. The purple line represents the recession indicator underlying the classifier, while the pink line represents the corresponding recession threshold. Gray shaded regions indicate recessions dated by the Economic and Social Research Institute (ESRI).

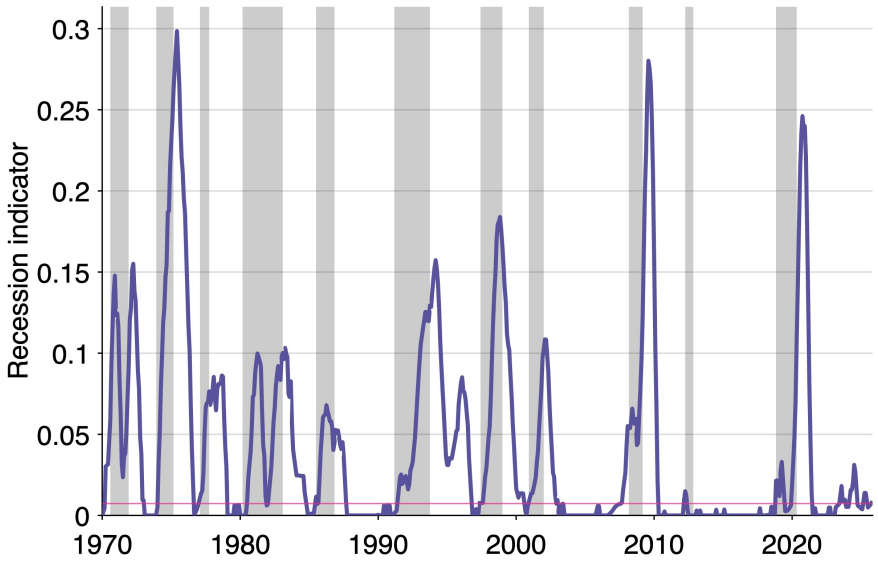


Figure 2.17: Classifier 6: Threshold 0.0072

Notes: The classifier specification corresponds to the sixth classifier reported in Table 2.4. The purple line represents the recession indicator underlying the classifier, while the pink line represents the corresponding recession threshold. Gray shaded regions indicate recessions dated by the Economic and Social Research Institute (ESRI).

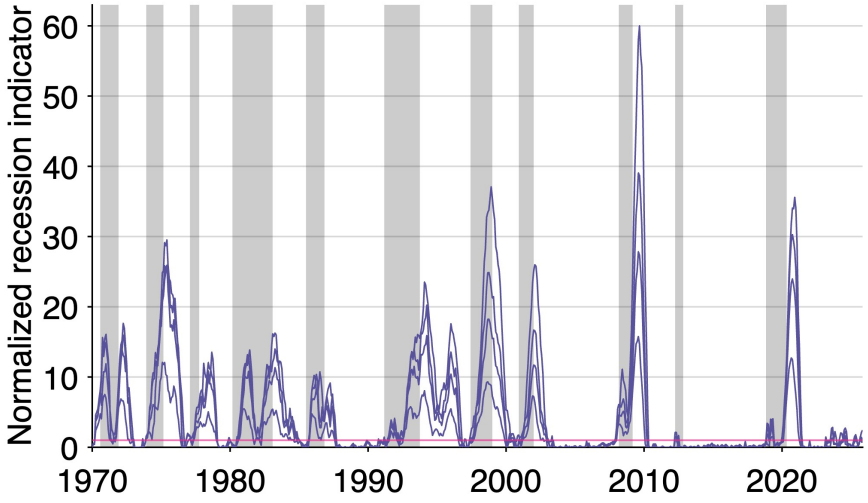


Figure 2.18: Full Set of All Six Recession Classifiers

Notes: The figure plots all classifiers reported in Table 2.4. The purple lines represent the recession indicators underlying the classifiers, while the pink lines represent the corresponding recession thresholds. Gray shaded regions indicate recessions dated by the Economic and Social Research Institute (ESRI).

2.7 Detecting Recessions

To build the recession detection algorithm, we follow Michaillat (2025) to aggregate the detection signals produced by the classifier ensemble. We then use the aggregated signals to compute the current recession risk.

2.7.1 Empirical Construction of Recession Probabilities

We follow Michaillat (2025) exactly. For each classifier k and recession j in the training sample, we compute the detection error

$$\varepsilon_{k,j} = d_{k,j} - s_j,$$

where $d_{k,j}$ is the classifier's detection date and s_j is the ESRI Japanese recession start date as shown in Table 2.2. For each classifier, these historical errors are summarized by their sample mean μ_k and standard deviation σ_k, which capture systematic anticipation or delay and timing precision, respectively.

When classifier k issues a real-time detection at date d_k, we treat the historical detection errors as an empirical forecast-error sample and approximate their distribution by a normal distribution with mean μ_k and variance σ_k^2, exactly as in Michaillat (2025). Letting

$$\varepsilon_k = d_k - s,$$

where s is the unknown start date of the current recession, the event $\{s \leq t\}$ is equivalent to $\{\varepsilon_k \geq d_k - t\}$. Therefore,

$$\Pr(s \leq t \mid d_k) = \Pr(\varepsilon_k \geq d_k - t).$$

This probability is computed using only information from the training sample.

Whenever an individual recession indicator crosses its threshold, we infer the probability that the recession has already started from the distribution of the detection error. If classifier k is on average exactly on time and the detection error is symmetric, then the recession has started with probability 0.5 when the classifier is activated. If the classifier tends to detect recessions early on average, this probability is below 0.5; if the classifier tends to detect recessions late on average, it is above 0.5. In the months following detection, this probability converges to one according to the cumulative distribution function of the detection error.

More formally, suppose that classifier k detects the J recessions in the training sample. Each detection j yields a detection date $d(k, j)$ and a detection error $\varepsilon(k, j)$. As in Michaillat (2025), we assume that the detection error $\varepsilon(k)$ is normally distributed with mean $\mu(k)$ and standard deviation $\sigma(k)$:

$$\varepsilon(k) \sim \mathcal{N}\Big(\mu(k), \sigma^2(k)\Big).$$

Then, for any date $t \geq d(k)$, the probability that the start date s of the new recession occurred before t, conditional on classifier k having detected a recession at date $d(k)$, is

$$P(k,t) = \Pr\Big(s < t \mid d(k), \mu(k), \sigma(k)\Big) \tag{2.21}$$

$$= \Pr\Big(d(k) - s > d(k) - t \mid \mu(k), \sigma(k)\Big) \tag{2.22}$$

$$= \Pr\Big(\varepsilon(k) > d(k) - t \mid \mu(k), \sigma(k)\Big). \tag{2.23}$$

Under the normal approximation, this becomes

$$P(k,t) = 1 - \Phi\left(\frac{d(k) - t - \mu(k)}{\sigma(k)}\right),$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution.

2.7.2 Recession Probability from Classifier Ensemble

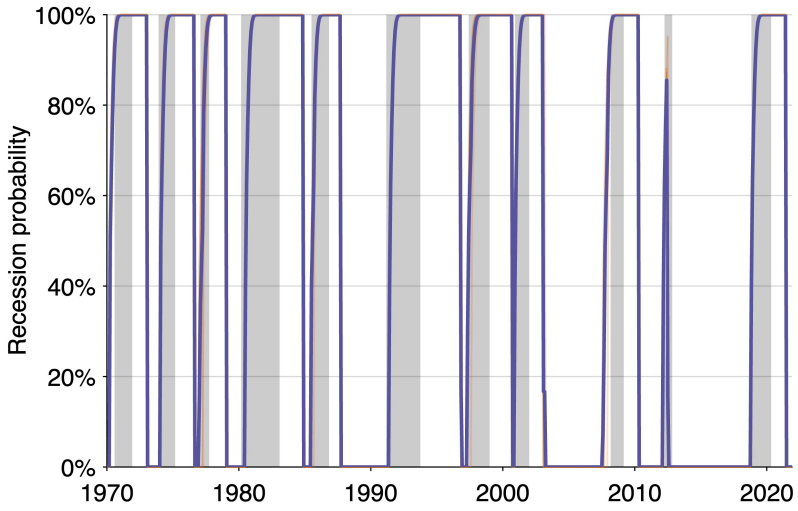


Figure 2.19: In-sample recession probability, January 1970–December 2021.

Notes: The figure plots the estimated probability of a recession over time. The thick purple line shows the average probability across the six classifiers in the selected ensemble. The thin orange lines show the probability from each classifier separately. Shaded areas indicate recessions dated by the ESRI.

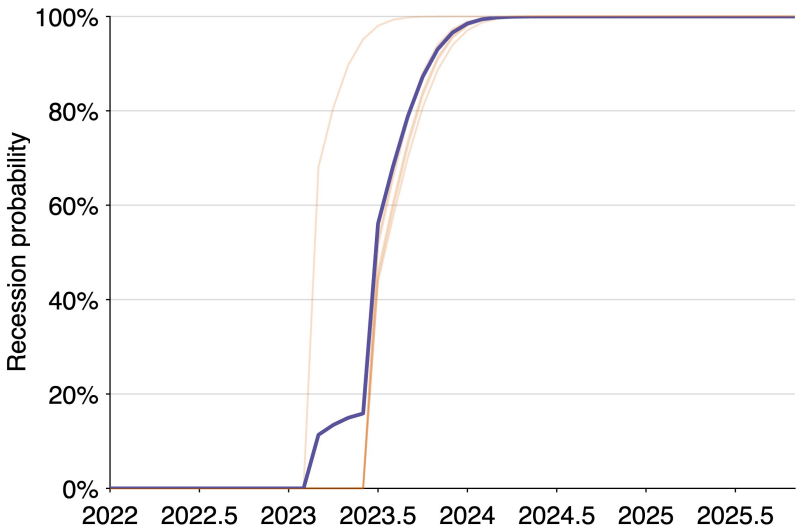


Figure 2.20: Out-of-sample recession probability, January 2022–December 2025

Lastly, we summarize the information from the $K = 6$ detection classifiers by constructing an ensemble recession probability. Each classifier produces a probability that a recession has started, and we average these probabilities to obtain a single aggregate measure. Figure 2.19 shows the resulting ensemble probability for the 1970–2021 training sample. For all 11 recessions, the probability rises in advance of the official start date, reaches 1 at the onset of the recession, and remains there until the recession has ended.

2.7.3 Application of Recession Detection to Current Data

Next, we apply the classifier estimated on the 1970-2021 training sample to the out-of-sample period 2022-2025 to assess current recession risk in the Japanese economy. The estimated recession probability first rises in 2023 and increases sharply after mid-2023, peaking in 2024 before easing somewhat. It rises again around mid-2024 and remains elevated through the end of the year. More recently, the probability begins to increase again around March 2025.

2.8 Evaluating the recession detection algorithm with backtests

To examine the reliability of the algorithm, we perform a series of out-of-sample backtests. In each exercise, the training sample is shortened and the algorithm is evaluated on recessions occurring after the training period. The results indicate that the detection algorithm performs robustly across different training windows. Moreover, the classifier ensembles obtained from these backtests consistently imply a meaningful probability that the Japanese economy entered a recession in 2025.

Table 2.5: Performance of the Algorithm on Backtests

Backtest	2015	2005	1995
Number of classifiers	10	14	9
Training sample detection error (months)			
Mean	0.9	0.4	-0.3
Standard deviation	2.7	2.4	2.1
Minimum	-4.2	-4.1	-3.2
Maximum	4.6	4.1	3.1
Testing sample detection error (months)			
Mean	0.7	-3.9	23.1
Standard deviation	0.0	5.0	25.5
Minimum	0.7	-10.6	-4.9
Maximum	0.7	0.6	63.8
False positives	0	1	2
False negatives	0	0	0
Training recessions	10	8	6
Testing recessions	1	3	5

Notes: The table reports the performance of the selected classifier ensembles in each backtest exercise. The training samples are January 1970–December 1994, January 1970–December 2004, and January 1970–December 2014. The testing periods are January 1995–December 2025, January 2005–December 2025, and January 2015–December 2025. Detection errors are measured in months relative to recession onset dates.

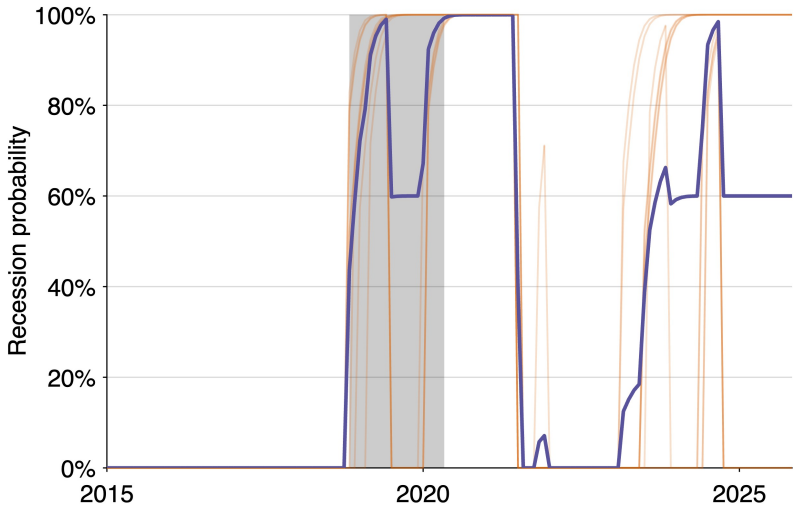


Figure 2.21: Out-of-sample Japan Recession Probability, January 2015–December 2025

2.8.1 Backtesting from 2015

We train the algorithm using data through December 2014. The training sample contains 10 recessions, as summarized in Table 2.5. This procedure selects 10 classifiers from the high-precision segment of the anticipation–precision frontier, reported in Table B.2. We then evaluate these classifiers out of sample using data from January 2015 through December 2025. During this test period, all 10 classifiers detect the 2020 recession with no false positives. The detection is also timely. On average, the recession is signaled only 0.7 months after its official onset.

2.8.2 Backtesting from 2005

We train the algorithm using data through December 2004. The training sample contains eight recessions, as summarized in Table 2.5. This procedure selects 14 classifiers from the high-precision segment of the anticipation–precision

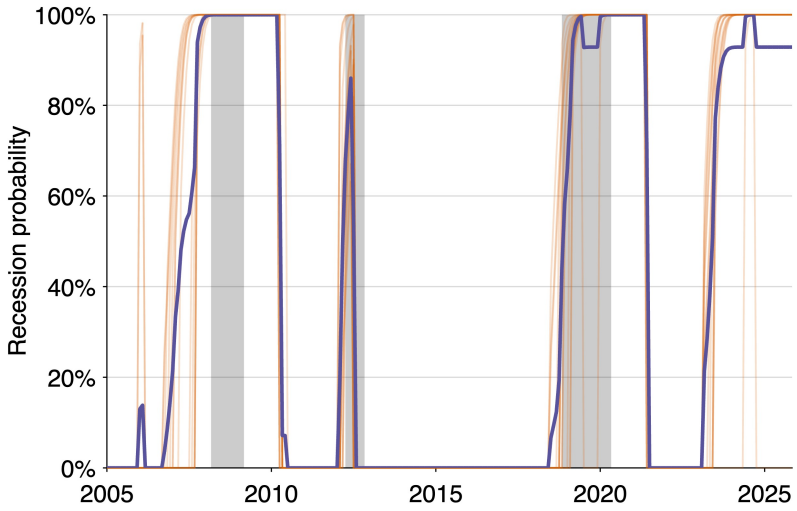


Figure 2.22: Out-of-sample Japan Recession Probability, January 2005–December 2025

frontier, reported in Table B.3. We then evaluate these classifiers out of sample using data from January 2005 through December 2025.

The out-of-sample performance is particularly striking, as the classifiers exhibit anticipatory detection since the mean detection error is -3.9 months, implying that, on average, the classifiers signal recessions nearly four months before the official recession onset dates. The standard deviation of the detection error is approximately five months, which remains reasonably small given the difficulty of predicting business-cycle turning points in real-time.

2.8.3 Backtesting from 1995

The backtesting exercise beginning in 1995 exhibits substantial noise, as shown in Figure 2.23, although the algorithm still successfully captures all recession episodes. However, the out-of-sample performance deteriorates considerably, with the mean detection error rising to 23.1 months, implying that recessions are identified nearly two years after their official onset on average. The exceptionally

large standard deviation further suggests that the backtest calibrated through 1995 becomes unstable and loses predictive reliability. Therefore, we conclude the backtesting analysis at this point.

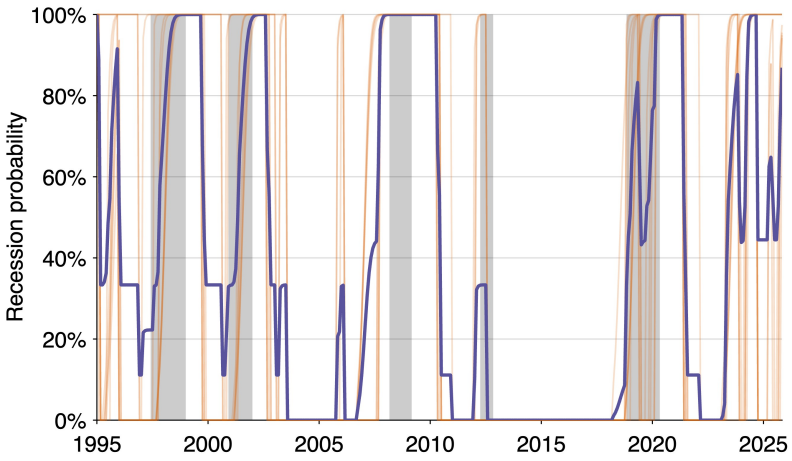


Figure 2.23: Out-of-sample Japan Recession Probability, January 1995–December 2025

2.8.4 Sensitivity Test: Excluding the 2012 recession episode

We conduct an additional robustness exercise excluding the 2012 recession episode because its underlying economic dynamics may differ from those of more standard cyclical recessions. The slowdown occurred in the aftermath of the Great East Japan Earthquake and the post-Lehman recovery period, and therefore reflects a combination of reconstruction effects, temporary supply disruptions, and broader cyclical weakening. Since the recession detection framework is designed to identify labor-market-based signals associated with typical business-cycle downturns rather than disaster-related shocks, the 2012 episode may represent a structurally atypical recession episode in the Japanese context. Excluding this episode allows us to examine the extent to which the detection results depend on this unusual historical event.

Consistent with this interpretation, even when the 2012 episode is included in the training sample, the model does not assign a near-100% recession probability to that episode.

Nevertheless, excluding the 2012 recession substantially worsens the overall detection performance. The average detection timing deteriorates from nearly contemporaneous detection in the baseline specification, with an average delay of only 0.06 months after onset, to roughly two months after recession onset. In addition, the predicted probability of a recession in 2025 declines from nearly 100% to around 80%.

Overall, these findings suggest that, despite its atypical origin, the 2012 episode still provides economically meaningful information about recession dynamics in Japan and helps the model identify weaker and more gradual downturns that are characteristic of the Japanese business cycle.

The exercise of excluding 2012 asks whether the results are stable to excluding a weak and ambiguous official recession. Performance weakens, not because the labor-market signal disappears, but because excluding 2012 changes the exact-count classifier-selection problem. The selected classifiers are re-optimized to generate 10 signals instead of 11, producing a different frontier. This shows that the method is informative but sensitive to how ambiguous recession episodes are treated.

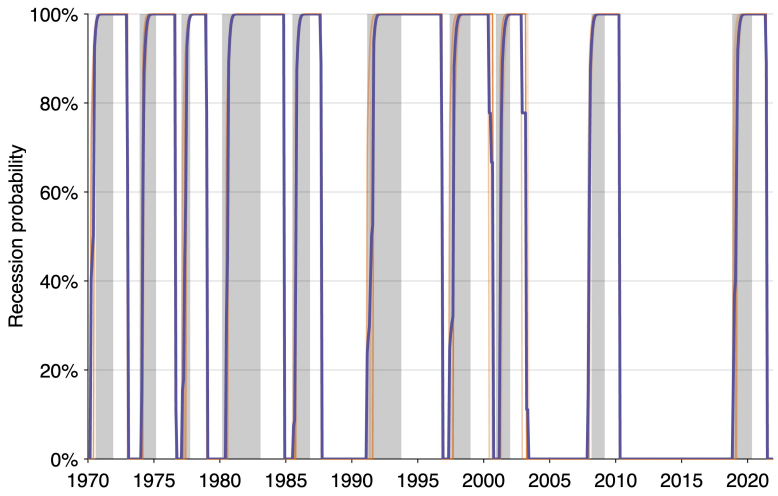


Figure 2.24: In-sample Recession Probability (excl.2012 recession), January 1970–December 2021

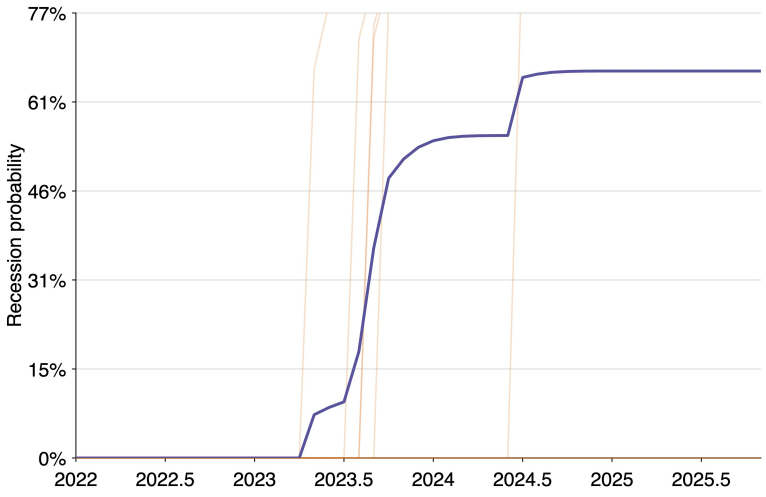


Figure 2.25: Out-of-sample Recession Probability, January 2022–December 2025

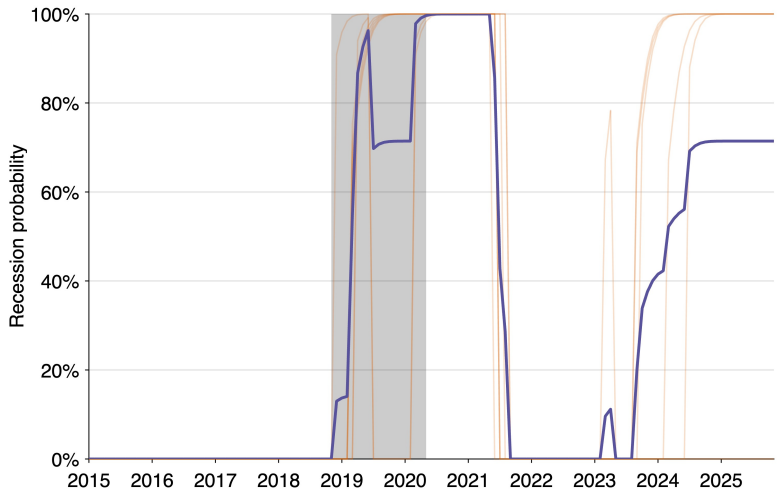


Figure 2.26: Out-of-sample Recession Probability, January 2015–December 2025

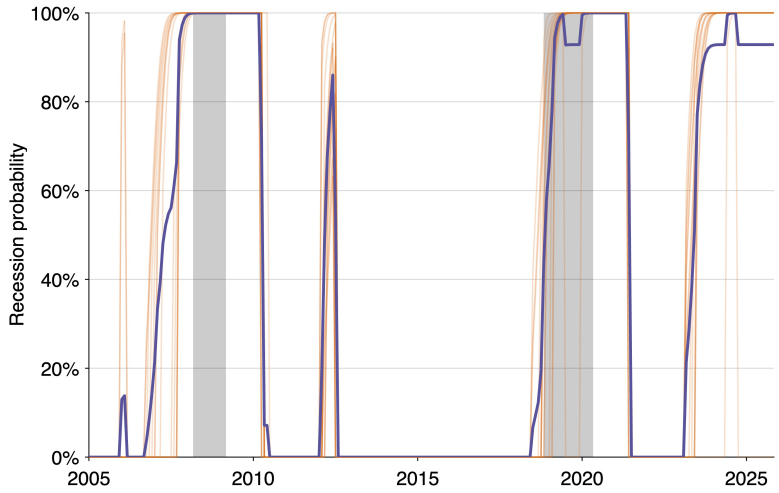


Figure 2.27: Out-of-sample Recession Probability (excl.2012 recession), January 2005–December 2025

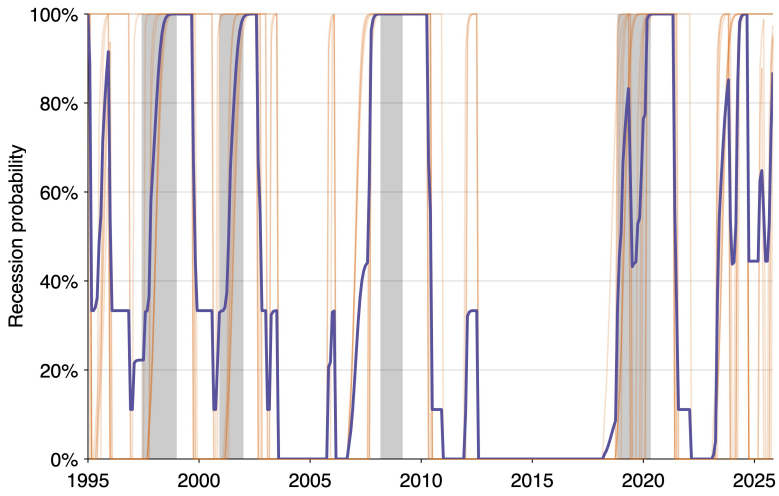


Figure 2.28: Out-of-sample Recession Probability (excl.2012 recession), January 1995–December 2025

2.9 Conclusion

This paper extends the algorithm developed by Michaillat (2025) to Japanese data. These results suggest that the algorithm provides a simple and effective framework for real-time recession detection in Japan. The algorithm detects recessions both more accurately and far earlier than official ESRI announcements, and it outperforms simpler benchmark rules such as the Sahm rule and the Michez rule. The algorithm is trained on data from January 1970 to December 2021, a period that includes 11 recessions. On the training sample, it performs very well. On average, the selected classifier ensemble signals recessions 0.06 months after their true onset, with a standard deviation of detection errors of 2.72 months.

These findings are also consistent with the evidence for the United States reported by Michaillat (2025), where the classifier ensemble substantially improves upon official recession dating delays. Between 1979 and 2021, the classifier ensemble identified recession starts on average 1.2 months after onset, compared to

6.3 months for the National Bureau of Economic Research (NBER) Business Cycle Dating Committee. Taken together, these results suggest that classifier-based recession detection provides a robust and generalizable framework for identifying business-cycle turning points across economies.

An important avenue for future research is to extend the sample further back in time, which would allow the analysis to incorporate additional recession episodes and provide a more stringent test of the algorithm's robustness. A longer historical sample may also improve the precision of the estimated classifiers by exposing them to a wider range of macroeconomic conditions.

More broadly, our results show that this classifier-based approach can be applied to other countries by calibrating thresholds and smoothing parameters to local labor market conditions, providing a practical framework for real-time recession detection across different economies.

Chapter 3

Recession Detection Using Real-time GDP Data

3.1 Introduction

Business-cycle turning points play a central role in macroeconomic analysis and policy decisions, yet recession dating remains difficult in real-time and lacks formal institutional recognition in many countries. While the United States relies on the National Bureau of Economic Research to identify recessions, many economies either announce turning points with substantial delays or rely on simplified heuristic rules such as two consecutive quarters of negative GDP growth. This paper asks whether real-time GDP announcements themselves can serve as a reliable measure of business-cycle turning points.

Building on Michaillat (2025), this paper answers this question by evaluating the performance of real-time GDP against official recession dates determined by the National Bureau of Economic Research (NBER). We construct 4,356 recession indicators using real-time GDP vintages with alternative smoothing methods and scaling variation. We then combine these indicators with different thresholds to obtain our classifiers. The selected 137,457 classifiers are statistically perfect in the sense that they identify all 12 historical recessions in the 1947–2021 training sample without generating false positives. Among these, a subset of classifiers lies on the anticipation–precision frontier. Restricting attention to the high-precision segment yields two classifiers with a standard deviation of detection errors below three months. The selected ensemble signals recessions, on average, 3.04 months after their true onset.

This paper contributes to the literature on real-time business-cycle dating by showing that real-time GDP announcements provide a reliable basis for recession dating across vintages. This suggests that discrepancies in prior work may reflect limitations of traditional dating methods in addition to data revisions. Overall, the results suggest that real-time GDP announcements contain substantial infor-

mation about business-cycle turning points and may provide a practical real-time proxy for NBER-style recession dating, particularly for countries without formal business-cycle dating committees.

3.2 Related literature

More broadly, this paper bridges the literature on real-time GDP measurement and the literature on recession-classification methods. While previous work studies either GDP-based recession dating or industrial-production-based classification algorithms separately, this paper shows that real-time GDP releases can support timely and stable recession detection within a recession-classification framework.

? shows that real-time GDP data contain meaningful information about business-cycle turning points, although recession dating remains difficult because initial GDP releases are noisy and subject to substantial revisions. Similarly, **?** show that business-cycle dating based on real GDP can differ considerably across data vintages, with some recessions appearing or disappearing as GDP estimates are revised over time. Together, these findings highlight the importance of evaluating recession indicators using real-time vintage data rather than ex post revised series. Motivated by this literature, this paper evaluates whether real-time GDP announcements provide timely and reliable signals of recession turning points.

Building on this literature, Michaillat (2025) evaluates whether recession-classification methods can improve real-time recession detection using industrial production data. Although the National Bureau of Economic Research relies heavily on product-market variables to date business cycles, the paper finds that industrial-production classifiers perform worse than labor-market classifiers in terms of both anticipation and precision. These findings suggest that not all output-based indicators are equally informative for recession detection and moti-

vate this paper's examination of whether broader output measures such as real-time GDP announcements can support more timely and stable recession detection.

3.3 Data

This section presents the data on business cycle dating and real-time GDP announcements used in the paper. The sample covers the period from January 1947 to December 2025.

3.3.1 real-time GDP Data

This paper uses real-time macroeconomic data from the Federal Reserve Bank of Philadelphia's real-time Data Set for Macroeconomists (RTDSM), which records National Income and Product Accounts (NIPA) variables as they were available to policymakers and researchers at each point in time. The data consist of historical "vintages" that preserve both initial releases and subsequent revisions, allowing the analysis to replicate the information available in real-time rather than relying on fully revised data. The vintages are constructed using publications from the Bureau of Economic Analysis (BEA), including the *Survey of Current Business* and historical NIPA tables. Quarterly vintages begin in 1947 and generally contain the BEA's advance estimate for the previous quarter, while monthly vintages track the sequence of advance, second, and third estimates. This structure makes it possible to evaluate recession detection using only information that would have been available to policymakers in real-time.

3.3.2 Business cycle dating

National Bureau of Economic Research (2023) identifies US recessions through its Business Cycle Dating Committee. The NBER does not identify recessions using a single economic indicator. Instead, it evaluates a broad range of macroeconomic variables to determine turning points in the business cycle. A business cycle peak marks the end of an economic expansion and the beginning of a contraction, while a trough marks the end of the contraction and the start of a recovery (NBER, 2004) . Following the NBER convention, the first month of a recession is defined as the month after the peak, and the last month is defined as the month of the trough. NBER recession dates play an important role in macroeconomic research because they provide the benchmark chronology used to evaluate business cycle fluctuations and recession forecasting models. However, these official dates are often announced with substantial delays, which limits their usefulness for real-time policymaking and motivates the development of recession detection algorithms.

Between January 1947 and December 2025, the National Bureau of Economic Research (NBER) dates 12 U.S. recessions, which are shown in Figure 3.1. However, recession start dates are identified only with a considerable delay: from 1979 to 2021, the NBER announced recessions an average of 6.3 months after their onset.

3.4 Real-time GDP

Figure 3.1 plots real-time announcements of U.S. real gross domestic product (GDP) from 1947 to 2025. The series exhibits strong long-run growth, with periods of slower growth and temporary declines during economic downturns. Be-

cause the data are based on first-release GDP announcements, the figure reflects the information that would have been available to policymakers and researchers in real-time rather than revised historical estimates. Gray shaded bars indicate recessions dated by the National Bureau of Economic Research (NBER). The figure highlights how GDP growth slows markedly during recession periods, particularly during the Great Recession and the COVID-19 recession.

The real-time GDP indicator is constructed in three steps: smoothing, curvature adjustment, and turning-point detection. First, quarterly real-time GDP data are smoothed using backward-looking simple moving averages over windows ranging from 0 to 4 quarters in order to reduce short-run volatility. Second, the smoothed series is transformed using a Box–Cox transformation with curvature parameters ranging from 0 to 1, allowing the indicator to vary between logarithmic and linear specifications. Finally, the real-time GDP indicator is constructed by comparing the transformed GDP series to its recent peak over backward-looking windows of 1 to 6 quarters. Because GDP is procyclical, the indicator measures the decline in GDP relative to its recent maximum, so larger values correspond to larger output declines and therefore a higher likelihood of recession. The procedure generates a large set of candidate real-time GDP indicators that differ according to their smoothing, curvature, and turning-point parameters.

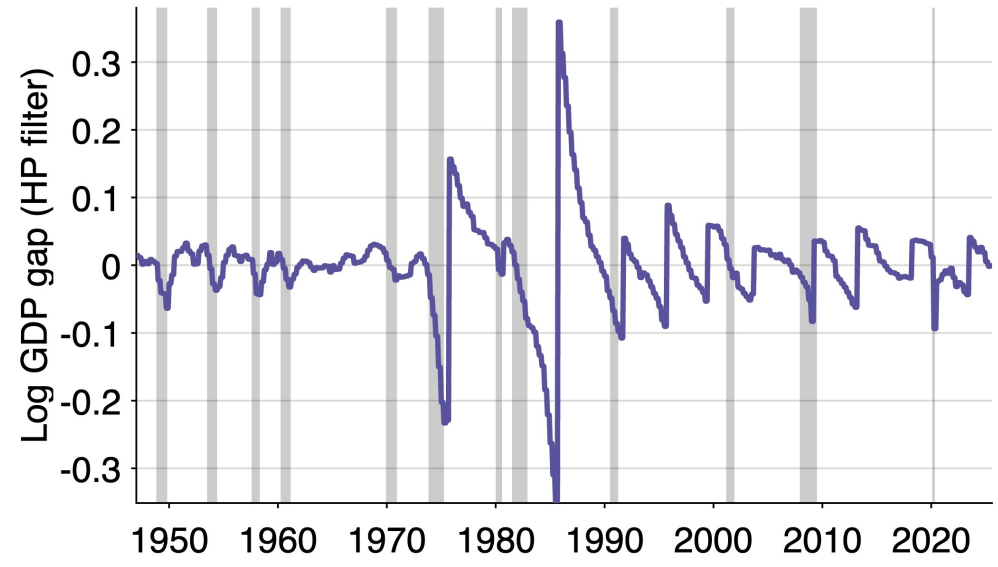


Figure 3.1: Real-Time Log GDP Gap from First-Release Real GDP

Notes: The figure plots the real-time log GDP gap in the United States from 1947–2025, computed using a Hodrick–Prescott filter. Positive values indicate that GDP is above trend, while negative values indicate that GDP is below trend. Gray shaded bars indicate recession periods dated by the NBER.

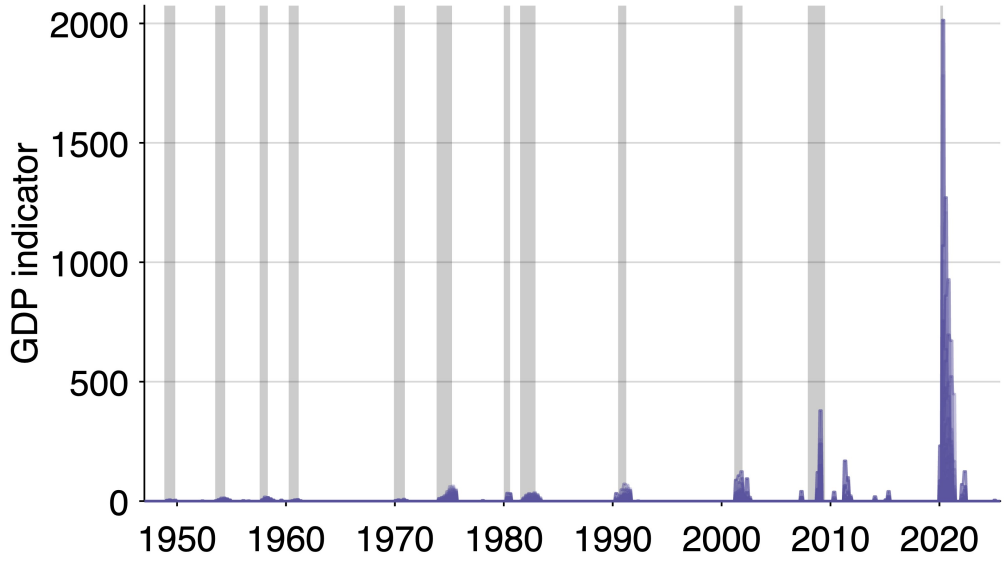


Figure 3.2: real-time GDP Indicator

Notes: The figure displays recession indicators constructed from real-time GDP vintages. Shaded areas denote NBER recessions. Higher values indicate larger GDP declines relative to recent peaks and a greater likelihood of recession.

3.5 Constructing and evaluating recession classifiers

After constructing the 4,356 recession indicators, we generate a large set of recession classifiers by applying threshold values to each indicator.

3.5.1 Detection methodology

Using 4,356 recession indicators together with 2,500 alternative thresholds ranging from 0.0001 to 0.25, we generate 137,457 perfect recession classifiers. A recession begins when the indicator crosses the threshold from below while the economy is currently in expansion. To avoid false recession signals caused by temporary movements around the threshold, the methodology keeps track of whether the economy is already in a recession or expansion state. In particular, the economy can only enter a new recession if it was previously in expansion, and it only returns to expansion once the indicator falls back to zero. This prevents short-lived rebounds in the indicator during recoveries from being incorrectly treated as new recessions. Although more computationally intensive than simple threshold rules, the procedure produces a more stable and economically intuitive classification of business-cycle turning points.

3.5.2 Selecting perfect classifiers

For each classifier, we evaluate its ability to identify official United States recession episodes over the sample period. We retain only classifiers that correctly identify all recessions without generating false positives or false negatives. A false negative occurs when a classifier fails to detect a true recession, whereas a false positive occurs when a classifier incorrectly signals a recession during an

expansion. Restricting attention to the high-precision segment, we then compare the mean and standard deviation of detection errors across classifiers to identify recession rules that deliver the strongest real-time turning-point performance.

3.5.3 Evaluating perfect classifiers

To select the classifiers included in the ensemble, we follow Michaillat (2025). In principle, a policymaker with mean–variance preferences over detection error would choose the frontier classifier k that minimizes

$$\mu(k) + \lambda\sigma(k),$$

where $\mu(k)$ and $\sigma(k)$ denote the mean and standard deviation of the detection error, and $\lambda > 0$ captures the weight placed on precision relative to anticipation. Because λ is not observed, we do not select a single classifier. Instead, as in Michaillat (2025), we retain all classifiers on the anticipation–precision frontier whose standard deviation of detection errors satisfies

$$\sigma(k) < 3 \text{ months}.$$

Under the normal approximation for detection errors, this restriction implies that the associated 95% confidence interval for the recession start date has width

$$4\sigma(k) < 12 \text{ months}.$$

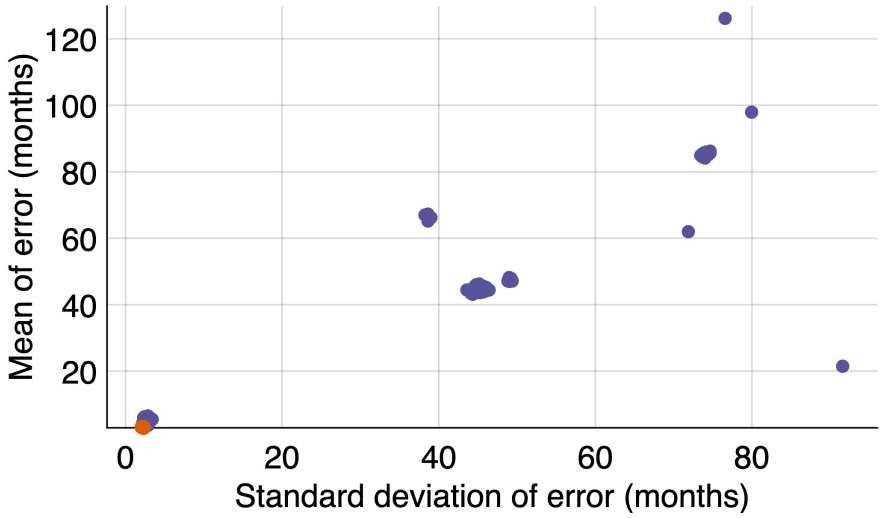
We therefore define the ensemble as

$$\mathcal{K}^* = \{k \in \mathcal{F} : \sigma(k) < 3\},$$

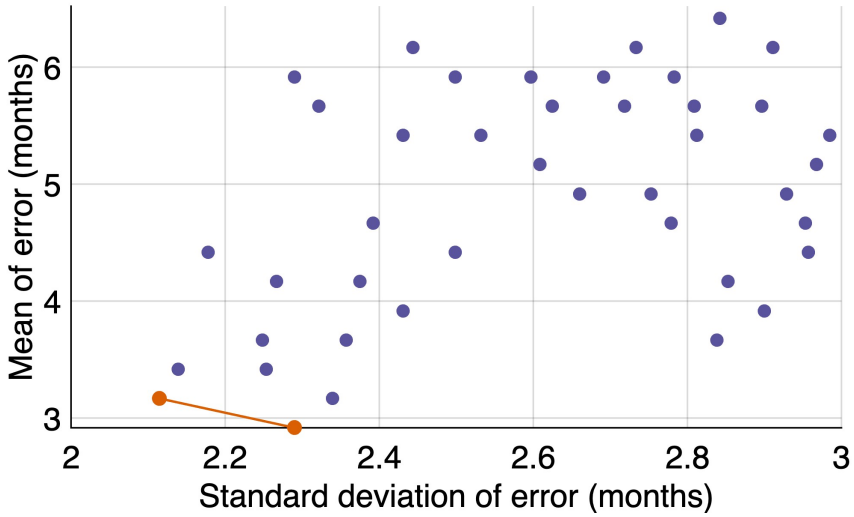
where $\mathcal{F}$ denotes the set of classifiers on the anticipation–precision frontier. This is the same selection rule used in Michaillat (2025): it focuses attention on the high-precision segment of the frontier and excludes classifiers whose implied timing uncertainty exceeds one year.

3.5.4 Finding the anticipation-precision frontier

We select classifiers along the anticipation–precision frontier, defined as those that minimize both the mean detection delay and its variability. This frontier highlights classifiers that achieve the best tradeoff between timely and reliable recession detection. Figure 3.3a plots the mean and standard deviation of detection errors across a large set of recession classifiers. The anticipation–precision frontier identifies the set of optimal classifiers that minimize detection delay for a given level of dispersion, or equivalently, minimize dispersion for a given level of delay. The downward-sloping frontier reveals a clear tradeoff: classifiers that achieve greater anticipation (more negative mean errors) exhibit substantially higher variability in detection timing, while classifiers with low dispersion detect recessions more consistently but with a modest delay.



(a) 4 Perfect Classifiers on the Anticipation–Precision Frontier



(b) 2 Perfect Classifiers satisfying $\sigma(k) < 3$ on the Anticipation–Precision Frontier

Figure 3.3: Anticipation and Precision of 137,457 Perfect Recession Classifiers in the United States, January 1947–December 2021

Notes: The figures plot the mean detection error against the standard deviation of detection errors for recession classifiers constructed using alternative indicators and thresholds. Each point represents a classifier. The orange frontier traces classifiers that maximize anticipation for a given level of dispersion, or equivalently maximize precision for a given level of anticipation. Panel (a) shows the full anticipation–precision frontier across all classifiers, while panel (b) focuses on the high-precision region satisfying $\sigma(k) < 3$. Detection errors are measured relative to recession start dates dated by the NBER.

The anticipation–precision frontier contains 2 classifiers out of the 137,457 statistically perfect classifiers. The most precise classifier has a mean detection error of 3.17 months and a standard deviation of detection errors of 2.11 months. The second frontier classifier detects recessions slightly earlier, with a mean detection error of 2.92 months, but at the cost of a modest reduction in precision, as reflected by a higher standard deviation of 2.29 months. Overall, both classifiers identify recessions 3.04 months after their official start dates while maintaining relatively precise timing. For comparison, Michaillat (2025) finds that the selected classifier ensemble for the United States detects recessions only 1.2 months after their onset.

3.5.5 Selecting the frontier classifiers with highest precision

There is no uniquely optimal classifier along the anticipation–precision frontier because the preferred tradeoff between early detection and precision depends on policymakers' preferences. Following Michaillat (2025), we therefore focus on the high-precision segment of the frontier by retaining all classifiers with a standard deviation of detection errors below 3 months. Under the assumption of normally distributed detection errors, this restriction implies a 95% confidence interval for recession start dates narrower than 12 months.

3.6 Detecting recessions

To build the recession detection algorithm, we follow Michaillat (2025) to aggregate the detection signals produced by the classifier ensemble. We then use the aggregated signals to compute the current recession risk.

Table 3.1: Classifier Performance Metrics

Smoothing method	Smooth. parameter	γ	β	Threshold ζ	Std. error	Mean error
SMA	2	0.0	1	0.0015	2.11	3.17
SMA	2	0.1	1	0.0027	2.29	2.92

Notes: Detection-error statistics are reported in months. `SMA` denotes simple moving average smoothing. The table reports classifiers selected from the anticipation–precision frontier subject to the restriction $\sigma(k) < 3$ months. The Box–Cox parameter γ governs the transformation of the recession indicator, while the turning parameter β determines the horizon used to compute turning points. A recession is identified when the resulting indicator crosses the threshold ζ from below. Detection errors are defined as the difference, in months, between estimated and official recession start dates dated by the NBER.

3.6.1 Empirical Construction of Recession Probabilities

Using the methodology of Michaillat (2025), we compute the detection error for each classifier k and recession j in the training sample.

$$\varepsilon_{k,j} = d_{k,j} - s_j,$$

where $d_{k,j}$ is the classifier's detection date and s_j is the NBER recession start date. For each classifier, these historical errors are summarized by their sample mean μ_k and standard deviation σ_k, which capture systematic anticipation or delay and timing precision, respectively.

When classifier k issues a real-time detection at date d_k, we treat the historical detection errors as an empirical forecast-error sample and approximate their distribution by a normal distribution with mean μ_k and variance σ_k^2, exactly as in Michaillat (2025). Letting

$$\varepsilon_k = d_k - s,$$

where s is the unknown start date of the current recession, the event $\{s \leq t\}$ is

equivalent to $\{\varepsilon_k \geq d_k - t\}$. Therefore,

$$\Pr(s \leq t \mid d_k) = \Pr(\varepsilon_k \geq d_k - t).$$

This probability is computed using only information from the training sample.

Whenever an individual recession indicator crosses its threshold, we infer the probability that the recession has already started from the distribution of the detection error. If classifier k is on average exactly on time and the detection error is symmetric, then the recession has started with probability 0.5 when the classifier is activated. If the classifier tends to detect recessions early on average, this probability is below 0.5; if the classifier tends to detect recessions late on average, it is above 0.5. In the months following detection, this probability converges to one according to the cumulative distribution function of the detection error.

More formally, suppose that classifier k detects the J recessions in the training sample. Each detection j yields a detection date $d(k, j)$ and a detection error $\varepsilon(k, j)$. As in Michaillat (2025), we assume that the detection error $\varepsilon(k)$ is normally distributed with mean $\mu(k)$ and standard deviation $\sigma(k)$:

$$\varepsilon(k) \sim \mathcal{N}\Big(\mu(k), \sigma^2(k)\Big).$$

Then, for any date $t \geq d(k)$, the probability that the start date s of the new recession occurred before t, conditional on classifier k having detected a recession

at date $d(k)$, is

$$P(k,t) = \Pr\Big(s < t \mid d(k), \mu(k), \sigma(k)\Big) \tag{3.1}$$

$$= \Pr\Big(d(k) - s > d(k) - t \mid \mu(k), \sigma(k)\Big) \tag{3.2}$$

$$= \Pr\Big(\varepsilon(k) > d(k) - t \mid \mu(k), \sigma(k)\Big). \tag{3.3}$$

Under the normal approximation, this becomes

$$P(k,t) = 1 - \Phi\left(\frac{d(k) - t - \mu(k)}{\sigma(k)}\right),$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution.

Lastly, we summarize the information from the $K = 2$ detection classifiers by constructing an ensemble recession probability. Each classifier produces a probability that a recession has started, and we average these probabilities to obtain a single aggregate measure. Figure 3.6 shows the resulting ensemble probability for the 1947–2021 training sample. For all 12 recessions, the probability rises in advance of the official start date, reaches 1 at the onset of the recession, and remains there until the recession has ended.

3.6.2 Application of Recession Detection to Current Data

Next, we apply the classifier estimated on the 1947–2021 training sample to the out-of-sample period 2022–2025 to assess recession risk in the United States economy. Figure 3.7 reports the estimated recession probability over this period. The classifier assigns essentially zero recession probability throughout 2022–2025, indicating no recession signal in the out-of-sample period.

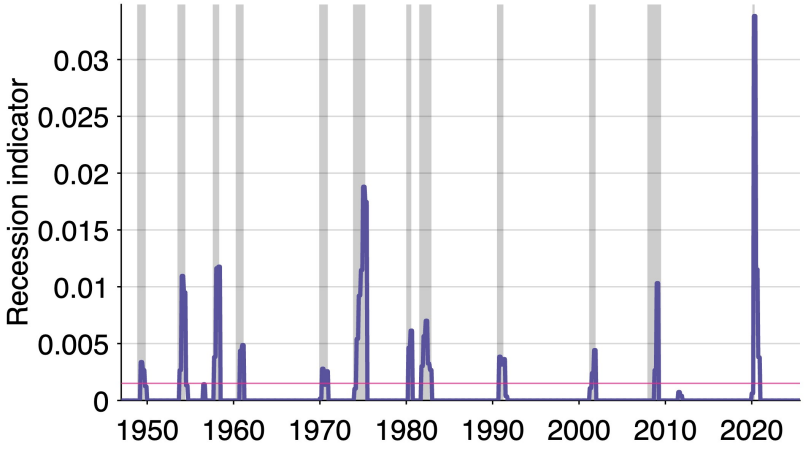


Figure 3.4: Classifier 1: Threshold 0.0015

Notes: The classifier specification corresponds to the first classifier reported in Table 2.4. The purple line represents the recession indicator underlying the classifier, while the pink line represents the corresponding recession threshold. Gray shaded regions indicate recessions dated by the NBER.

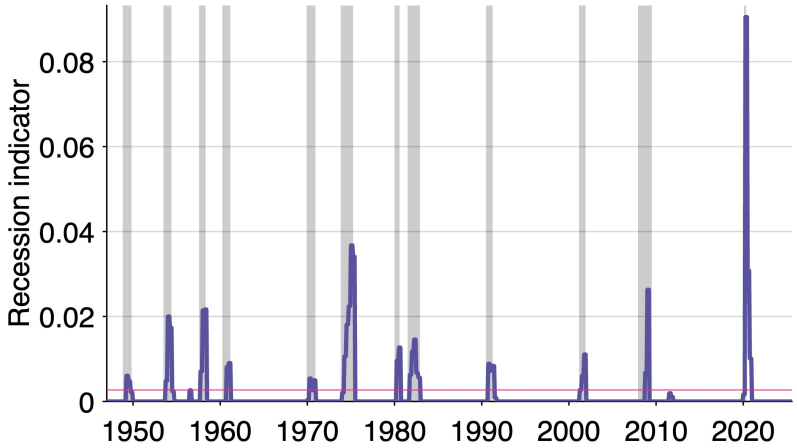


Figure 3.5: Classifier 2: Threshold 0.0027

Notes: The classifier specification corresponds to the second classifier reported in Table 2.4. The purple line represents the recession indicator underlying the classifier, while the pink line represents the corresponding recession threshold. Gray shaded regions indicate recessions dated by the NBER.

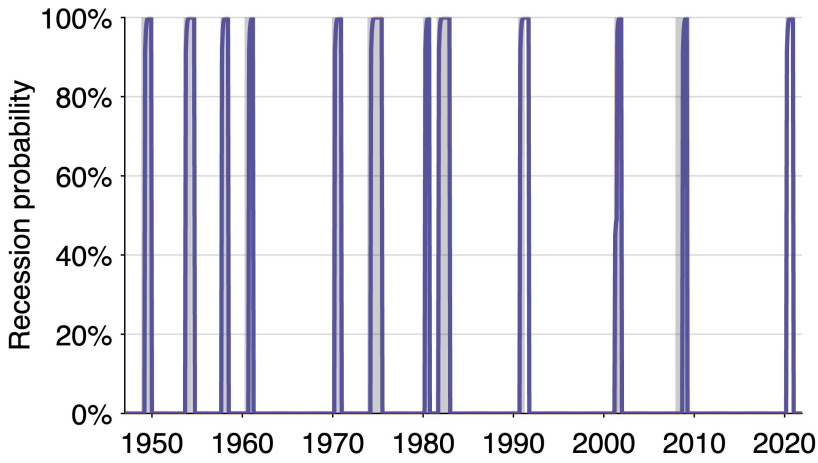


Figure 3.6: In-sample Recession Probability, January 1947–December 2021.

Notes: The figure plots the estimated probability of a recession over time. The thick purple line shows the average probability across the two classifiers in the selected ensemble. The thin orange lines show the probability from each classifier separately. Shaded areas indicate recession periods.

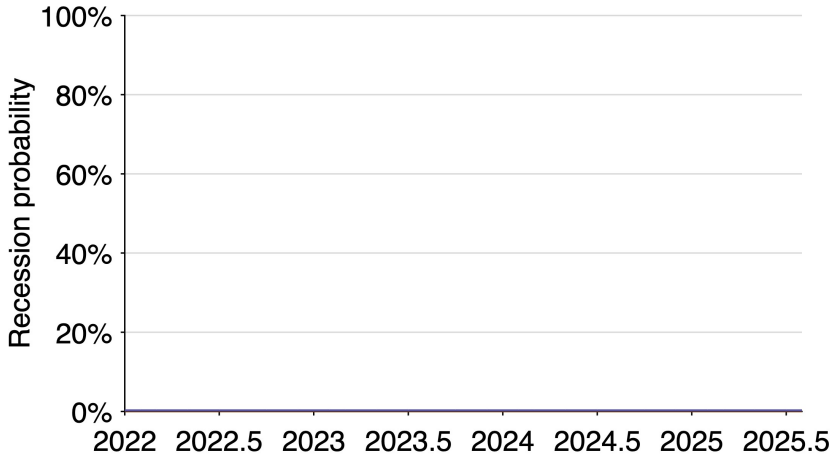


Figure 3.7: Out-of-sample Recession Probability, January 2022–December 2025

3.7 Evaluating the recession detection algorithm with backtests

To examine the reliability of the algorithm, we perform a series of out-of-sample backtests. In each exercise, the training sample is shortened, classifiers are selected using only information available within that training window, and performance is evaluated on recessions occurring after the training period. The results show that the algorithm remains broadly robust across several training windows. However, performance deteriorates when the training sample ends in 2005, suggesting that recession dynamics embedded in historical GDP vintages may not remain fully stable across subsequent business-cycle episodes. We therefore stop the backtesting exercise at the 2005 window.

The full-sample result shows that real-time GDP announcements contain enough information to recover historical turning points. The out-of-sample exercise is more demanding as it asks whether the same classifier can be selected before later recession episodes are observed. The weaker performance therefore reflects uncertainty in classifier selection across changing business-cycle environments, rather than a lack of recession information in GDP itself.

3.7.1 Backtesting from 2015

We train the algorithm using data through December 2014. The training sample contains 11 recessions, as summarized in Table 3.2. This procedure selects 5 classifiers from the high-precision segment of the anticipation–precision frontier, reported in Table C.1. We then evaluate these classifiers out of sample using data from January 2015 through December 2025. During this test period, all 5 classifiers detect the 2020 recession with no false positives. The detection is also

Table 3.2: Out-of-Sample Backtesting Results

Metric	**2015**
Number of classifiers	5
Training sample detection error (months)	
Mean	4.7
Standard deviation	2.0
Minimum	2.2
Maximum	9.4
Testing sample detection error (months)	
Mean	1.0
Standard deviation	0.0
Minimum	1.0
Maximum	1.0
False positives	0
False negatives	0
Training recessions	11
Testing recessions	1

Note: The table reports the performance of the selected classifier ensembles in each backtest exercise. The training sample is Jan 1947–Dec 2014. The testing period is January 2015–December 2025. Detection errors are measured in months relative to recession onset dates.

timely. On average, the recession is signaled only 1 month after its official onset.

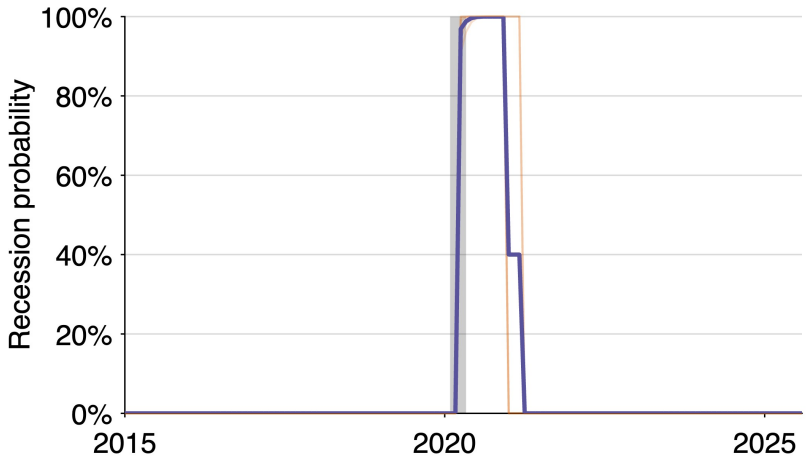


Figure 3.8: Out-of-sample U.S. Recession Probability, January 2015–December 2025

3.7.2 Backtesting from 2005

We train the algorithm using data through December 2004. The training sample contains ten recessions. This procedure selects one classifier from the high-precision segment of the anticipation–precision frontier, reported in Table C.2. We then evaluate the selected classifier out of sample using data from January 2005 through December 2025.

The out-of-sample performance deteriorates relative to earlier backtesting exercises. Although the classifier exhibits anticipatory detection, with a mean detection error of approximately -5.5 months, it also generates four false positives during the testing period. This implies that the classifier signals recessions nearly five months before the official recession onset dates on average, but at the cost of reduced precision. The standard deviation of the detection error is approximately 3.5 months, which remains relatively moderate given the difficulty of identifying business-cycle turning points in real-time. Nevertheless, the increase in false pos-

itives suggests that classifiers calibrated on earlier postwar recession dynamics may not generalize fully to later business-cycle episodes. We therefore stop the backtesting exercise at the 2005 window.

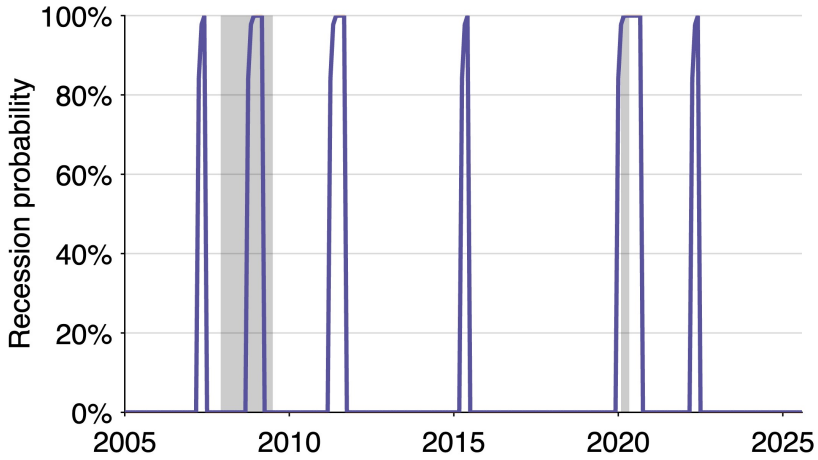


Figure 3.9: Out-of-sample U.S. Recession Probability, January 2005–December 2025

3.8 Conclusion

A natural question is why one should rely on real-time GDP rather than labor-market indicators to proxy business-cycle turning points. Labor-market indicators are often timely and informative, particularly when employment conditions deteriorate before output declines are fully recorded. However, GDP remains the broadest summary measure of aggregate economic activity and is closely tied to the concept of business-cycle turning points used in official recession dating. For this reason, evaluating whether real-time GDP announcements alone can approximate NBER-style recession dates provides a useful benchmark. The purpose is not to replace labor-market indicators, but rather to examine whether real-time GDP itself contains sufficient information to systematically identify turning points in economic activity.

The results in this paper suggest that real-time GDP announcements contain substantial information about business-cycle turning points and can closely approximate official recession dates using only information available at each point in time. While the selected classifiers reproduce historical recession episodes with a high degree of accuracy in sample, the weaker out-of-sample performance indicates that the relationship between GDP dynamics and recession timing may vary across business-cycle episodes. Nevertheless, the findings demonstrate that real-time GDP can still provide a transparent and operational framework for real-time recession dating, particularly in economies where official business-cycle chronologies are unavailable or announced with substantial delays.

Appendix A

This appendix presents supplementary material for Chapter 1 on the Beveridgean unemployment gap with part-time employment, including notation relative to Michaillat and Saez (2021a), additional U.S. and Japanese results, and robustness tables.

A.1 Notation relative to Michaillat and Saez (2021)

This paper follows the Beveridgean unemployment-gap framework of Michaillat and Saez (2021a), but uses slightly different notation. In Michaillat and Saez (2021a), the social value of nonwork is denoted by ζ and recruiting cost by κ. In this paper, I denote the same objects by z and c:

$$z \equiv \zeta, \qquad c \equiv \kappa.$$

In Michaillat and Saez (2021a), the tangency condition is stated as follows:

> **Proposition 1.** In a Beveridge diagram, efficiency is achieved at the point where the Beveridge curve is tangent to an isowelfare curve. Hence, the efficient unemployment rate is implicitly defined by
>
> $$v'(u) = -\frac{1-\zeta}{\kappa},$$

where $\zeta < 1$ is the social value of nonwork and $\kappa > 0$ is the recruiting cost.

Using the notation of this paper, this condition becomes

$$v'(u) = -\frac{1-z}{c}. \tag{A.1}$$

With an isoelastic Beveridge curve, Michaillat and Saez (2021a) then obtain the following efficient-tightness condition:

Proposition 2. Consider a point on the Beveridge curve with tightness θ, Beveridge elasticity ϵ, social value of nonwork ζ, and recruiting cost κ. Tightness is inefficiently high if

$$\theta > \frac{1-\zeta}{\kappa\epsilon},$$

inefficiently low if

$$\theta < \frac{1-\zeta}{\kappa\epsilon},$$

and efficient if

$$\theta = \frac{1-\zeta}{\kappa\epsilon}.$$

Under the notation used in this paper, the full-time benchmark is

$$\theta^* = \frac{1-z}{c\epsilon}. \tag{A.2}$$

The part-time extension replaces the full-time employment term 1 with effective employment input,

$$\alpha + (1-\alpha)\left[\gamma + (1-\gamma)z\right],$$

where α is the full-time employment share and γ is the relative-hours parameter.

The efficient-tightness condition with part-time employment is therefore

$$\theta^* = \frac{\alpha + (1-\alpha)\left[\gamma + (1-\gamma)z\right] - z}{c\epsilon}. \tag{A.3}$$

When $\alpha = 1$, equation (A.3) collapses to the full-time benchmark in equation (A.2).

A.2 Additional results

A.2.1 Additional results for the United States

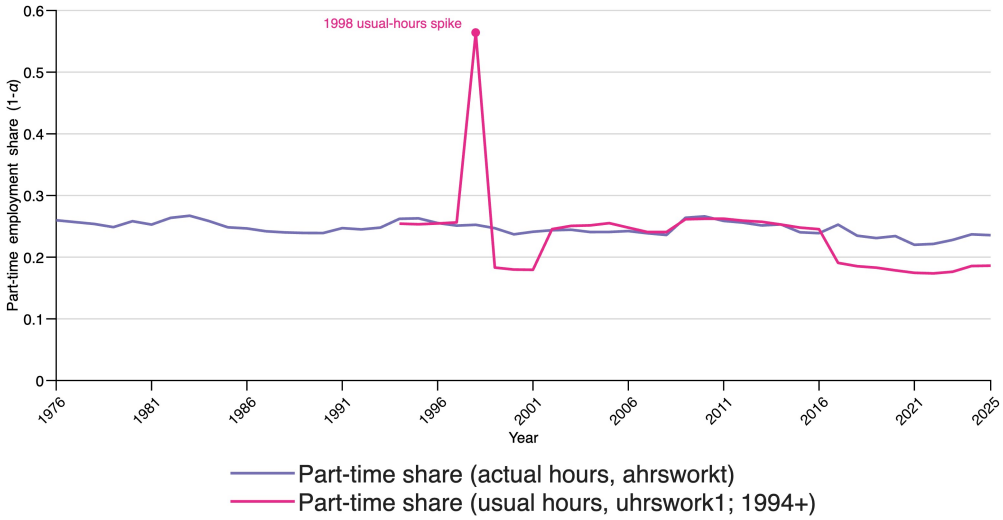
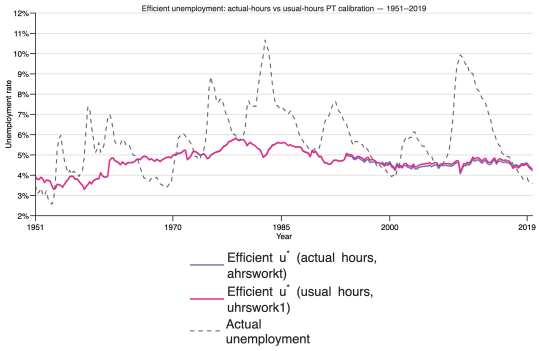
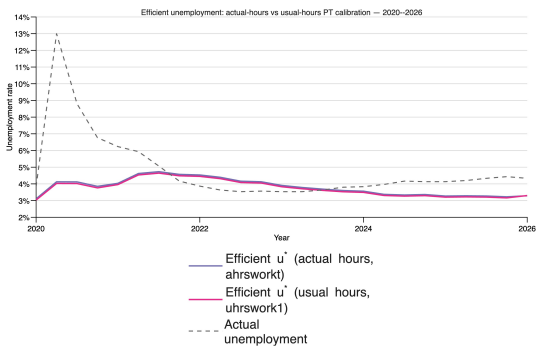


Figure A.1: Part-time employment shares based on actual and usual hours in the CPS ASEC.

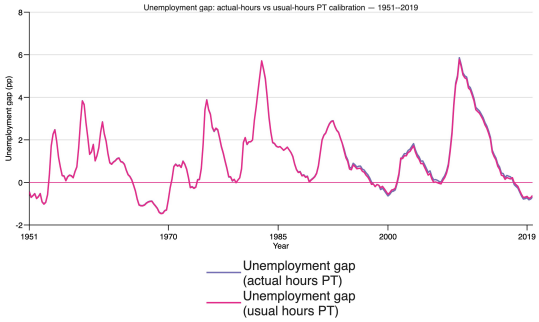
Notes: The figure compares part-time employment shares constructed using actual hours worked and usual hours worked. Workers are classified as part-time if their reported hours are fewer than 35 per week. The usual-hours measure produces a large spike in 1998 that is absent from the actual-hours measure.



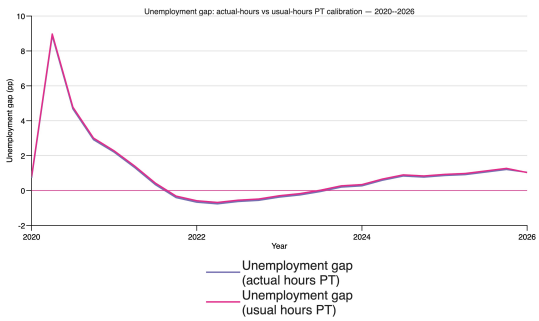
(a) Efficient unemployment rate, 1951–2019



(b) Efficient unemployment rate, 2020–2026



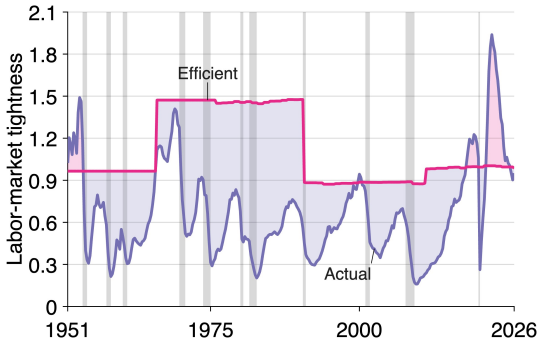
(c) Unemployment gap, 1951–2019



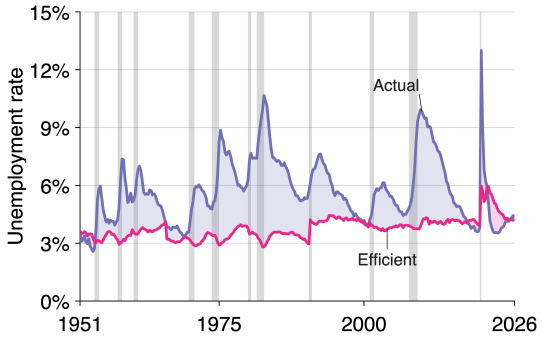
(d) Unemployment gap, 2020–2026

Figure A.2: Robustness to actual-hours and usual-hours part-time classifications in the United States.

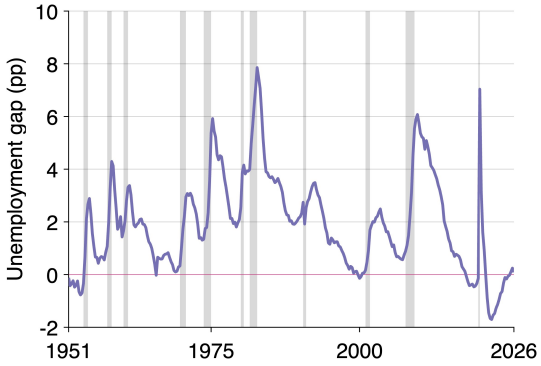
Notes: The figure compares efficient unemployment rates and unemployment gaps constructed using part-time employment classifications based on actual hours worked and usual hours worked. Workers are classified as part-time if reported hours are fewer than 35 per week. The actual-hours measure uses `ahrsworkt`, while the usual-hours measure uses `uhrswork1`.



(a) Efficient labor-market tightness with total part-time employment, 1951–2026



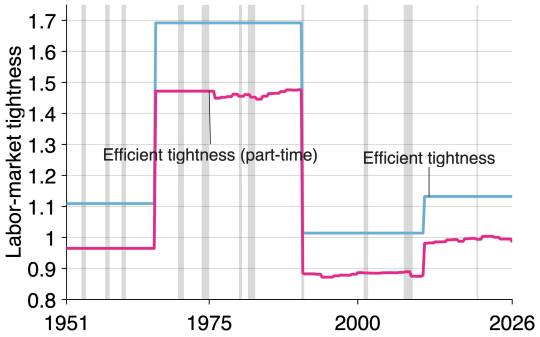
(b) Efficient unemployment rate with total part-time employment, 1951–2026



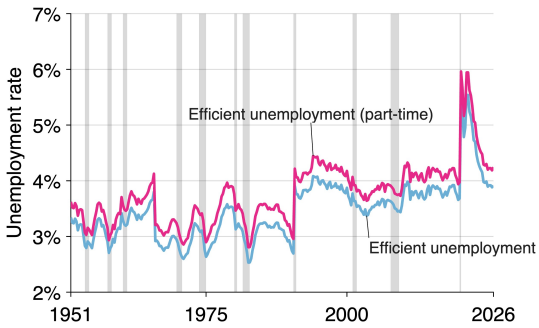
(c) Unemployment gap with total part-time employment, 1951–2026

Figure A.3: Efficient labor-market tightness, efficient unemployment, and unemployment gap in the United States, 1951–2026.

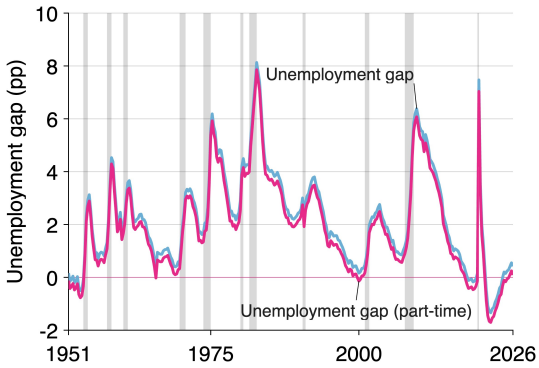
Notes: The figure reports efficient labor-market tightness, the efficient unemployment rate, and the unemployment gap under the extended framework that incorporates total part-time employment. The unemployment gap is measured as the difference between the actual unemployment rate and the efficient unemployment rate implied by the model.



(a) Efficient labor-market tightness comparison, 1951–2026



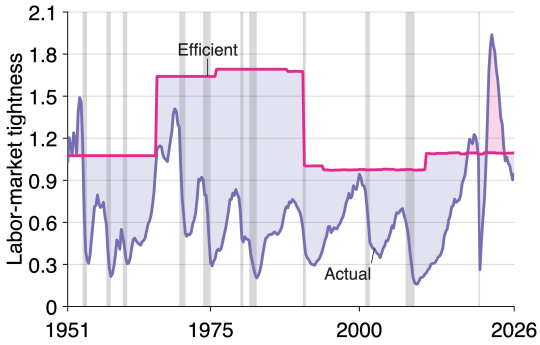
(b) Efficient unemployment rate comparison, 1951–2026



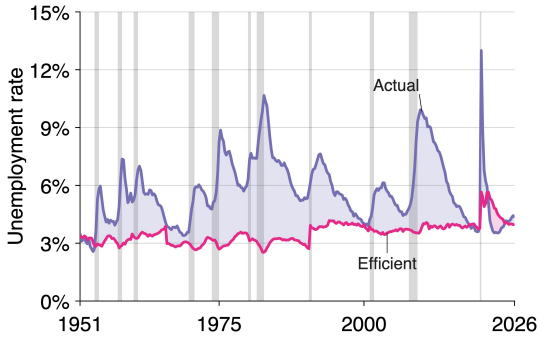
(c) Unemployment gap comparison, 1951–2026

Figure A.4: Comparison of efficient labor-market tightness, efficient unemployment, and unemployment gaps with and without total part-time employment in the United States, 1951–2026.

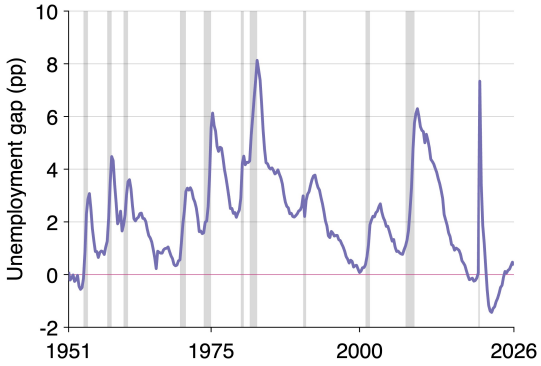
Notes: The figure compares the baseline framework of Michaillat and Saez (2021a), which treats employment as homogeneous, with the extended framework that incorporates total part-time employment.



(a) Efficient labor-market tightness with involuntary part-time employment, 1951–2026



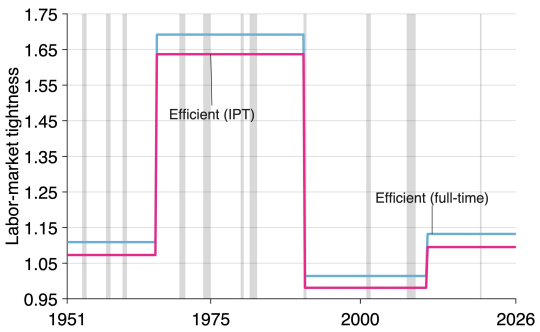
(b) Efficient unemployment rate with involuntary part-time employment, 1951–2026



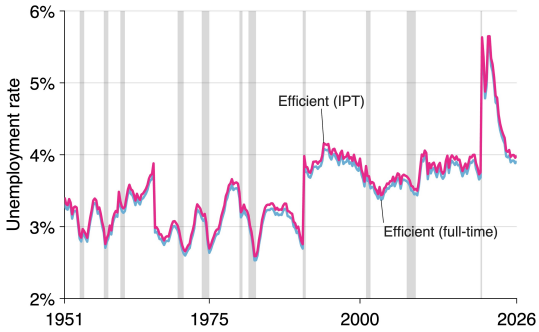
(c) Unemployment gap with involuntary part-time employment, 1951–2026

Figure A.5: Efficient labor-market tightness, efficient unemployment, and unemployment gap in the United States, 1951–2026.

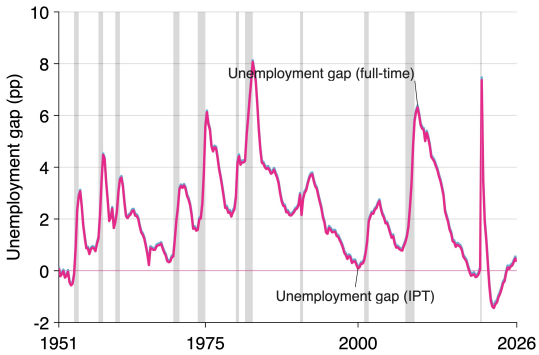
Notes: The figure reports efficient labor-market tightness, the efficient unemployment rate, and the unemployment gap under the extended framework that incorporates involuntary part-time employment. The unemployment gap is measured as the difference between the actual unemployment rate and the efficient unemployment rate implied by the model.



(a) Efficient labor-market tightness comparison, 1951–2026



(b) Efficient unemployment rate comparison, 1951–2026

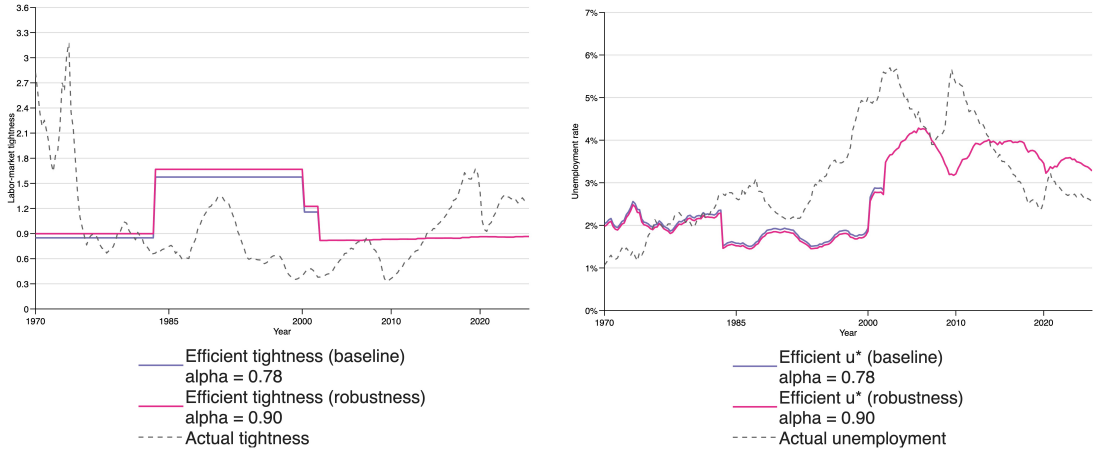


(c) Unemployment gap comparison, 1951–2026

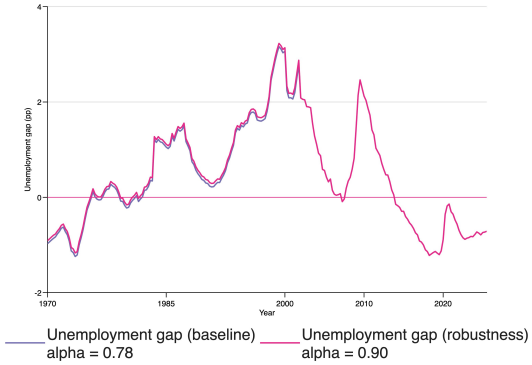
Figure A.6: Comparison of efficient labor-market tightness, efficient unemployment, and unemployment gaps with and without involuntary part-time employment in the United States, 1951–2026.

Notes: The figure compares the baseline framework of Michaillat and Saez (2021a), which treats employment as homogeneous, with the extended framework that incorporates involuntary part-time employment.

A.2.2 Additional Japanese results



(a) Efficient labor-market tightness **(b)** Efficient unemployment rate



(c) Unemployment gap

Figure A.7: Robustness to the pre-2002 full-time employment share in Japan.

Notes: The figure compares the baseline calibration, which sets $\alpha = 0.78$ before 2002, with an alternative calibration using $\alpha = 0.90$ before 2002. From 2002 onward, both calibrations use the observed OECD part-time employment share.

Table A.1: Summary Statistics of Labor-Market Tightness in Japan

Statistic	Actual	Efficient (Full-time)	Efficient (Part-time)
Mean	0.965	1.338	1.073
SD	0.525	0.299	0.334
Min	0.329	0.940	0.818
Max	3.179	1.744	1.575
N	223	223	223

Notes: The sample covers 223 quarterly observations from 1970Q1 to 2025Q3. The full-time specification assumes $\alpha = 1$. The part-time specification assumes $\alpha = 0.78$ before 2002 and uses the OECD part-time employment share thereafter. Parameters are $z = 0.26$, $c = 0.92$, and $\gamma = 0.56$.
Source: OECD Database, Statistics Bureau of Japan, and author's calculations.

Table A.2: Summary Statistics of Actual and Efficient Unemployment Rates in Japan

Statistic	Actual	Efficient (Full-time)	Efficient (Part-time)
Mean	3.16%	2.28%	2.71%
SD	1.20%	0.58%	0.95%
Min	1.07%	1.40%	1.51%
Max	5.70%	3.26%	4.28%
N	223	223	223

Notes: The sample covers 223 quarterly observations from 1970Q1 to 2025Q3. The full-time specification assumes $\alpha = 1$. The part-time specification assumes $\alpha = 0.78$ before 2002 and uses the OECD part-time employment share thereafter. Parameters are $z = 0.26$, $c = 0.92$, and $\gamma = 0.56$.
Source: OECD Database, Statistics Bureau of Japan, and author's calculations.

Table A.3: Summary Statistics of the Unemployment Gap in Japan

Statistic	Gap (Full-time)	Gap (Part-time)
Mean	0.88 pp	0.45 pp
SD	1.07 pp	1.09 pp
Min	-1.11 pp	-1.24 pp
Max	3.28 pp	3.16 pp
N	223	223

Notes: The unemployment gap is defined as the difference between the actual unemployment rate and the efficient unemployment rate. The sample covers 223 quarterly observations from 1970Q1 to 2025Q3. The full-time specification assumes $\alpha = 1$. The part-time specification assumes $\alpha = 0.78$ before 2002 and uses the OECD part-time employment share thereafter. Parameters are $z = 0.26$, $c = 0.92$, and $\gamma = 0.56$.
Source: OECD Database, Statistics Bureau of Japan, and author's calculations.

A.2.3 Robustness test tables

Table A.4: Summary Statistics for Efficient Unemployment and Beveridge-Elasticity Uncertainty

	United States (1951Q1–2019Q4; $N = 276$)	Japan (1970Q1–2025Q2; $N = 223$)
Point estimate u^* (percent)		
Mean	4.66	2.72
Min	3.25	1.51
Max	5.84	4.30
Sensitivity to 95% confidence interval for ε		
Avg. gap, lower bound vs. baseline (pp)	0.6	0.33
Avg. gap, upper bound vs. baseline (pp)	0.5	0.28
Max. deviation from baseline (pp)	1.3	0.55
Full envelope: min–max (percent)	2.85–6.55	1.11–4.65

Notes: The shaded band varies the Beveridge elasticity within its 95% confidence interval at each quarter; all other sufficient statistics are held at baseline ($z = 0.26$, $c = 0.92$). *Source*: Author's calculations.

Table A.5: Summary Statistics for Efficient Unemployment and Social Value of Nonwork Uncertainty

	United States (1951Q1–2019Q4; $N = 276$)	Japan (1970Q1–2025Q2; $N = 223$)
Point estimate u^* (percent)		
Mean	4.66	2.72
Min	3.25	1.51
Max	5.84	4.30
Sensitivity to $\zeta \in [0.03, 0.49]$		
Avg. gap, lower bound vs. baseline (pp)	0.6	0.42
Avg. gap, upper bound vs. baseline (pp)	1.0	0.70
Max. deviation from baseline (pp)	1.3	1.11
Full envelope: min–max (percent)	2.81–7.15	1.25–5.41

Notes: The shaded band varies z over its calibration range; ϵ and c are held at their baseline estimates. *Source*: Author's calculations.

Table A.6: Summary Statistics for Efficient Unemployment and Recruiting-Cost Uncertainty

	United States (1951Q1–2019Q4; $N = 276$)	Japan (1970Q1–2025Q2; $N = 223$)
Point estimate u^* (percent)		
Mean	4.66	2.72
Min	3.25	1.51
Max	5.84	4.30
Sensitivity to $c \in [\frac{2}{3} \times 0.92, \frac{4}{3} \times 0.92]$		
Avg. gap, lower bound vs. baseline (pp)	0.9	0.60
Avg. gap, upper bound vs. baseline (pp)	0.8	0.53
Max. deviation from baseline (pp)	1.2	0.95
Full envelope: min–max (percent)	2.61–6.83	1.14–5.13

Notes: The shaded band varies c over its calibration range; ϵ and z are held at their baseline estimates. *Source*: Author's calculations.

Table A.7: Summary Statistics for Inverse-Optimum Beveridge Elasticity

	United States (1951Q1–2019Q4; $N = 276$)	Japan (1970Q1–2025Q3; $N = 223$)
Mean of ϵ^*	1.42	0.84
Min of ϵ^*	0.48	0.23
Max of ϵ^*	4.35	2.06
Mean of calibrated ϵ	0.91	0.63
Inside 95% CI (% of quarters)	23.0	24.2
Below CI (% of quarters)	16.0	25.6
Above CI (% of quarters)	60.9	50.2

Notes: ϵ^* is the inverse-optimum Beveridge elasticity implied by observed labor-market tightness, holding $z = 0.26$, $c = 0.92$, and part-time parameters at baseline. The calibrated U.S. elasticity reported here, $\epsilon = 0.91$, is the benchmark value from the five-break pre-COVID calibration in Michaillat and Saez (2021a), used for comparison with the original framework. It differs from the regime-specific elasticities reported in Table 1.5, which are estimated using the full sample and a different break structure.
Source: Author's calculations.

Table A.8: Summary Statistics for Inverse-Optimum Social Value of Nonwork

	United States (1951Q1–2019Q4; $N = 276$)	Japan (1970Q1–2025Q3; $N = 223$)
Mean of z^*	0.41	0.27
Min of z^*	−0.48	−1.77
Max of z^*	0.86	0.83
Calibrated z	0.26	0.26
Inside [0.03, 0.49] (% of quarters)	42.0	42.2
Below range (% of quarters)	12.0	22.0
Above range (% of quarters)	46.0	35.9
$z^* < 0$ (% of quarters)	12.0	Negative z^*

Notes: Negative values of z^* occur when observed labor-market tightness is high enough that efficiency would require nonwork to have negative social value. The U.S. minimum occurs in 1969Q1 and the Japan minimum occurs in 1973Q4. U.S. maximum in 2009Q4; Japan maximum in 1999Q2.
Source: Author's calculations.

Table A.9: Summary Statistics for Inverse-Optimum Recruiting Cost

	United States (1951Q1–2019Q4; $N = 276$)	Japan (1970Q1–2025Q3; $N = 223$)
Mean of c^*	1.46	1.34
Min of c^*	0.46	0.25
Max of c^*	4.75	4.11
Calibrated c	0.92	0.92
Inside $[\frac{2}{3} \times 0.92, \frac{4}{3} \times 0.92]$ (% of quarters)	42.0	46.2
Below range (% of quarters)	7.0	13.9
Above range (% of quarters)	51.0	39.9

Notes: c^* is the inverse-optimum recruiting cost, holding ϵ, z, and part-time parameters at baseline.
Source: Author's calculations.

Appendix B

This appendix presents supplementary material for Chapter 2 on recession detection in Japan using labor market data, including extended backtest results, classifier ensembles, and recession-probability figures.

Table B.1: Detected Recession Starts by Threshold for the Main Ensemble (continued)

Ensemble Member	Threshold	Recession	Official Start	Official Trough	Detection Start	Error (Months)
			Threshold = 0.0054			
4	0.0054	1	1970.58	1971.92	1970.25	-4
4	0.0054	2	1973.92	1975.17	1974.08	2
4	0.0054	3	1977.08	1977.75	1977.08	0
4	0.0054	4	1980.17	1983.08	1980.50	4
4	0.0054	5	1985.50	1986.83	1985.50	0
4	0.0054	6	1991.17	1993.75	1991.42	3
4	0.0054	7	1997.42	1999.00	1997.33	-1
4	0.0054	8	2000.92	2002.00	2000.92	0
4	0.0054	9	2008.17	2009.17	2007.67	-6
4	0.0054	10	2012.25	2012.83	2012.17	-1
4	0.0054	11	2018.83	2020.33	2018.83	0
			Threshold = 0.0073			
5	0.0073	1	1970.58	1971.92	1970.25	-4
5	0.0073	2	1973.92	1975.17	1974.08	2
5	0.0073	3	1977.08	1977.75	1977.00	-1
5	0.0073	4	1980.17	1983.08	1980.50	4
5	0.0073	5	1985.50	1986.83	1985.50	0
5	0.0073	6	1991.17	1993.75	1991.42	3
5	0.0073	7	1997.42	1999.00	1997.33	-1
5	0.0073	8	2000.92	2002.00	2000.92	0
5	0.0073	9	2008.17	2009.17	2007.67	-6
5	0.0073	10	2012.25	2012.83	2012.17	-1
5	0.0073	11	2018.83	2020.33	2018.83	0
			Threshold = 0.0072			
6	0.0072	1	1970.58	1971.92	1970.25	-4
6	0.0072	2	1973.92	1975.17	1974.08	2
6	0.0072	3	1977.08	1977.75	1977.00	-1
6	0.0072	4	1980.17	1983.08	1980.50	4
6	0.0072	5	1985.50	1986.83	1985.50	0
6	0.0072	6	1991.17	1993.75	1991.42	3
6	0.0072	7	1997.42	1999.00	1997.33	-1
6	0.0072	8	2000.92	2002.00	2000.92	0
6	0.0072	9	2008.17	2009.17	2007.58	-7
6	0.0072	10	2012.25	2012.83	2012.17	-1
6	0.0072	11	2018.83	2020.33	2018.83	0

Notes: Continued from Table 2.3. The table reports recession detection dates for Japan under alternative ensemble thresholds. Official start dates correspond to the benchmark recession chronology dated by the Economic and Social Research Institute (ESRI). Detection start denotes the first month in which the recession indicator crosses the specified threshold. Detection errors are measured as the difference, in months, between the detected and official recession start dates. Negative values indicate early detection, while positive values indicate delayed detection.

B.1 Additional Results for 1985-2015 Backtests

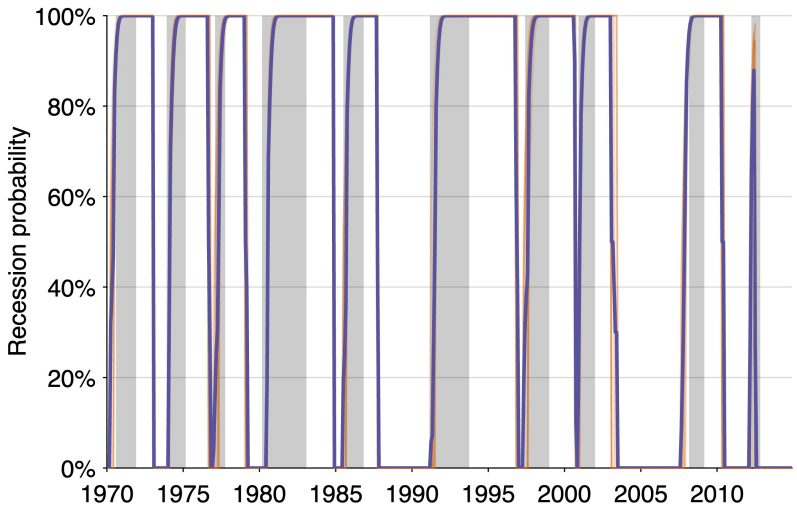


Figure B.1: Out-of-sample Japan recession probability, January 2015–December 2025

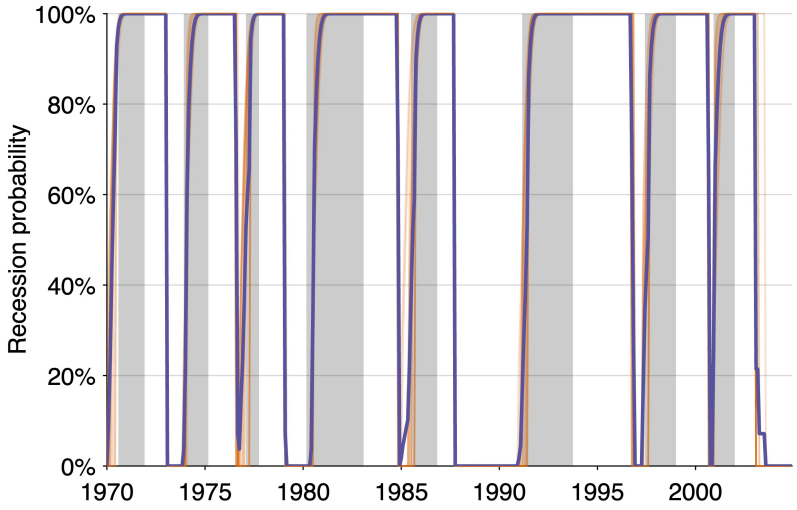


Figure B.2: Out-of-sample Japan recession probability, January 2005–December 2025

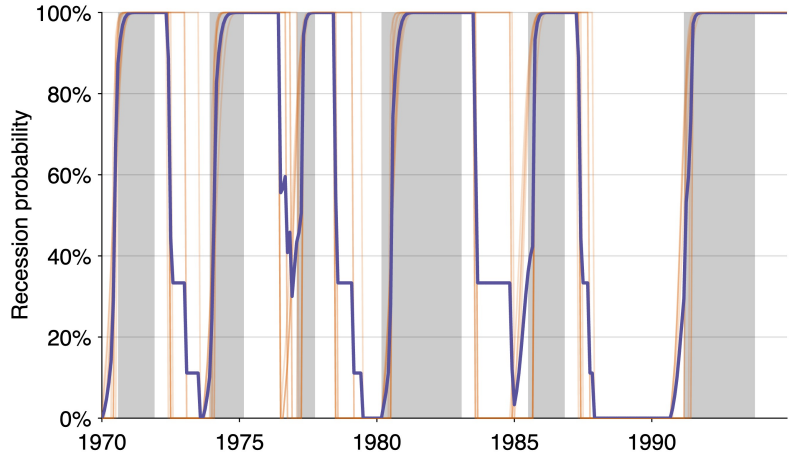


Figure B.3: Out-of-sample Japan recession probability, January 1995–December 2025

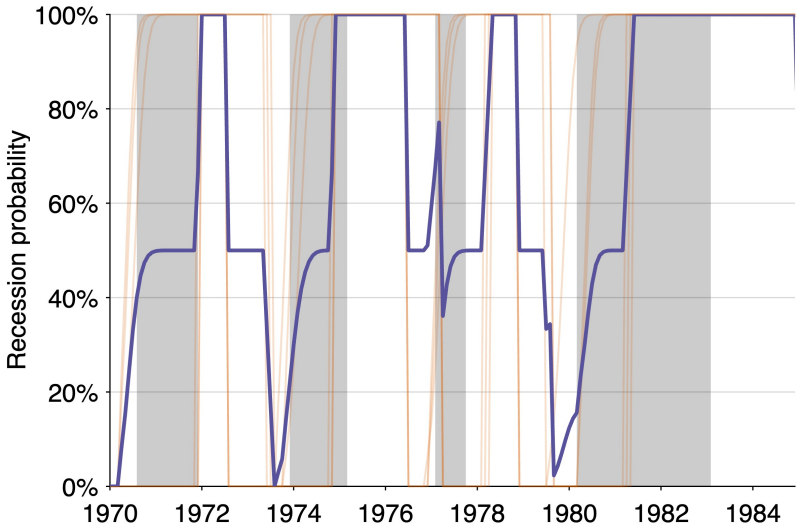


Figure B.4: Out-of-sample Japan recession probability, January 1985–December 2025

Table B.2: Classifier Ensemble Selected in the 2015 Backtest

	Smoothing method	Smoothing parameter	Curving parameter	Turning parameter	Mixing method	Mixing parameter	Threshold value	Standard error	Mean error
(1)	SMA	5	0.5	11	linear	1.0	0.0034	2.315	2.2
(2)	SMA	5	0.6	11	linear	1.0	0.0023	2.508	2.1
(3)	SMA	5	0.7	11	linear	1.0	0.0014	2.530	2.0
(4)	SMA	4	0.3	11	linear	1.0	0.0080	2.617	1.5
(5)	SMA	6	0.8	8	linear	0.9	0.0005	2.638	1.2
(6)	SMA	7	0.0	9	linear	0.9	0.0043	2.693	0.5
(7)	SMA	5	0.1	9	linear	0.8	0.0056	2.821	0.2
(8)	SMA	5	0.0	9	linear	0.9	0.0081	2.844	0.1
(9)	SMA	5	0.1	9	linear	0.8	0.0054	2.865	-0.3
(10)	SMA	5	0.0	9	linear	0.8	0.0073	2.871	-0.4

Notes: The table reports the classifier ensemble selected in the 2015 backtest. The ensemble is restricted to classifiers satisfying a high-precision criterion based on the standard deviation of detection errors. Detection-error statistics are reported in months. `SMA` denotes simple moving average smoothing. The Box–Cox parameter γ governs the transformation of the recession indicator, the turning parameter β determines the turning horizon, and indicators are combined linearly using weight δ. A recession is identified when the resulting indicator crosses the threshold ζ from below. Detection errors are defined as the difference, in months, between estimated and official recession start dates dated by the Economic and Social Research Institute (ESRI).

Table B.3: Classifier Ensemble Selected in the 2005 Backtest

	Smoothing method	Smoothing parameter	Curving parameter	Turning parameter	Mixing method	Mixing parameter	Threshold value	Standard error	Mean error
(1)	SMA	4	0.6	11	linear	1.0	0.0026	1.728	2.625
(2)	SMA	4	0.8	8	linear	0.8	0.0012	1.732	2.500
(3)	SMA	4	0.7	8	linear	0.8	0.0016	1.996	2.375
(4)	SMA	4	0.6	8	linear	0.8	0.0024	2.278	2.250
(5)	SMA	4	1.0	8	minmax	0.9	0.0001	2.288	0.625
(6)	SMA	5	0.1	9	linear	0.8	0.0054	2.345	0.500
(7)	SMA	5	0.0	9	linear	0.8	0.0072	2.395	0.375
(8)	SMA	6	0.5	8	linear	0.7	0.0009	2.421	0.125
(9)	SMA	6	0.3	8	linear	0.6	0.0018	2.449	0.000
(10)	SMA	5	0.0	9	linear	0.6	0.0054	2.643	-0.375
(11)	SMA	6	0.5	8	linear	0.4	0.0005	2.739	-0.500
(12)	SMA	5	0.0	9	minmax	0.9	0.0009	2.781	-0.625
(13)	SMA	5	0.0	9	linear	0.3	0.0027	2.817	-1.750
(14)	SMA	5	0.0	9	linear	0.2	0.0019	2.958	-2.000

Notes: The ensemble comprises the classifiers selected from the anticipation–precision frontier in the 2005 backtest. The ensemble is restricted to classifiers satisfying a high-precision criterion based on the standard deviation of detection errors. Detection-error statistics are reported in months. `SMA` denotes simple moving average smoothing. The Box–Cox parameter γ governs the transformation of the recession indicator, the turning parameter β determines the turning horizon, and indicators are combined using either linear weighting or the minimum–maximum (`minmax`) method with weight δ. A recession is identified when the resulting indicator crosses the threshold ζ from below. Detection errors are defined as the difference, in months, between estimated and official recession start dates dated by the Economic and Social Research Institute (ESRI).

B.2 Additional results for the 1985–2015 back-tests excluding the 2012 recession episode

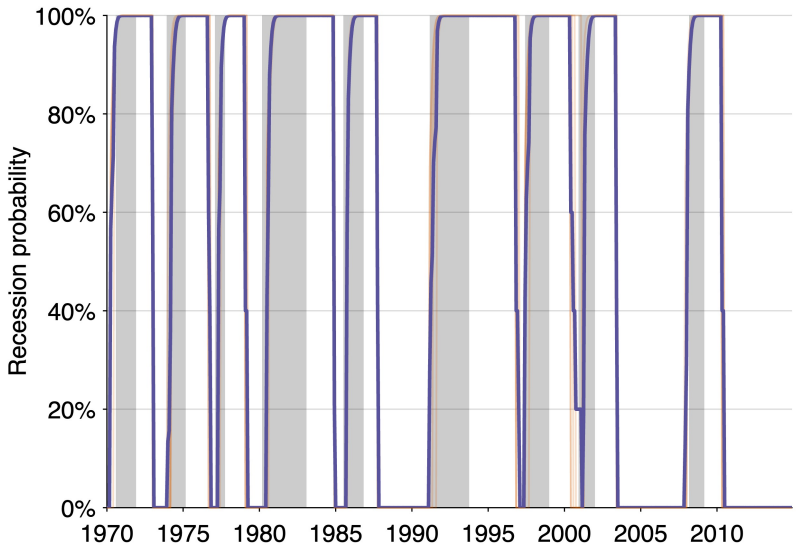


Figure B.5: In-sample recession probability, January 1970–December 2014

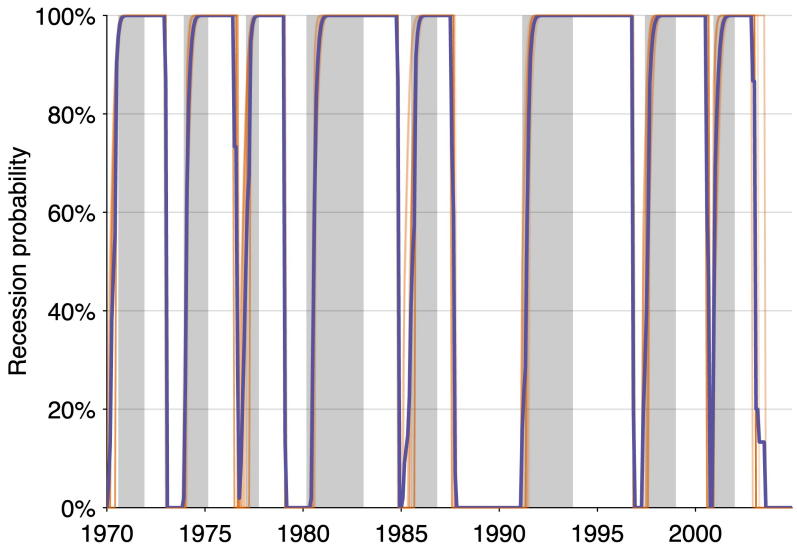


Figure B.6: In-sample recession probability, January 1970–December 2004

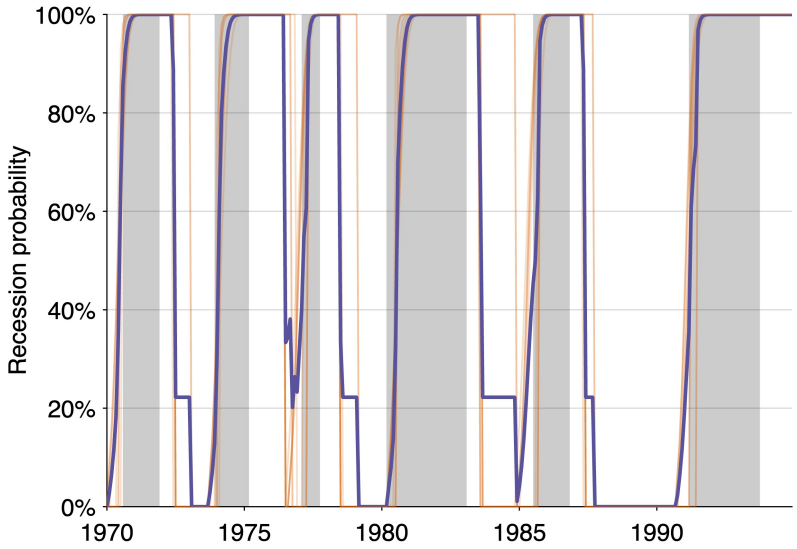


Figure B.7: In-sample recession probability, January 1970–December 1994

Appendix C

This appendix presents supplementary material for Chapter 3 on recession detection using real-time GDP data, including additional anticipation–precision frontier results, classifier ensembles from the backtests, and in-sample recession-probability figures.

C.1 Additional results

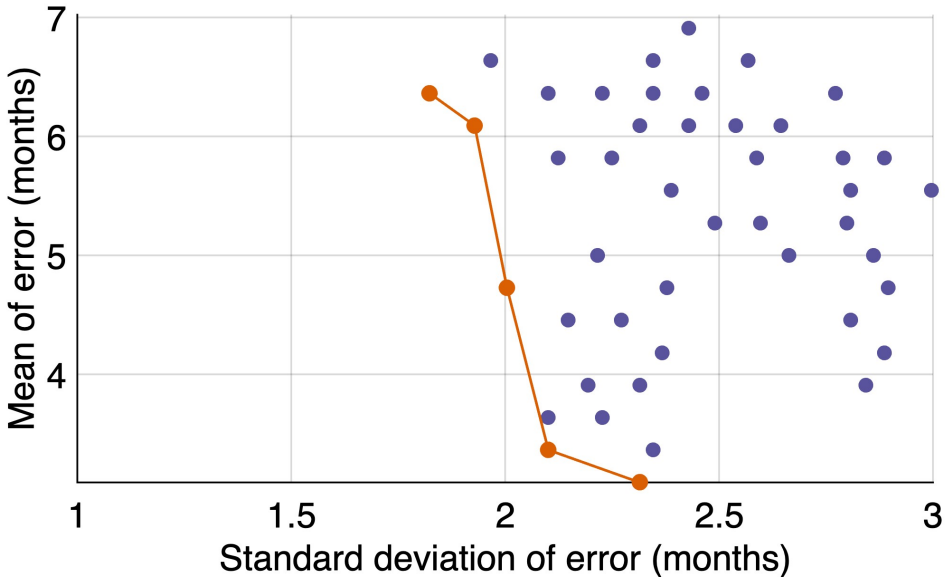


Figure C.1: Anticipation Precision Frontier 2015

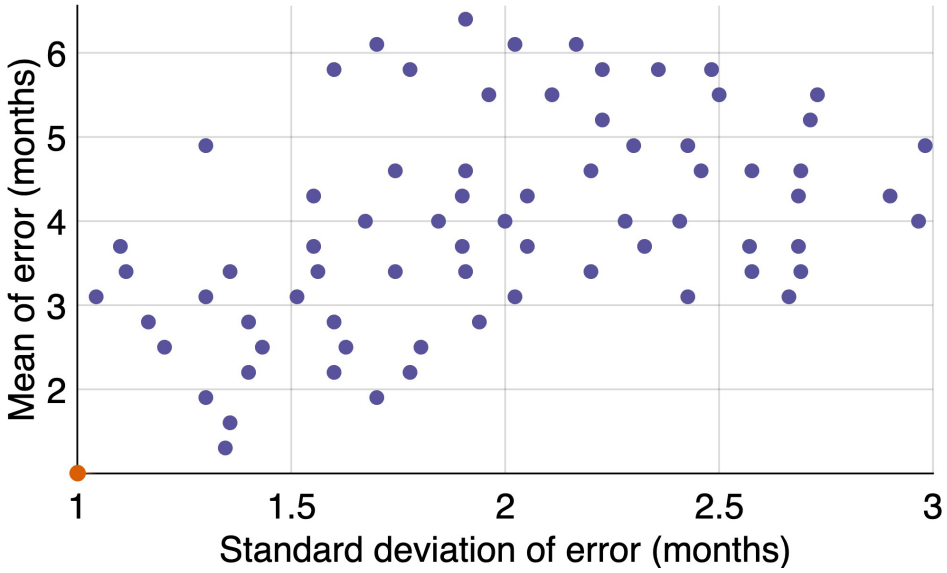


Figure C.2: Anticipation Precision Frontier 2005

Table C.1: Classifier Ensemble Selected in the 2015 Backtest

	Smooth		Curving	Turning	Threshold	Standard	Mean
Classifier	Method	Param	Method	Param	Value	Error	Error
(1)	SMA	3	0.1	6	0.0039	1.823	6.364
(2)	SMA	3	0.0	6	0.0014	1.928	6.091
(3)	SMA	2	0.2	6	0.0131	2.004	4.727
(4)	SMA	2	0.0	3	0.0015	2.101	3.364
(5)	SMA	2	0.1	3	0.0027	2.314	3.091

Table C.2: Classifier Ensemble Selected in the 2005 Backtest

	Smooth		Curving	Turning	Threshold	Standard	Mean
Classifier	Method	Param	Method	Param	Value	Error	Error
(1)	SMA	1	0.0	6	0.0009	1.000	1.000

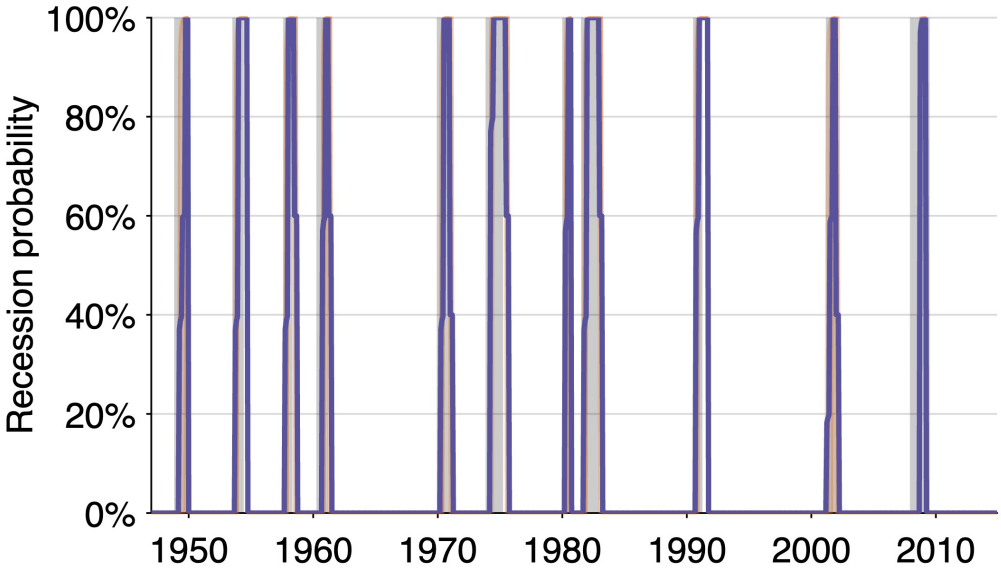


Figure C.3: In Sample Recession Probability, January 1947 - December 2014

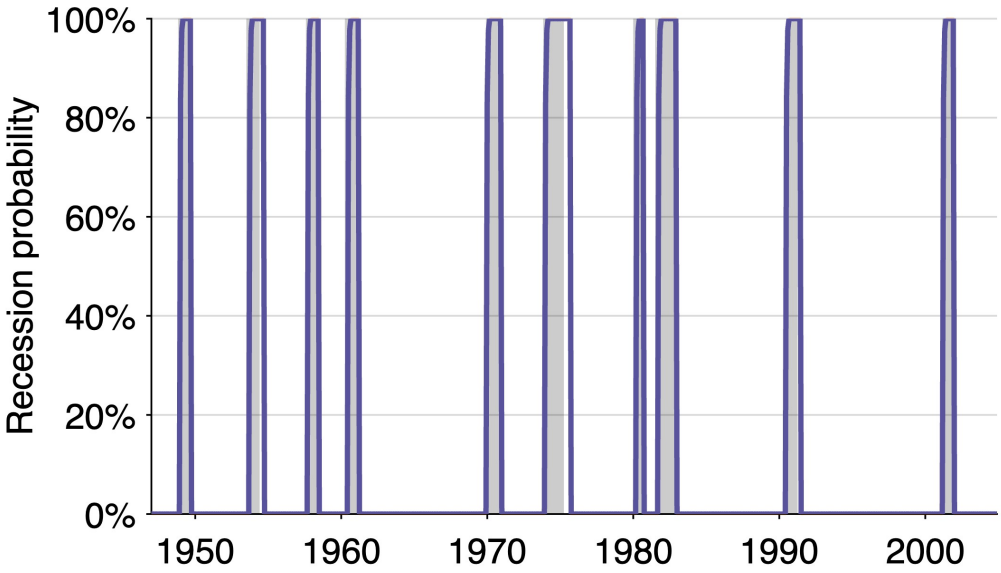


Figure C.4: In Sample Recession Probability, January 1947 - December 2004

Bibliography

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59(3):817–858.

Bai, J. and Perron, P. (1998). Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1):47–78.

Bai, J. and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1):1–22.

Barlevy, G., Faberman, R. J., Hobijn, B., and Şahin, A. (2023). The shifting reasons for beveridge-curve shifts. *NBER Working Paper 31783*.

Barnichon, R. (2010). Building a composite help-wanted index: Dataset. *Economics Letters*, 109(3):175–178.

Bell, D. N. and Blanchflower, D. G. (2019). The well-being of the overemployed and the underemployed and the rise in depression in the uk. *Journal of Economic Behavior & Organization*, 161:180–196.

Blanchflower, D. G. and Levin, A. T. (2015). Labor market slack and monetary policy. *NBER Working Paper 21094*.

Booth, A. L. and Ours, J. C. V. (2008). Job satisfaction and family happiness: The part-time work puzzle. *The Economic Journal*, 118(526):77–79.

Borgschulte, M. and Martorell, P. (2018). Paying to avoid recession: Using reenlistment to estimate the cost of unemployment. *American Economic Journal: Applied Economics*, 10(3):101–127.

Borowczyk-Martins, D. and Lalé, E. (2020). The ins and outs of involuntary part-time employment. *Labour Economics*, 67.

Brunner, A. D. and Kamin, S. B. (1996). Determinants of the 1991-1993 japanese recession: Evidence from a structural model of the japanese economy. *Japan and the World Economy*, 8(4):363–399.

Chen, Z., Iqbal, A., and Lai, H. (2011). Forecasting the probability of us recessions: A probit and dynamic factor modelling approach. *Canadian Journal of Economics/Revue canadienne d'économique*, 44(2):651–672.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annual Review of Economics*, 1:451–87.

Crump, R. K., Giannone, D., and Lucca, D. O. (2020a). Reading the tea leaves of the us business cycle—part one. Technical report, Federal Reserve Bank of New York.

Crump, R. K., Giannone, D., and Lucca, D. O. (2020b). Reading the tea leaves of the us business cycle—part two. Technical report, Federal Reserve Bank of New York.

Economic and Social Research Institute (2014). About ESRI. Accessed: 2026-05-11.

Fornaro, P. (2016). Forecasting us recessions with a large set of predictors. *Journal of Forecasting*, 35(6):477–492.

Hamilton, J. D. (2011). Calling recessions in real time. *International Journal of Forecasting*, 27(4):1006–1026.

Hashimoto, K., Ono, Y., and Schlegl, M. (2023). Structural unemployment, underemployment, and secular stagnation. *Journal of Economic Theory*, 209.

Hoshi, T. and Kashyap, A. K. (2004). Japan's financial crisis and economic stagnation. *Journal of Economic Perspectives*, 18(1):3–26.

Japan Times (2025). Japan might be in a technical recession already. *The Japan Times*. Accessed: 2026-05-11.

Kikuchi, S., Kitao, S., and Mikoshiba, M. (2021). Who suffers from the covid-19 shocks? labor market heterogeneity and welfare consequences in japan. *Journal of the Japanese and International Economies*, 59:3–26.

Landais, C., Michaillat, P., and Saez, E. (2018a). A macroeconomic approach to optimal unemployment insurance: Applications. *American Economic Journal: Economic Policy*, 10(2):182–216.

Landais, C., Michaillat, P., and Saez, E. (2018b). A macroeconomic approach to optimal unemployment insurance: Theory. *American Economic Journal: Economic Policy*, 10(2):152–181.

Liu, J., Wu, S., and Zidek, J. V. (1997). On segmented multivariate regression. *Statistica Sinica*, (7):497–525.

Mas, A. and Pallais, A. (2017). Labor supply and the value of non-work time: Experimental estimates from the field. *NBER Working Paper 23906*.

Mathers, C. D. and Shofield, D. J. (1998). The health consequences of unemployment: the evidence. *Med J Aust*, 168(4):178–182.

Michaillat, P. (2012). Do matching frictions explain unemployment? not in bad times. *American Economic Review*, 102(4):1721–1750.

Michaillat, P. (2014). A theory of countercyclical government multiplier. *American Economic Journal: Macroeconomics*, 6(1):190–217.

Michaillat, P. (2025). Recession detection using classifiers on the anticipation-precision frontier. *ArXiv*, pages 1324–1329.

Michaillat, P. and Saez, E. (2015). Aggregate demand, idle time, and unemployment. *Quarterly Journal of Economics*, 130(2):507–569.

Michaillat, P. and Saez, E. (2019). Optimal public expenditure with inefficient unemployment. *Review of Economic Studies*, 86(3):1301–1331.

Michaillat, P. and Saez, E. (2021a). Beveridgean unemployment gap. *Journal of Public Economics Plus*, 2:100009.

Michaillat, P. and Saez, E. (2021b). Resolving New Keynesian anomalies with wealth in the utility function. *Review of Economics and Statistics*, 103(2):197–215.

Michaillat, P. and Saez, E. (2022a). An economical business-cycle model. *Oxford Economic Papers*, 74(2):382–411.

Michaillat, P. and Saez, E. (2022b). $u^* = \sqrt{uv}$. arXiv:2206.13012.

Michaillat, P. and Saez, E. (2025). Has the recession started? *Oxford Bulletin of Economics and Statistics*, 87(6):1047–1058.

Miyahara, T. and Betschka, L. (2026). Can the sahm rule indicator signal recession in oecd countries? *OECD Statistics Working Papers*, 2026(01).

Moser, K., Fox, A., and Jones, D. (1984). Unemployment and mortality in the opcs longitudinal study. *The Lancet*, 324(8415):1324–1329.

National Bureau of Economic Research (2023). The NBER's business cycle dating procedure.

National Bureau of Economic Research (2024). U.s. business cycle expansions and contractions.

Ng, S. (2014). Boosting recessions. *Canadian Journal of Economics/Revue canadienne d'économique*, 47(1):1–34.

Piger, J. (2019). Turning points and classification. In *Macroeconomic Forecasting in the Era of Big Data: Theory and Practice*, pages 585–624. Springer.

Pissarides, C. A. (2009). Comments on "a new method for estimating time variation in the nairu" by william t. dickens. *Understanding Inflation and the Implications for Monetary Policy: A Phillips Curve Retrospective. MIT Press, Cambridge*, pages 235–242.

Sahm, C. (2019). Direct stimulus payments to individuals. In Boushey, H., Nunn, R., and Shambaugh, J., editors, *Recession Ready: Fiscal Policies to Stabilize the American Economy*, chapter 3. Brookings Institution, Washington, DC.

Stock, J. H. and Watson, M. W. (2014). Estimating turning points using large data sets. *Journal of Econometrics*, 178:368–381.

U.S. Bureau of Labor Statistics (2024a). Part-time workers. *U.S. Bureau of Labor Statistics*. Persons who usually work fewer than 35 hours per week are classified as part-time workers. `https://www.bls.gov/cps/definitions.htm`.

U.S. Bureau of Labor Statistics (2024b). Persons at work part time for economic reasons. *U.S. Bureau of Labor Statistics*. Persons working part time because of slack work, business conditions, or inability to find full-time work. `https://www.bls.gov/cps/definitions.htm`.

U.S. Bureau of Labor Statistics (2026). Alternative measures of labor underutilization for states. *U.S. Bureau of Labor Statistics*. Accessed June 12, 2026. `https://www.bls.gov/lau/stalt.htm`.

Valletta, R. G., Bengali, L., and van der List, C. (2020). Cyclical and market determinants of involuntary part-time employment. *Journal of Labor Economics*, 38(1):67–93.

Villena-Roldán, B. (2010). Aggregate implications of employer search and recruiting selection. Documentos de Trabajo 271, Centro de Economía Aplicada, Universidad de Chile.

Yago, K. (2020). Before the 'locomotive' runs: the impact of the 1973–1974 oil shock on japan and the international financial system. *Cambridge University Press*, pages 418–435.

Yao, Y.-C. (1988). Estimating the number of change-points via schwarz' criterion. *Statistics & Probability Letters*, 6(3):181–189.