NBER WORKING PAPER SERIES

THE DUAL U.S. LABOR MARKET UNCOVERED

Hie Joo Ahn
Bart Hobijn
Ayşegül Şahin

Working Paper 31241
http://www.nber.org/papers/w31241

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2023

This material is based upon work supported by the National Science Foundation under Grant No. SES-2048713. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, Federal Reserve Bank of Chicago, Federal Reserve Board of Governors, or of any of the other institutions the authors are affiliated with. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Dual U.S. Labor Market Uncovered
Hie Joo Ahn, Bart Hobijn, and Ayşegül Şahin
NBER Working Paper No. 31241
May 2023
JEL No. J20,J6

ABSTRACT

Aggregate U.S. labor market dynamics are well approximated by a dual labor market supplemented with a third, predominantly, home-production segment. We uncover this structure by estimating a Hidden Markov Model, a machine-learning method. The different market segments are identified through (in-)equality constraints on labor market transition probabilities. This method yields time series of stocks and flows for the three segments for 1980-2021. Workers in the primary sector, who make up around 55 percent of the population, are almost always employed and rarely experience unemployment. The secondary sector, which constitutes 14 percent of the population, absorbs most of the short-run fluctuations, both at seasonal and business cycle frequencies. Workers in this segment experience six times higher turnover rates than those in the primary tier and are ten times more likely to be unemployed than their primary counterparts. The tertiary segment consists of workers who infrequently participate in the labor market but nevertheless experience unemployment when they try to enter the labor force. Our individual-level analysis shows that observable demographic characteristics only explain a small part of the cross-individual variation in segment membership. The combination of the aggregate and individual-level evidence we provide points to dualism in the U.S. labor market being an equilibrium division of labor, under labor market imperfections, that minimizes adjustment costs in response to predictable seasonal as well as unpredictable business cycle fluctuations.

Hie Joo Ahn
Board of Governors
of the Federal Reserve System
20th & C Street, NW.
Washington, D.C. 20551
HieJoo.Ahn@frb.gov

Bart Hobijn
Federal Reserve Bank of Chicago
230 S. LaSalle St.
Chicago, IL 60604
bart.hobijn@barthobijn.net

Ayşegül Şahin
Department of Economics
University of Texas at Austin
2225 Speedway
Austin, TX 78712
and NBER
aysegul.sahin@austin.utexas.edu

1 Introduction

We show that U.S. labor market dynamics are well characterized by a Dual Labor Market (DLM) supplemented with a tertiary home-production sector that consists of those who only infrequently participate. We uncover this dual labor market structure for the U.S. by estimating a Hidden Markov Model (HMM) with inequality restrictions using histories of *all* individuals in the Current Population Survey (CPS) for 1980-2021. Our paper is the first one that adopts this novel approach for the analysis of dualism in the U.S. labor market. This stark characterization sheds light on various puzzling features of the U.S. labor market and has distinct policy implications.

The DLM Hypothesis was first posited by Doeringer and Piore (1970), who argued that a useful characterization of the U.S. labor market is that of one segmented into a *primary* and a *secondary* tier. Jobs in the primary tier generally have low turnover, pay high wages, come with benefits, offer potential for job advancement, and provide job security. Jobs in the secondary tier have high turnover, pay low wages, come with limited benefits, offer few career opportunities, and provide little job security (Piore, 1970). After a flurry of papers about dualism in the '70s and '80s,[1] the DLM Hypothesis fell into disfavor among macroeconomists during the Neoclassical Renaissance of the '80s and '90s. As early critics put it, theories of the DLM are "... too varied, incomplete, and amorphous "(Cain, 1975) to be captured in a set of microfounded first principles that explain the reasons for the endogenous emergence of discontinuous segments in the labor market (Wachter, 1974).

Recent analyses of dualism in the labor market in developed economies have mainly focused on Europe (Costain *et al.*, 2010; Bentolila *et al.*, 2019) and ignored dualism in the U.S.. This is because the institutional reasons for dualism in European labor markets, such as tiered contracts (Bentolila *et al.*, 2019), size-dependent policies (Guner *et al.*, 2008), and unionization (Berger *et al.*, 1980) are much less applicable in the U.S.. However, as the theories by Bulow and Summers (1986), Albrecht and Vroman (1992), and Saint-Paul (1997) point out, dualism can emerge even in the absence of such institutional arrangements and structures. For example, as a result of frictions, the existence of efficiency wages and, more generally, due to the nature of demand fluctuations in different segments of the economy

While the regulatory and institutional differences between segments in European labor markets allow for a clear identification of which workers are in the primary and secondary tiers, the absence of such differences in the U.S. makes this identification nontrivial.[2] Moreover, the

[1]See, for example Reich *et al.* (1973), Harrison and Sum (1979), Berger *et al.* (1980), and Dickens and Lang (1985).

[2]Some authors have used occupation as a proxy for the labor market segment workers are in (e.g. McNabb

original DLM hypothesis focuses specifically on those participating in the labor market. Since these insightful studies in the early 1970s, the U.S. labor market experienced several important changes in demographics and labor supply behavior which amplified the importance of the participation margin for labor market trends and fluctuations.[3] Motivated by the rising importance of the participation margin, we augment the DLM hypothesis with a tertiary home production sector when bringing it to the data.

We apply an unsupervised machine learning method, that involves estimating an HMM, to identify which respondents in the CPS are part of the primary, secondary, and tertiary sectors over the period 1980-2021. Each of these segments themselves consist of four hidden states: employed, short-term unemployed, long-term unemployed, and non-participants. Based on more than 10 million individual labor market histories from 1980 to 2021, we identify these twelve distinct labor market states and their dynamics within the three labor market segments. We also estimate the probabilities of belonging to each of the three labor market segments for each respondent in the CPS.

Our analysis builds on a small, but growing, literature that aims to identify a limited set of worker types to capture the relevant aspects of macro heterogeneity in the U.S. labor market. (Hall and Kudlyak, 2019; Gregory *et al.*, 2021; Shibata, 2019). Our method differs from those used in these papers in four important ways. First, our hidden states have a direct economic interpretation that stems from the identifying restrictions we impose that are based on the DLM Hypothesis.[4] Second, in contrast to the models of Hall and Kudlyak (2019), Gregory *et al.* (2021) and Shibata (2019), we estimate monthly *time series* of the stocks and flows for each of the hidden states. Therefore, we can analyze seasonality, business cycle properties, and long-run trends in the three labor market segments we identify. Third, for our identification we use detailed labor force status data in the CPS to inform our hidden states.[5] Fourth, the method yields individual-level results that aggregate to the monthly labor market stocks and flows published by the Bureau of Labor Statistics (BLS).

and Psacharopoulos, 1981).

[3]The complex interaction of trend and cyclical factors after the Great Recession required policymakers to make *"... difficult judgments about the magnitudes of the cyclical and structural influences affecting labor market variables, including labor force participation"* (Yellen, 2014). The sudden and drastic drop in participation at the onset of the pandemic in 2020 has made these judgments even more important in the wake of the COVID-19 Recession. It has led policymakers to consider the unemployment rate corrected for changes in labor force participation as a measure of labor market slack (Powell, 2021).

[4]A useful analogy is the structural VAR literature. The structural VAR approach helps recover the structural estimate of parameters of interest by imposing restrictions to a reduced-form model that are informed by economic theory.

[5]While detailed labor force status data have been exploited to assess aggregate labor market conditions, we are the first to systematically introduce refined labor force states, such as part-time for economic reasons, discouraged or marginally attached, into a unified statistical framework used to identify underlying heterogeneity in the labor market.

The aggregate results show that the U.S. labor market is well characterized by three distinct tiers. Workers, in the *primary* segment, who make up around 55 percent of the population, are almost always employed and they very rarely experience unemployment. They also seamlessly move from non-participation to employment unlike workers in the secondary and tertiary sectors. Labor market frictions are basically irrelevant for these *primary* sector workers. The *secondary* sector, which constitutes 14 percent of the population, exhibits high turnover and high unemployment and absorbs most of the short-run fluctuations in the labor market, at both seasonal and business cycle frequencies. Workers in this sector are six times more likely to move between labor market states than those in the primary tier and are ten times more likely to be unemployed than their primary counterparts. The *tertiary* sector includes workers who are only loosely attached to the labor force and has a very low employment-to-population ratio. These workers tend to experience unemployment when they enter the labor force from non-participation but do not share the high job-loss rate of secondary workers. These large differences between the three tiers of the labor market imply that average stocks and flow rates, which are commonly used to quantitatively discipline macro-labor models, are not at all reflective of individual labor market experiences and outcomes.

The stark contrast between the three segments means that each segment contributes to different aspects of aggregate outcomes. The primary sector accounts for more than 80% of employment and participation but its contribution to the unemployment rate is much smaller. Only about a quarter of aggregate unemployment is due to the incidence of unemployment in the primary market. What is probably the most striking finding is that the secondary market, which makes up only 11.9% of total employment, accounts for 61% of unemployment in the economy and almost two thirds of unemployment fluctuations over the business cycle. Our findings also relate to the puzzling nature of the Beveridge curve in the U.S. (see for example, Daly *et al.* (2012); Elsby *et al.* (2015a)). We find that shifts in the Beveridge curve are also mostly accounted for by the secondary sector. The primary sector Beveridge curve exhibits a tight and mostly stable relationship between unemployment and vacancies. Labor market dynamism, for which we use a new measure of flows per capita, is also highly uneven, with the secondary market accounting for half of the turnover in the economy. Moreover, two of the most notable long-run labor market trends in the U.S., i.e. the trend decline in the unemployment rate and the decline in labor market dynamism, are mostly accounted by changes in the secondary sector.

We combine our estimates of individual-level posterior probabilities of being in each segment for each of the 10 million individuals in the CPS from 1980-2021 with other variables in the CPS that are not used in our estimation. This allows us to examine the potential reasons for dualism in the U.S. labor market. We find some evidence consistent with life-cycle effects as well as discrimination. However, observable demographic characteristics only explain a small part of the cross-individual variation in segment membership. Moreover, their significance has been declining over time. Consistent with the efficiency wage theories of dualism, analyzed in Bulow and Summers (1986), Albrecht and Vroman (1992), and Saint-Paul (1997), jobs in the primary sector are for high-skilled service occupations for which output is hard to monitor, are more stable, pay higher wages, and have higher returns to schooling and experience. The combination of the aggregate and individual-level evidence we provide points to dualism in the U.S. labor market being an equilibrium division of labor, under labor market imperfections, that minimizes adjustment costs in response to predictable seasonal as well as unpredictable business cycle fluctuations.

The rest of this paper is structured as follows. In Section 2 we discuss the details of our methodology in the context of the literature and explain how the DLM provides a way to consider many dimensions of micro and macro heterogeneity at the same time. In section 3, we describe how we distinguish the primary, secondary, and tertiary markets in the context of an HMM and how we resolve the practical challenge of estimating the model with many parameters and observations subject to the identifying restrictions we impose. We present our results in two parts. In Section 4 we show how the primary, secondary, and tertiary tiers are very different from each other as well as from the aggregate labor market. We quantify the importance of each of the three segments for the trends and cycles in commonly analyzed aggregates in Section 5. We present robustness and model comparisons in Section 6. In Section 7 we analyze the individual-level evidence and discuss what it implies about the possible causes of labor-market dualism in the U.S..

2 Importance and Identification of Macro Heterogeneity

The division of the population into the labor market states of employed, unemployed, and non-participants is the common classification system used to analyze macroeconomic outcomes in the labor market, including in the CPS. While these categories capture very important differences in workers' labor market experiences, they are too coarse to characterize many different aspects of individual and aggregate labor market outcomes.

A growing recent literature has emphasized the importance of different subcategories of persons within the three labor market states for individual and aggregate outcomes. These include heterogeneity among the unemployed that accounts for the duration distribution of unemployment (van den Berg and van Ours, 1996; Hornstein, 2012; Ahn and Hamilton, 2020a;

Kroft *et al.*, 2016; Mueller and Spinnewijn, 2023), heterogeneity in the type of jobs for the employed to account for the tenure distribution (Hall, 1982; Hyatt and Spletzer, 2016) as well as worker turnover (Pries, 2004; Pries and Rogerson, 2021), and heterogeneity among different categories of non-participants and unemployed to account for fluctuations in matching efficiency (Hall and Schulhofer-Wohl, 2018; Sedláček, 2016; Abraham *et al.*, 2020). All these studies have the common implication that a more accurate description of individual-level labor market histories as well as macro-level labor market dynamics requires the identification and measurement of broad subcategories of the three coarse labor market states. We refer to these subcategories as "Macro Heterogeneity."

Relatedly, commonly used Diamond-Mortensen-Pissarides search and matching framework (e.g. Pissarides, 1985; Mortensen and Pissarides, 1994) imply that flows between employment and unemployment are Markovian in that multi-period transition probabilities are compounded one-period transition probabilities, where the latter are calibrated from the data. As Kudlyak and Lange (2017) and Morchio (2020) point out, this is neither the case for employment-unemployment flows in the data nor for flows across the participation margin. As a result, such models do not fit individual multi-period transition probabilities between labor market states. One approach to fit these individual histories is to represent them as a mixture of different unobserved first-order Markov processes. Such a mixture approach is not only useful to match individual-level evidence, Ferraro (2018) and Gregory *et al.* (2021) show that mixtures of commonly-used models help us better understand the sources of persistence and asymmetries in aggregate labor market dynamics.[6]

These insights provide the following research challenge for the identification of Macro Heterogeneity: Develop a method to find a parsimonious representation of individual and aggregate labor market dynamics in terms of a mixture of a limited number of hidden first-order Markov processes that each have a clear economic interpretation. This method, by definition, involves classifying individuals at each point in time into untagged hidden labor market states. Because it does not use prior information about who belongs to which group, it is a form of *unsupervised* machine learning.

The method we use to tackle this challenge is a Hidden Markov Model (HMM), which is a statistical tool that estimates latent states and their dynamics from data on categorical sequences. We use this method because it has four important advantages over earlier studies of

[6]The crucial insight is that a higher-order Markov process can be characterized as a mixture of first-order Markov processes. See Granger and Morris (1976) for an example of this for ARMA processes. This insight has been applied by Ferraro (2018) and Gregory *et al.* (2021). Ferraro (2018) shows that the dynamics of a mixture of Mortensen and Pissarides (1994) models can have very rich dynamics. Gregory *et al.* (2021) show the same for a mixture of Menzio and Shi (2011) models.

Macro Heterogeneity in the U.S. labor market (Hall and Kudlyak, 2019; Shibata, 2019; Gregory *et al.*, 2021).

The first is that it allows us to impose specific identifying restrictions, guided by the Dual Labor Market Hypothesis posited in Doeringer and Piore (1970), across the hidden states we identify. These restrictions are (in-)equality constraints on the persistence of and turbulence between labor force states. They assure us that the hidden states are interpretable as the primary, secondary, and tertiary sectors of the labor market.

The implementation of an HMM with inequality restrictions is appealing since it mimics the use of similar restrictions for the identification of economically meaningful shocks in Structural Vector-Autoregression (SVAR) models.[7] However, it has not yet been applied in labor economics since its implementation is numerically challenging.

We estimate the HMM with Maximum Likelihood via the Expectation-Maximization (EM) algorithm (Dempster *et al.*, 1977). Our main methodological contribution is to show that this can be done through a generalization of the Baum-Welch (BW) algorithm, introduced by Baum *et al.* (1970) and Welch (2003), that is commonly used for the estimation of HMMs. The difference is that in our method the M-step involves the numerical maximization of the expectation of the complete-data likelihood function subject to the identifying (in-)equality restrictions we impose. We show that this is feasible because it can be split up into a set of well-behaved convex maximization problems for which efficient numerical methods are available.

The second is that our method allows us to estimate monthly time-varying stocks of and transition probabilities between the hidden states. This yields a set of stocks and flows for the hidden states that is conceptually identical to those published monthly by the BLS for the labor force states of employment, unemployment, and non-participation and studied extensively in many papers (e.g Marston, 1976; Blanchard and Diamond, 1990; Shimer, 2012; Barnichon and Nekarda, 2012; Elsby *et al.*, 2015b). This allows us to analyze the importance of the segmentation of the labor market for seasonality, business cycle fluctuations, and long-run trends.

The third advantage is that our HMM allows us to use an extensive set of twenty nine nuanced answers in the CPS about the types of and reasons for employment, unemployment, and non-participation that respondents provide. These include part-time versus full-time employment, the reasons for unemployment, as well as the intent to and reason for not looking for a job when not participating, among others. We use the variation in labor market outcomes between different groups of individuals like those part-time employed for economic reasons versus full-time employed, those temporarily unemployed versus ones having been laid off, or marginally

[7] For example, Stock and Watson (2001), Christiano *et al.* (2006), and Baumeister and Hamilton (2015).

attached non-participants versus those not wanting a job, to enhance our assessment of the likelihood of which of the three labor market segments they are part of.

The final advantage is that, based on the reported labor market histories in the CPS, our method provides estimates of the posterior probability that a respondent is part of each of the three respective labor market segments. To make sure these individual-level estimates aggregate to the three-state stocks and flows published by the BLS and analyzed in other studies, we assume, just like in the published data, that missing observations are random and that workers do not make any classification errors when they report whether they are employed, unemployed, or not participating in the labor market.[8] The individual-level posterior probabilities are additional variables for all CPS respondents that can be used to assess both the incidence of segment membership by demographic group as well as the impact of segment membership on labor market outcomes, like industry and occupation of employment, earnings, as well as hours worked and tenure.

3 A Dual Labor Market in an HMM

In this section we describe the structure of the HMM we estimate. We then discuss the restrictions we impose to distinguish the three market segments. We explain how we estimate the model and how we obtain posterior probabilities of each CPS respondent's segment membership. Finally, we discuss the identification of the parameters in the model.

3.1 Structure of the Dual Labor Market HMM

The structure of the HMM we estimate is guided by both the aim to estimate the stocks and flows in the segments of the DLM as well as by the specific structure of the CPS data we use for that purpose.[9] We focus on our benchmark specification, which is known as a Non-Homogenous Hidden Markov Model (NHMM). While it has been applied in other fields,[10] the application to U.S. labor market data is new to our paper.

Our specification consists of three labor market tiers: A primary (P), secondary (S), and tertiary (T). Each of these segments themselves consist of four hidden states: employed (EM), short-term unemployed (UMS), long-term unemployed (UML), and non-participants (NM),

[8] There is an extensive literature on such classification errors (Abowd and Zellner, 1985; Blanchard and Diamond, 1990; Feng and Hu, 2013; Elsby *et al.*, 2015b; Ahn and Hamilton, 2020b) and a large degree of disagreement about their importance.

[9] The specific version of the data we use are from Flood *et al.* (2020)

[10] For example, for the analysis of rainfall patterns (Hughes *et al.*, 1999).

Table 1: Hidden states in HMM

State	Description
EP	Primary employed
ES	Secondary employed
ET	Tertiary employed
UPS	Primary short-term unemployed
UPL	Primary long-term unemployed
USS	Secondary short-term unemployed
USL	Secondary long-term unemployed
UTS	Tertiary short-term unemployed
UTL	Tertiary long-term unemployed
NP	Primary non-participant
NS	Secondary non-participant
NT	Tertiary non-participant

where $M \in \{P, S, T\}$ denotes the market segment. The resulting twelve hidden states are listed in Table 1.

Our goal is to classify persons, who are categorized as either employed, unemployment, or not-in-the-labor-force, into a set of refined hidden states based on their responses to the CPS about their labor market status. In the context of the HMM, these responses are called *emissions*, because they are observable signals that respondents "send" about the hidden state they are in.

The HMM consists of two layers. The first is the stochastic process that drives the evolution of the hidden state for each individual that aggregates to the flows and stocks in the labor market. We denote the hidden labor market state of individual i by $\ell_{i,t} \in L$, where L is the set of twelve hidden labor market states. It follows a first-order Markov process in that the transition probabilities satisfy

$$q_{l,l',t} = P\left(\ell_{i,t} = l' \mid \ell_{i,t-1} = l; t\right) = P\left(\ell_{i,t} = l' \mid \ell_{i,t-1} = l, \cap_{k=2}^{\infty} \ell_{i,t-k} = l_{t-k}; t\right), \tag{1}$$

where $(l, l') \in L \times L$. The argument $t = 1, \ldots, T$ reflects that they vary over time.[11] These transition probabilities are the *flow* rates between the different hidden states in our model. These flow rates determine the evolution of the stocks of individuals in the each hidden state. These stocks are the unconditional probabilities of an individual being in state $l \in L$ in month t. We denote them by

$$\delta_{l,t} = P\left(\ell_{i,t} = l; t\right). \tag{2}$$

[11] Earlier applications of HMMs in empirical studies of labor market data, such as Boeschoten *et al.* (2019) and Shibata (2019), do not allow for time variation in transitions probabilities and do not consider (in-)equality constraints as identifying assumptions.

The advantage of the assumption that the hidden states follow a first-order Markov process is that this makes the hidden states interpretable as states in a generalized theoretical model of the labor market in which transitions between the states follow a first-order Markov process, as in the seminal model by Mortensen and Pissarides (1994). At first glance, this assumption might seem restrictive. However, because there are more hidden states than the three observed categories of employment, unemployment, and non-participation, the observed categories are a mixture of the underlying hidden states and mixtures of first-order Markov processes can have a wide range of non-Markovian properties.

The second layer of the HMM is the stochastic process that determines the information the emissions provide about the hidden state that an individual is in. We denote the emission of individual $i = 1, \ldots, n$ in month t by $x_{i,t} \in X$, where X is the set of possible emissions that we discuss in more detail below. The relationship between the emissions and the hidden states is known as the emission model.

The main assumption behind the emission model in an HMM is that the probability of a particular emission only depends on the current hidden state. This conditional-independence assumption yields the following expression for the emission probability

$$\omega_{x,l,t} = P\left(x_{i,t} = x \mid \ell_{i,t} = l; t\right), \text{ where } x \in X \text{ and } l \in L. \tag{3}$$

Here, the argument t captures that the emission probabilities in our model vary over time.

We include in the set of emissions, X, information about the labor force status, i.e. employed, unemployed, or non-participant, the type of employment, the reason for unemployment, the duration of unemployment, whether or not non-participants completed a seasonal or temporary job, and information about labor-force attachment. This results in 29 different possible emissions, listed in Table 2.

The emissions distinguish between unemployed of different durations. This might seem like a violation of the conditional-independence assumption because to report having been unemployed for several months seems to imply that one was unemployed in the previous month. This, however, is not the case. Unemployed respondents in the CPS report how long they have been searching for a job rather than the duration of their unemployment spell. Many respondents in the survey report to be employed or out of the labor force during the period for which they later report to have been searching for a job (Elsby *et al.*, 2011).

To summarize, we have a panel of incomplete observed 16-month long labor market histories across individuals that sends an imperfect signal about in which of the 12 hidden labor market states they are in at each point in time. We use the HMM described above to estimate the

Table 2: Observed emissions in HMM

Emission	Description
M	Labor market state not reported in the CPS
EX	Employed, no other detail
EPE	Employed, part-time for economic reason
ENW	Employed, absent for other reasons
UTL5	Unemployed on temporary layoff, duration < 5w
UTL14	Unemployed on temporary layoff, duration < 14w
UTL26	Unemployed on temporary layoff, duration < 26w
UTLLT	Unemployed on temporary layoff, duration > 26w
UTJ5	Unemployed temporary job ended, duration < 5w
UTJ14	Unemployed temporary job ended, duration < 14w
UTJ26	Unemployed temporary job ended, duration < 26w
UTJLT	Unemployed temporary job ended, duration > 26w
UJL5	Unemployed job loser, duration < 5w
UJL14	Unemployed job loser, duration < 14w
UJL26	Unemployed job loser, duration < 26w
UJLLT	Unemployed job loser, duration > 26w
UX5	Unemployed n.e.c., duration < 5w
UX14	Unemployed n.e.c., duration < 14w
UX26	Unemployed n.e.c., duration < 26w
UXLT	Unemployed n.e.c., duration > 26w
NTJDW	Non-participant who ended temporary job and discouraged worker
NTJMA	Non-participant who ended temporary job, not discouraged but marginally attached
NTJNA	Non-participant who ended temporary job, recently searched but not available for work
NTJNS	Non-participant who ended temporary job, no previous job search but want a job
NTJDNW	Non-participant who ended temporary job, does not want a job
NDW	Non-participant and discouraged worker
NMA	Non-participant, not discouraged but marginally attached
NNA	Non-participant, recently searched but not available for work
NNS	Non-participant, no previous job search but want a job
NDNW	Non-participant, does not want a job

following time series: (i) the share of individuals in each of the states, $\delta_{j,t}$, i.e. the equivalent of the stocks, (ii) the transition probabilities between the latent states $q_{l',l,t}$, i.e. the flow rates, and (iii) the emission probabilities, $\omega_{x,l,t}$. We denote the vector with all these parameters as $\boldsymbol{\theta}$, the vector with the observed history of emissions for individual i as $\boldsymbol{x}_i$, and the vector with the unobserved path of underlying hidden states as $\boldsymbol{\ell}_i$.

3.2 Distinguishing the Three Labor Market Segments

The economic usefulness of methods that identify latent states, like our hidden Markov states, depends on their interpretability. The goal of our analysis is for the hidden states of our model to correspond to the primary and secondary segments from the DLM theory and for the third sector to capture those less attached to the labor force. In our benchmark model these three segments are distinguished by inequality constraints and zero restrictions on the monthly transition probabilities between employment, unemployment, and non-participation.

Inequality restrictions on transition probabilities

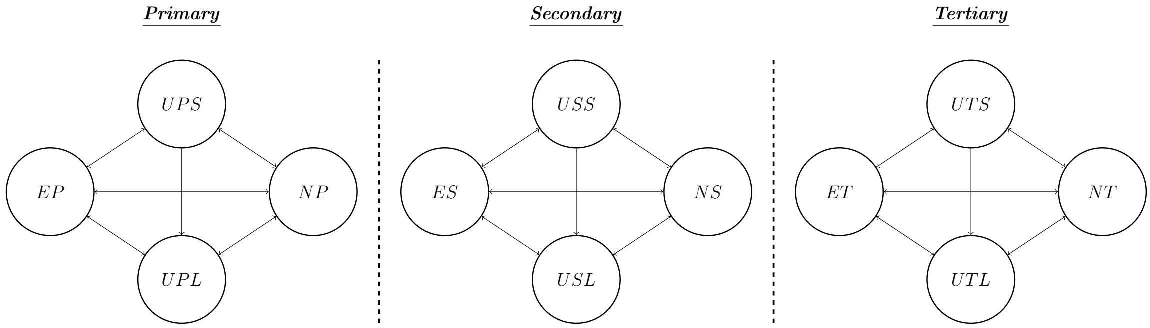
The first type of restrictions captures the differences in relative turnover rates between labor market segments from the DLM Hypothesis.

The hypothesis is that the primary tier of the labor market is characterized by a higher level of employment stability than the secondary and tertiary tiers. Employment stability is a key attribute that distinguishes the primary market from the secondary market (Doeringer and Piore, 1970; Piore, 1970; Berger *et al.*, 1980; Dickens and Lang, 1985). According to Wachter (1974), one of the hypotheses defining the dual labor market is that workers in the secondary sector experience a pattern of job instability. Similarly, Bentolila *et al.* (2019) mention that the main feature of dual labor market in Europe is the coexistence of open-ended and fixed-term contracts. The former guarantee job security, while the latter make a job last only for a short period of time. The inequality restriction we impose captures the gist of this aspect of the DLM theory. To have our parameter estimates satisfy this property, we impose the restrictions that

$$q_{EP,EP,t} \geq q_{ES,ES,t} + 0.05 \text{ and } q_{EP,EP,t} \geq q_{ET,ET,t} + 0.05, \text{ for all } t. \tag{4}$$

The main focus of the existing literature on dual labor markets are persons who are in the labor force and are either employed or unemployed. But the data from the CPS that we use cover the whole population, among which there is substantial heterogeneity in labor force attachment and labor supply elasticities (Krusell *et al.*, 2017; Mui and Schoefer, forthcoming).

Figure 1: Description of the three labor market segments.



In particular, there is a large group of people who are only very loosely attached to the labor force. It is this group of people who are part of the tertiary market. We capture their loose attachment to the labor market imposing additional inequality restrictions on the persistence of non-participation in the tertiary market segment. These restrictions take the form

$$q_{NT,NT,t} \geq q_{NP,NP,t} + 0.05 \text{ and } q_{NT,NT,t} \geq q_{NS,NS,t} + 0.05, \text{ for all } t. \tag{5}$$

In addition, our model specification includes more than one hidden type of unemployment in each market. To assure that the interpretation of the hidden unemployment states we uncover matches their labels, we assume that long-term unemployment is more persistent than short-term unemployment. That is

$$q_{UML,UML,t} \geq q_{UMS,UMS,t} + 0.05 \text{ where } M \in \{P, S, T\}, \text{ for all } t, \tag{6}$$

and also impose that persons can only flow from short- to long-term unemployment and not vice-versa, i.e.

$$q_{UML,UMS,t} = 0 \text{ where } M \in \{P, S, T\}, \text{ for all } t. \tag{7}$$

The four constraints above assure us that each of the hidden states we identify has a clear economic interpretation in the context of the DLM with a tertiary home production sector.

Zero restrictions on transition probabilities

The second type of restrictions is guided by another assumption in the DLM Hypothesis. Namely, that there is very limited mobility between labor market segments. Because of the very short histories reported in the CPS we approximate this assumption by the restriction that respondents do not switch market tiers during the 16-month period they are in the sample. This translates into a set of zero restrictions on the transition probabilities that capture that there are no flows between the primary, secondary, and tertiary markets. Figure 1 summarizes the market structure we estimate using our HMM.

3.3 Restrictions on Emission Probabilities

A second set of restrictions that we impose is related to the connection between the emissions and underlying hidden states. The restrictions are guided by our goal to uncover the stocks and flows in the segments of the DLM with a tertiary home production sector that are consistent with aggregate stocks and flows published by the BLS.

No classification errors: Zero restrictions on emission probabilities

To assure our estimates align with published statistics, we assume that there are no classification errors. That is, we impose that respondents correctly report their labor market status of employment, unemployment, and non-participation. In that case, the probability that their emission does not correspond to their hidden labor market status is zero. We impose these zero restrictions on the emission probabilities for all months in our sample.

Random missing values

Consistent with the methodology the BLS constructs the published statistics, we impose that missing values for the emissions, $x_{i,t}$, are random. That is, the probability that a respondent does not report any emissions does not depend on the hidden state they are in. This way of treating the missing values means that no information is gleaned from whether an observation is missing or not.[12] This assumption is, by definition, true for the 8-month reporting gap in the CPS during which respondents drop out of the sample and we assume it also holds for the eight months they are in the sample.

3.4 Estimation

We estimate our model for all respondents, $i = 1, \ldots, n$, in the CPS from 1980-2021. The resulting sample size is $n = 10,178,593$ individual 4-8-4 labor market histories. The total number of parameters is 84,168, which is 167 for each of the 504 months in the sample.[13]

[12] Alternatively, one can treat missing observations as being in a fourth observable state and include it in the model through the emission probabilities. This is how Ahn and Hamilton (2020b) treat missing values in their analysis of measurement error in the CPS without using an explicit HMM.

[13] These 167 parameters are: 33 transition probabilities, 123 emission probabilities, and 11 shares of the population in the hidden states.

Because the estimation involves a large number of observations, n, and parameters, $\dim(\boldsymbol{\theta})$, direct maximization of the likelihood function is not feasible. However, it can be accomplished through the application of the BW algorithm (Baum *et al.*, 1970; Welch, 2003), commonly used in machine learning and estimation of HMMs. This is a specific case of the EM algorithm (Dempster *et al.*, 1977). The particular form of the algorithm we use exploits the panel data structure (Maruotti, 2011) of the CPS and takes into account the identifying (in-)equality restrictions on the parameters.

The likelihood function, $\mathcal{L}(\boldsymbol{\theta})$, we maximize is the joint probability of observing the paths, $\{\boldsymbol{x}_i\}_{i=1}^{n}$, for a given vector of model parameters

$$\mathcal{L}(\boldsymbol{\theta}) = \prod_{i=1}^{n} P(\boldsymbol{x}_i;\boldsymbol{\theta})^{w_i} = \prod_{i=1}^{n}\left[\sum_{\ell_{i,t_i+15}\in L} P(\boldsymbol{x}_i \cap \ell_{i,t_i+15};\boldsymbol{\theta})\right]^{w_i} = \prod_{i=1}^{n}\left[\sum_{\ell\in L}\alpha_{i,15}(\ell;\boldsymbol{\theta})\right]^{w_i} \quad (8)$$

Here w_i is the sample weight for individual i.[14]

$$\alpha_{i,k}(\ell;\boldsymbol{\theta}) = P(x_{i,t_i},\ldots,x_{i,t_i+k}\cap \ell_{i,t_i+k}=\ell). \quad (9)$$

It is the joint probability of the observed data from t_i through t_i+k *and* individual i being in the latent state $\ell \in L$ at $t = t_i + k$.

In principle, the computation of $\alpha_{i,k}(\ell;\boldsymbol{\theta})$ requires the summation over all possible paths of the latent state between t_i and t_i+k, which quickly becomes infeasible. However, the BW algorithm uses that $\alpha_{i,k}(\ell;\boldsymbol{\theta})$ can be calculated using a forward recursion. For the specific case of the CPS data with missing values, this recursion is of the form

$$\alpha_{i,0}(l) = \delta_{l,t}\left((1-\eta_{i,t_i}) + \eta_{i,t_i}\omega_{x_{i,t_i},l,t_i}\right), \text{ and} \quad (10)$$

$$\alpha_{i,k}(l') = \sum_{l\in L}\alpha_{i,k-1}(l)\, q_{l,l',t_i+k}\left((1-\eta_{i,t_i+k}) + \eta_{i,t_i+k}\omega_{x_{i,t_i+k},l',t_i+k}\right) \quad (11)$$

where $\eta_{i,t}$ is the indicator function for non-missing observations which ensures that missing observations are integrated out of the fitted path, consistent with the assumption that they are random.

[14] Because an individual appears in the likelihood for her/his whole 16 periods labor market history, no matter whether observations are missing or not, w_i is, in principle, the sampling weight of individuals conditional on them reporting their labor market state for at least one out of eight interviews. However, such a weight is not provided for the CPS data. Therefore, we approximate it by their average cross-sectional weight across all the 8 months in sample. That is, w_i is the average number of persons the individual represents across the 8 rotations in which they are interviewed.

As with any application of the EM algorithm, it involves iteratively updating the parameters to monotonically increase the likelihood function. Each iteration involves two steps. An E-step and an M-step. These steps for the estimation of a panel-data HMM have been described in Maruotti (2011) and Shibata (2019). For this reason, we leave the details for Appendix A and focus on two specific aspects we use in the rest of our analysis: (*i*) how the E-step provides estimates of the posterior probabilities that each of the respondents in the CPS is in a particular segment of the labor market, and (*ii*) how we implement the identifying (in-)equality restrictions on the parameters in the M-step.

The starting point for the EM algorithm is the complete-data log-likelihood function, which is the log of the likelihood function for the case in which all data, i.e. $\{\boldsymbol{x}_i, \boldsymbol{\ell}_i\}_{i=1}^n$, are observed. If we had data on the hidden state, we could construct the dummy variables

$$u_{i,t,l} = \mathbb{1}\left(\ell_{i,t} = l\right) \text{ and } v_{i,t,l,l'} = \mathbb{1}\left(\ell_{i,t-1} = l \cap \ell_{i,t} = l'\right). \tag{12}$$

Given these indicator functions, the complete-data log-likelihood function equals

$$\begin{aligned} \ln \mathcal{L} \;=\; & \sum_{i=1}^{n} w_i \left\{ \sum_{l \in L} u_{i,t_i,l} \ln \delta_{l,t_i} + \sum_{k=1}^{15} \sum_{l' \in L} \sum_{l \in L} v_{i,t_i+k,l,l'} \ln q_{t_i+k,l,l'} \right. \\ & \left. + \sum_{k=0}^{15} \eta_{i,t_i+k} \sum_{l \in L} u_{i,t_i+k,l} \ln \omega_{x_{i,t_i+k},l,t_i+k} \right\}. \end{aligned} \tag{13}$$

Individual-level posterior probabilities from E-step

In the E-step, the expectation of the complete-data log-likelihood conditional on the observed data $\boldsymbol{x} = \{\boldsymbol{x}_i\}_{i=1}^n$ and parameter vector $\boldsymbol{\theta}$ is calculated.

Taking the conditional expectation of (13) involves replacing $u_{i,t_i+k,l}$ and $v_{i,t_i+k,l,l'}$ with their conditional expectations, which we denote by $\hat{u}_{i,t_i+k,l}$ and $\hat{v}_{i,t_i+k,l,l'}$ respectively. They are calculated using the Forward-Backward recursions, part of the BW algorithm, described in Appendix A.

For our analysis it is important to realize that these conditional expectations are not only useful for the implementation of the BW algorithm. They also allow us to do individual-level analyses of our results. The reason is that $\hat{u}_{i,t_i+k,l}$ can be interpreted as the posterior probability that a person is in a particular hidden state at time $t_i + k$, i.e.

$$\hat{u}_{i,t_i+k,l} = E\left[\mathbb{1}\left(\ell_{i,t_i+k} = l\right) \mid \boldsymbol{x}_i, \boldsymbol{\theta}\right] = P\left(\ell_{i,t_i+k} = l \mid \boldsymbol{x}_i, \boldsymbol{\theta}\right) \text{ for } l \in L. \tag{14}$$

Therefore, the BW algorithm does not only yield a set of parameter estimates but also provides posterior probabilities of the stocks for each of the individuals in the data for these estimates. Note that the algorithm does not classify individuals in a particular hidden state at each point in time. Instead, their classification is a probabilistic assessment based on the limited information revealed by a person's labor market history from the 4-8-4 survey structure of the CPS.

Our focus, in particular, is on the posterior probability that a respondent is part of one of the three market segments. This probability is given by

$$P_i(M) = \sum_{l \in \{EM, UMS, UML, NM\}} P(\ell_{i,t} = l \mid \boldsymbol{x}_i, \boldsymbol{\theta}), \text{ where } M \in \{P, S, T\}. \tag{15}$$

Because we impose the restriction that individuals cannot flow from one market segment to another, this probability is constant over time.[15]

Our estimation procedure thus yields two additional variables for each respondent in the CPS that reflect the posterior probabilities that she or he is part of the primary or secondary segment of the labor market.[16]

Imposing identifying zero- and inequality restrictions in M-step

The use of zero- and inequality constraints on the transition and emission probabilities is at the heart of our identification strategy to provide specific economic meaning to the hidden states we uncover. In the M-step the expectation of the complete-data likelihood function is maximized with respect to the parameters subject to these restrictions. In the absence of these restrictions, the M-step yields a well-known closed-form solution that is easy to solve, even in the case of a very large number of parameters (e.g. Maruotti, 2011). However, this is not the case under the constraints that we impose. Zero restrictions on the transition and emission probabilities are easily imposed in the maximization problem. The challenge is how to deal with the inequality constraints, especially in light of the large number of parameters we estimate.

One approach of dealing with inequality constraints in the BW algorithm is to transform the problem to one that has a closed-form solution (e.g. Levinson *et al.*, 1983; Otterpohl, 2002). This, however, is not feasible for the large number of parameters and restrictions in our model specification. Instead, we use that the maximization problem in the M-step can be split up into $3T$ sub-problems. Each of these involves the calculation of a Weighted Analytic Center and can be easily solved using the numerical method introduced in Andersen *et al.* (2011).[17] Even though our specification has a large number of parameters and constraints, we are able

[15] See Appendix A for a proof.

[16] The probability that the respondent is in the tertiary market is implied by the first two by the constraint that the probabilities add up to one.

[17] We discuss the details of this approach in Appendix A.

to impose the identifying restrictions in the M-step by reframing the maximization problem as a set of much smaller, well behaved, maximization problems for particular subsets of the parameters.

3.5 Identification and Reliability of Classification

While similar NHMM's have been applied and estimated in other fields (Hughes *et al.*, 1999), sufficient conditions for identification have only been established for the homogeneous case with time-invariant transition and emission probabilities.[18]

Intuitively, identification in our model can be thought of in the context of the method of moments. In each month, we have 167 parameters; monthly transitions between these emissions imply 812 transition probabilities, which can be interpreted as empirical moments.[19] Thus, on a period-by-period basis we have many more empirical moments than parameters. Applications of NHMM's of the type we analyze use bootstrapping methods to quantify the reliability of the estimates as well as establish local identification of the parameters. We follow that practice in the rest of our analysis when we present our results. Our bootstrap consists of 1000 draws from the model at the estimated parameters.

What is central to our method is the classification of individuals into different labor-market segments, which hinges on the information content of the emissions that we use, which is formalized by Petrie (1969). The main insight from Petrie (1969) is that if those in different hidden states are equally likely to report the same emissions, then the emissions would not help to classify individuals in different hidden states. Formally, this requires that, in any month t, the 29×12 matrix with emission probabilities $\omega_{x,l,t}$'s has full column rank. This is indeed the case for our parameter estimates.

In line with the results in Petrie (1969), what is crucial for the classification of individuals is the fact that our estimates imply very different probabilities of people reporting emissions depending on what market segment they are in. This can be seen from Table 3. It reports the average emission probabilities, $\bar{\omega}_{x,l}$, over our sample period. For example, on average, only 1 percent of those employed in the primary sector report to be part-time for economic reasons (element (EPE,EP) in the table) while about a third of those employed in the secondary segment

[18] A simple generalization of the application of the sufficient conditions in Allman *et al.* (2009) applied by Shibata (2019) implies that our model is identified in this homogeneous case. We discuss that case in Section 6.

[19] There are 33 transition probabilities in each month. While these matrices are 4×4, due to adding up constraints and the assumption that there is no transition from long-term to short-term unemployment, the number of parameters is 33. The emission probability matrix is a 12×29 matrix but since markets are perfectly segmented and there is no measurement error, the total number of emission probabilities is $3 \times 2 + 6 \times 15 + 3 \times 9 = 123$. In addition we need to estimate the share of the population in each of the 12 hidden states which adds another 11 parameters to estimate. This implies a total of 167 parameters for each month.

report they are ((EPE,ES) in the table).

To illustrate the information that the algorithm distills from the differences between the columns in Table 3, consider Table 4. It contains four examples of how the probabilities that the model assigns to CPS respondents being part of the three market segments evolve depending on their reported history of emissions. All examples are for respondents in the sample from January 2005 through April 2006.

Example I is for someone who reports to be employed, not part-time for economic reasons and not absent from work, for all eight months she or he is in the survey. In the first month in the sample the model assigns an 89.2 percent probability this person is in the primary market and 7.4 percent and 3.4 for the secondary and tertiary markets respectively. Because of a lack of history in the first month, these are the unconditional probabilities of someone who reports this emission being in each of the market segments. As the individual continues to report the same type of employment for the subsequent months the likelihood that she or he is in the primary market increases for two reasons. The first is that longer employment spells are more likely in the primary market. The second is that those in the primary segment are more likely to report they are not absent from work and do not work part-time for economic reasons. The combination of these two effects yields a posterior probability, based on the whole reported history, of 99.7 percent that the person is in the primary market.

Example II shows the information that the extended 29 emissions provide in addition to the three basic labor-force statuses of employed, unemployed, and non-participation. Just like in Example I, the respondent reports to be employed in all eight months that they are in the sample. However, in the middle six months they report to be part-time employed for economic reasons. Even though this respondent is always employed, the algorithm assigns them with almost certainty to the secondary segment after their eight months in the sample. This is because those who report to be part-time employed for economic reasons are likely to flow into unemployment and non-participation and are therefore more likely to be in the secondary tier.

Example III shows that those persistently are nonparticipants who do not want a job are classified in the tertiary sector. While the emissions in the second four months reduce uncertainty about the respondent's sector, the first four months still provide an accurate assessment.

Example IV shows a mixed employment, unemployment, and non-participation history that results in more uncertainty about which segment the respondent belongs to. The advantage of our method is that it does not force us to classify this person in a particular segment. Instead, for both the aggregate and individual-level results we present in the next sections, we simply weigh respondents by their posterior probabilities across the three segments. Thus, we use the estimated individual-level posterior probabilities to construct aggregates for each of the market segments and analyze how they differ in terms of the levels of their unemployment and participation. The respondent in Example IV would thus be counted as 0.112 in the primary sector, 0.792 in the secondary, and 0.092 in the tertiary sector.

Example IV illustrates the importance of the second four months in the CPS's sampling structure. The algorithm gains additional information from the fact that the respondent is still a non-participant at the 5$^{\text{th}}$ month in the sample, in January 2006, after having been out of the sample for 8 months. This is especially important for those classified in the secondary sector, because that sector involves instability and reliably inferring instability from the data requires many observations.[20] In Example IV, the 4-8-4 reported mixed labor market history in the CPS leaves some uncertainty about which sector the respondent is in.

In our results, Example IV is the exception to the rule. Like Examples I-III, most respondents are reliably classified in one of the three segments in the sense that the mode of the posterior probabilities across the three market segments is well above 90%. To formally analyze how reliably the model classifies individuals, we introduce an additional summary statistic. This statistic is based on the distance between the estimated posterior distribution across labor market segments for each individual and the uniform distribution. To see why this distance captures the reliability of the classification of an individual, it is useful to write the estimated posterior distribution across the three markets as the triple

$$\left\{P_i\left(P\right), P_i\left(S\right), P_i\left(T\right)\right\}. \tag{16}$$

If the data provide no information on which of the segments an individual is in then this tuple is the uniform distribution $\{1/3, 1/3, 1/3\}$. If, on the other hand, the data perfectly pin down the tier then this triple is either $\{1, 0, 0\}$, $\{0, 1, 0\}$, or $\{0, 0, 1\}$, depending on whether the individual is in the primary, secondary, or tertiary market respectively. To provide a measure pf the degree of information the model provides about the segment membership of individual i in the sample, we calculate the rescaled distance of the posterior distributions from the non-informative, uniform, case. This measure is given by

$$D_i = \frac{\sqrt{9}}{\sqrt{6}}\sqrt{\sum_{M\in\{P,S,T\}}\left(P_i\left(M\right) - 1/3\right)^2} \in [0, 1]. \tag{17}$$

It is zero if the model does not provide any information about the segment membership of individual i and one if it is fully informative. Figure 2 shows the distribution of D_i across all

[20]Consistent with Example IV, estimates of the model based on only four observations result in a lower share of workers in the secondary sector.

Table 3: Average emission probabilities

State Emission	EP	ES	ET	UPS	UPL	USS	USL	UTS	UTL	NP	NS	NT
EX	98.71 (0.00)	68.13 (0.03)	93.40 (0.03)	-	-	-	-	-	-	-	-	-
EPE	1.01 (0.00)	29.42 (0.03)	2.17 (0.02)	-	-	-	-	-	-	-	-	-
ENW	0.29 (0.00)	2.45 (0.01)	4.43 (0.02)	-	-	-	-	-	-	-	-	-
UTL5	-	-	-	29.43 (0.12)	0.62 (0.02)	8.00 (0.05)	0.16 (0.01)	3.98 (0.08)	0.27 (0.02)	-	-	-
UTL14	-	-	-	4.57 (0.05)	13.34 (0.07)	4.12 (0.03)	0.38 (0.01)	0.73 (0.03)	0.97 (0.04)	-	-	-
UTL26	-	-	-	0.37 (0.02)	6.58 (0.05)	1.03 (0.02)	0.60 (0.01)	0.36 (0.02)	0.36 (0.02)	-	-	-
UTLLT	-	-	-	0.43 (0.02)	3.84 (0.04)	0.76 (0.01)	1.72 (0.02)	0.45 (0.03)	0.25 (0.02)	-	-	-
UTJ5	-	-	-	9.29 (0.07)	0.33 (0.02)	5.35 (0.04)	0.16 (0.01)	0.89 (0.04)	0.11 (0.01)	-	-	-
UTJ14	-	-	-	0.66 (0.02)	3.22 (0.04)	2.69 (0.03)	1.85 (0.02)	0.29 (0.02)	0.26 (0.02)	-	-	-
UTJ26	-	-	-	0.16 (0.01)	1.38 (0.03)	0.57 (0.01)	1.73 (0.02)	0.13 (0.01)	0.10 (0.01)	-	-	-
UTJLT	-	-	-	0.39 (0.02)	0.29 (0.01)	0.21 (0.01)	5.04 (0.04)	0.51 (0.03)	0.09 (0.02)	-	-	-
UJL5	-	-	-	31.99 (0.12)	2.90 (0.04)	9.55 (0.05)	1.21 (0.02)	1.55 (0.05)	0.18 (0.01)	-	-	-
UJL14	-	-	-	1.51 (0.03)	33.35 (0.10)	6.82 (0.04)	4.03 (0.03)	0.81 (0.03)	0.44 (0.02)	-	-	-
UJL26	-	-	-	0.19 (0.01)	19.66 (0.09)	1.31 (0.02)	6.03 (0.04)	0.45 (0.02)	0.24 (0.02)	-	-	-
UJLLT	-	-	-	0.93 (0.03)	11.64 (0.07)	0.45 (0.01)	28.57 (0.07)	1.73 (0.04)	0.26 (0.02)	-	-	-
UX5	-	-	-	16.27 (0.10)	0.51 (0.01)	34.06 (0.07)	1.48 (0.02)	71.33 (0.15)	6.90 (0.08)	-	-	-
UX14	-	-	-	2.12 (0.04)	1.68 (0.03)	18.58 (0.06)	9.66 (0.05)	7.72 (0.09)	54.26 (0.17)	-	-	-
UX26	-	-	-	0.50 (0.02)	0.47 (0.02)	3.79 (0.03)	7.36 (0.04)	0.73 (0.03)	23.90 (0.15)	-	-	-
UXLT	-	-	-	1.18 (0.03)	0.19 (0.01)	2.70 (0.02)	30.02 (0.07)	8.34 (0.11)	11.39 (0.08)	-	-	-
NTJDW	-	-	-	-	-	-	-	-	-	0.03 (0.00)	0.07 (0.00)	0.00 (0.00)
NTJMA	-	-	-	-	-	-	-	-	-	0.04 (0.00)	0.05 (0.00)	0.00 (0.00)
NTJNA	-	-	-	-	-	-	-	-	-	0.01 (0.00)	0.01 (0.00)	0.00 (0.00)
NTJNS	-	-	-	-	-	-	-	-	-	0.30 (0.01)	0.61 (0.01)	0.03 (0.00)
NTJDNW	-	-	-	-	-	-	-	-	-	1.13 (0.02)	0.86 (0.01)	0.15 (0.00)
NDW	-	-	-	-	-	-	-	-	-	0.88 (0.01)	3.57 (0.02)	0.07 (0.00)
NMA	-	-	-	-	-	-	-	-	-	1.41 (0.02)	6.94 (0.02)	0.12 (0.00)
NNA	-	-	-	-	-	-	-	-	-	0.39 (0.01)	1.93 (0.01)	0.06 (0.00)
NNS	-	-	-	-	-	-	-	-	-	9.96 (0.04)	19.04 (0.04)	1.29 (0.00)
NDNW	-	-	-	-	-	-	-	-	-	85.86 (0.05)	66.92 (0.04)	98.29 (0.00)

Notes: - Average probability of observed emission conditional on being in state over sample. No-classification-error restrictions are indicated by '-'. Bootstrapped standard errors in parentheses.

Table 4: Inference of market segments based on emissions

Date	Emission	P(P)	P(S)	P(T)
	Example I			
2005-01	Employed-not PTER+no other absence	89.2	7.4	3.4
2005-02	Employed-not PTER+no other absence	92.6	4.9	2.5
2005-03	Employed-not PTER+no other absence	94.8	3.3	1.9
2005-04	Employed-not PTER+no other absence	96.4	2.2	1.4
2006-01	Employed-not PTER+no other absence	98.9	0.9	0.2
2006-02	Employed-not PTER+no other absence	99.3	0.6	0.1
2006-03	Employed-not PTER+no other absence	99.5	0.4	0.1
2006-04	Employed-not PTER+no other absence	99.7	0.3	0.1
	Example II			
2005-01	Employed-not PTER+no other absence	89.2	7.4	3.4
2005-02	Employed-PTER	31.5	66.3	2.2
2005-03	Employed-PTER	1.7	98.2	0.1
2005-04	Employed-PTER	0.1	99.9	0.0
2006-01	Employed-PTER	0.0	100.0	0.0
2006-02	Employed-PTER	0.0	100.0	0.0
2006-03	Employed-PTER	0.0	100.0	0.0
2006-04	Employed-not PTER+no other absence	0.0	100.0	0.0
	Example III			
2005-01	Nonparticipants who do not want a job	4.4	7.2	88.4
2005-02	Nonparticipants who do not want a job	2.3	3.2	94.5
2005-03	Nonparticipants who do not want a job	1.1	1.5	97.4
2005-04	Nonparticipants who do not want a job	0.5	0.7	98.8
2006-01	Nonparticipants who do not want a job	0.0	0.1	99.8
2006-02	Nonparticipants who do not want a job	0.0	0.1	99.9
2006-03	Nonparticipants who do not want a job	0.0	0.0	100.0
2006-04	Nonparticipants who do not want a job	0.0	0.0	100.0
	Example IV			
2005-01	Employed-not PTER+no other absence	89.2	7.4	3.4
2005-02	U-Temporary job ended-less than 5 weeks	60.3	39.6	0.0
2005-03	Nonparticipants who do not want a job	45.1	54.8	0.0
2005-04	Nonparticipants who do not want a job	45.9	54.0	0.1
2006-01	Nonparticipants who do not want a job	10.4	88.7	0.8
2006-02	Nonparticipants who do not want a job	11.2	86.9	2.0
2006-03	Nonparticipants who do not want a job	12.0	83.5	4.5
2006-04	Nonparticipants who do not want a job	11.2	79.6	9.2

Source: Current Population Survey and authors' calculations.
Notes: Imputed probabilities of being in primary, secondary, or tertiary market segment for hypothetical emissions history.

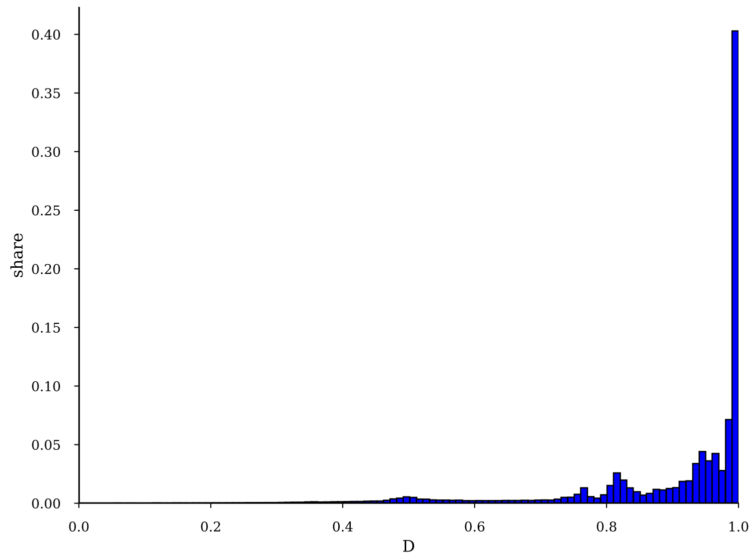


Figure 2: Distribution of D_i across CPS respondents

Source: CPS and authors' calculations.

CPS respondents for our baseline model. It shows that $D_i \geq 0.99$ for more than 40 percent of them and $D_i > 0.95$ for more than half of the respondents. Thus, the model is able to reliably classify the bulk of the individuals in different segments which is also evident in the distribution of the three posterior probabilities is shown in Figure B.1.

The benefit of using 29 emissions, rather than 3, in terms of the reliability with which CPS respondents are classified is reflected in the sample average of D_i, i.e. $\bar{D}$. It is 0.914 for our benchmark model with 29 emissions and 0.895 for the same specification using only 3 of them. This difference is highly statistically significant with a p-value of 0.00.[21]

4 Characteristics of each of the market segments

The most important finding from our analysis is that the U.S. labor market can be thought of as being comprised of three distinct segments, each of which is very different from the aggregate. The stark differences between the market tiers manifest themselves in the average outcomes over time, amount of turnover, as well as business cycle and seasonal fluctuations. In this section we highlight the main differences between the market segments along these dimensions. We quantify turnover using a new measure, i.e. the average number of flows between employment,

[21] Additionally, we report the shares of the population in each segment over time, $\delta_{l,t}$'s in Figure B.2. None of the inequality constraints end up binding in our estimation as can be seen in Table 6 which reports the average of estimated transition probabilities, $q_{l,l',t}$ and time-series of transitions probabilities between latent states,that are reported in Appendix B.

Table 5: Labor market aggregates by segment

	Primary	Secondary	Tertiary	Total
Share of population	54.46	13.75	31.79	100.00
	(1.11)	(1.08)	(0.87)	(0.00)
Unemployment rate	2.07	26.45	19.92	6.62
	(0.47)	(3.22)	(5.41)	(0.01)
Labor-force participation rate	97.16	72.92	8.84	65.77
	(0.62)	(3.27)	(1.54)	(0.01)
Employment-to-population ratio	95.15	53.55	7.05	61.42
	(0.80)	(3.92)	(1.40)	(0.01)
Flows per capita	0.50	3.20	0.62	0.91
	(0.07)	(0.31)	(0.10)	(0.00)

Source: Current Population Survey and authors' calculations.
Notes: Average of reported statistics over sample period for each market segment and total civilian non-institutionalized population 16-years and over. Flows per capita are annual flows between E, U, and N per person. Bootstrapped standard errors in parentheses.

unemployment, and non-participation per person.

Primary sector

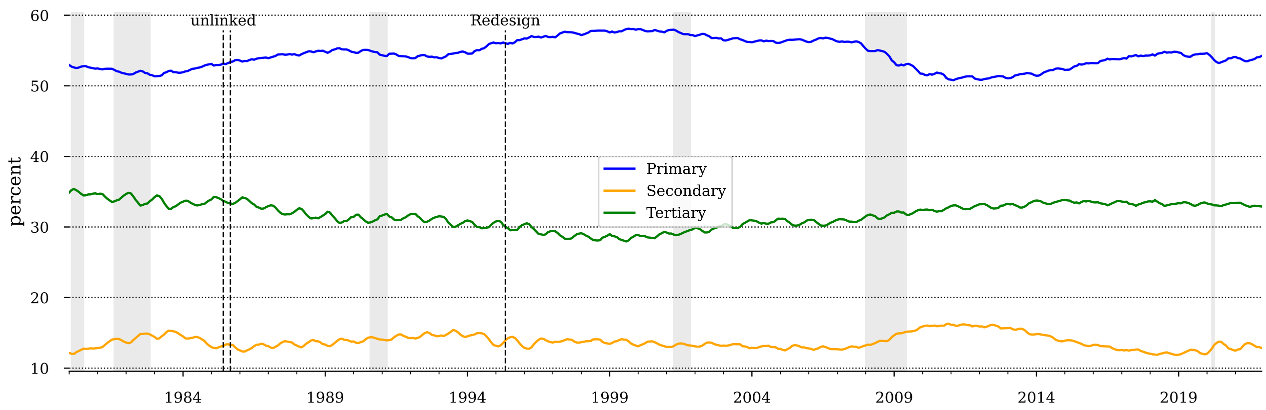
The primary sector consists of 54.5% of the population as can be seen from the top line of Table 5, which provides averages for the main labor market aggregates for each segment as well as for the labor market as a whole.[22] This share does not vary much over our sample period. This can be seen from Figure 3, which plots the time series of the estimated shares of the population in each segment.

Table 5 also shows that the primary sector is characterized by a low unemployment rate, high labor force participation rate (LFPR), and, consequently, a high employment-to-population ratio (EPOP). Moreover, at 0.5 flows per person per year, turnover in the primary sector is half that in the overall labor market. The blue bars in Figure 4 show the composition of these flows. They help put the low unemployment rate and high EPOP in the primary sector in context.

Labor market frictions are basically irrelevant for primary sector workers. These workers are almost always employed and very rarely experience unemployment. When they become unemployed, it is generally for a very short period, because their job-finding rates are much higher than those of others. This can be seen by comparing the 1-month flow rates from US and UL to E in Table 6 for primary sector with that in the overall labor market. They more or less seamlessly move from non-participation to employment and vice versa. In many ways one can think of the labor market experiences of these primary-sector workers as being captured

[22]The estimated time series on which Table 5 is based are plotted in Appendix Figures B.3 and B.4.

Figure 3: Share of population in each labor market segment.



well by those in standard Real Business Cycle (RBC) models (Cooley and Prescott, 1995).

Moreover, the unemployment rate in the primary segment fluctuates much less over the business cycle than in the labor market as a whole. This can be seen from the standard deviation, $\sigma(x)$ of the HP-filtered unemployment rate in Table 7. The standard deviation of cyclical component of the primary-sector unemployment rate is about half of that in the total labor market. That said, the business cycle correlation, $\rho(x_t, Y_t)$, shows that the unemployment rate in the primary sector is highly countercyclical.

The cyclical fluctuations in the primary-sector LFPR are also different. The magnitude of the cyclical fluctuations in primary-sector participation is smaller than that of the secondary sector's LFPR and the primary-sector LFPR is procyclical rather than countercyclical as is the case for the secondary and tertiary sectors.

Just like over the business cycle, the seasonal fluctuations in the primary segment of the labor market are subdued compared to the total as can be seen from the standard deviation of the seasonal components, $\sigma_{seas}(x)$ of the primary-sector labor market aggregates in Table 7.

Secondary sector

The secondary sector is almost the polar opposite of the primary, except for the fact that it also has a high LFPR. The unemployment rate in the secondary sector is more than ten times higher than in the primary sector and almost four times that of the labor market as a whole. Most notably, workers in the sector seem to be in a constant state of flux, as reflected by their flows per capita being six times higher than that of those in the primary sector. As Figure 4 shows, flow rates in the secondary sector are elevated for all six types of flows.

Table 6: 1-month transition probabilities in different market segments

from	segment to	Primary	Secondary	Tertiary	Total
E	E	97.92	84.89	71.65	95.41
		(0.00)	(0.02)	(0.05)	(0.00)
	US	0.73	6.80	1.83	1.49
		(0.00)	(0.02)	(0.02)	(0.00)
	UL	0.03	0.79	0.13	0.13
		(0.00)	(0.01)	(0.00)	(0.00)
	N	1.32	7.52	26.39	2.97
		(0.00)	(0.02)	(0.05)	(0.00)
US	E	51.35	31.73	18.18	34.19
		(0.14)	(0.07)	(0.15)	(0.06)
	US	8.39	31.14	8.26	24.12
		(0.07)	(0.07)	(0.08)	(0.05)
	UL	33.19	8.13	27.51	15.39
		(0.12)	(0.05)	(0.16)	(0.05)
	N	7.06	28.99	46.05	26.30
		(0.08)	(0.07)	(0.18)	(0.05)
UL	E	22.44	13.55	15.36	16.19
		(0.10)	(0.06)	(0.14)	(0.05)
	US	0.00	0.00	0.00	0.00
		(0.00)	(0.00)	(0.00)	(0.00)
	UL	70.13	63.21	63.99	65.38
		(0.11)	(0.08)	(0.16)	(0.06)
	N	7.43	23.24	20.65	18.43
		(0.07)	(0.07)	(0.15)	(0.05)
N	E	44.87	14.11	1.78	5.02
		(0.08)	(0.04)	(0.00)	(0.01)
	US	2.04	13.29	0.65	2.04
		(0.02)	(0.03)	(0.00)	(0.00)
	UL	1.94	6.88	0.12	0.95
		(0.02)	(0.02)	(0.00)	(0.00)
	N	51.14	65.71	97.45	91.99
		(0.08)	(0.05)	(0.00)	(0.01)

Source: Current Population Survey and authors' calculations.
Notes: Average 1-month transition probabilities between hidden states by segment and total. Bootstrapped standard errors in parentheses.

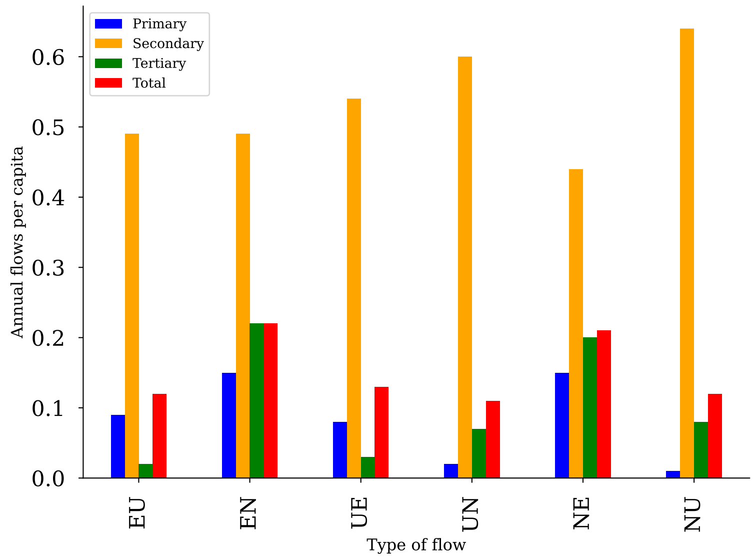


Figure 4: Average annual flows per capita by origin and destination.

Source: CPS and authors' calculations.

Notes: Averages taken over 1980-2021.

Table 7: Business cycle and seasonality statistics by labor market segment

measure	statistic	Primary	Secondary	Tertiary	Total
Unemployment rate	$\sigma(x)$	0.52	2.58	2.48	0.95
	$\rho(x_t, x_{t-1})$	0.71	0.78	0.81	0.79
	$\rho(x_t, Y_t)$	-0.74	-0.62	-0.49	-0.73
	$\sigma_{seas}(x)$	0.21	1.27	1.00	0.35
Labor-force participation rate	$\sigma(x)$	0.20	1.10	0.34	0.28
	$\rho(x_t, x_{t-1})$	0.61	0.81	0.67	0.77
	$\rho(x_t, Y_t)$	0.24	-0.28	-0.14	0.15
	$\sigma_{seas}(x)$	0.09	0.95	0.56	0.47
Employment-to-population ratio	$\sigma(x)$	0.62	1.99	0.37	0.75
	$\rho(x_t, x_{t-1})$	0.67	0.73	0.72	0.78
	$\rho(x_t, Y_t)$	0.68	0.47	0.21	0.65
	$\sigma_{seas}(x)$	0.24	1.22	0.42	0.53
Flows per capita	$\sigma(x)$	0.06	0.13	0.02	0.05
	$\rho(x_t, x_{t-1})$	0.51	0.54	0.36	0.55
	$\rho(x_t, Y_t)$	-0.59	-0.17	-0.11	-0.59
	$\sigma_{seas}(x)$	0.04	0.17	0.07	0.07

Source: Current Population Survey and authors' calculations.
Business-cycle variables - $\sigma(x)$: standard deviation of HP-filtered cyclical gap from quarterly seasonally adjusted data. $\rho(x_t, x_{t-1})$: first-order autocorrelation of HP-cyclical gap of variable. $\rho(x_t, Y_t)$: correlation of HP-cyclical gap of variable with that of GDP. HP-filter applied with smoothing parameter of 1600. $\sigma_{seas}(x)$: standard deviation of seasonal component. Seasonal component is difference between not-seasonally adjusted and seasonally adjusted montly time series.

Contrary to workers in the primary sector, labor market frictions are very relevant for workers in the secondary sector. Their labor market experience is characterized by intermittent periods of employment. They frequently move between labor market states and experience unemployment and non-participation spells very often. Because of the importance of labor market frictions for the outcomes of workers in the secondary sector, Diamond-Mortensen-Pissarides style search models are most applicable to this segment of the labor market.

Compared to their primary-sector counterparts, persons in the secondary market bear much more of the risk associated with business cycle fluctuations in the labor market. This can be seen by comparing the first and second columns of Table 7. The magnitude of cyclical fluctuations, $\sigma(x)$, in unemployment is five times higher in the secondary than in the primary segment. A notable difference is that, contrary to participation in the primary segment, the secondary-sector LFPR is countercyclical. That is, labor force participation in the secondary market tends to increase during economic downturns unlike the primary-sector LFPR. Differential cyclical behavior of participation in the primary and secondary sectors is informative about the difficulty of simultaneously accounting for the observed cyclical behavior of participation and unemployment rates (Veracierto (2008)). Our findings imply that the mild procyclicality of the aggregate LFPR is an outcome of procyclical and countercyclical forces which affect individuals differently; consistent with Krusell *et al.* (2017) who emphasize the role of heterogeneity in modeling the joint cyclical behavior of unemployment and participation.

The difference between the secondary and primary segments in the magnitude of seasonal fluctuations is even larger than over the business cycle as we can show in Table 7. Depending on the labor-market statistic, seasonal fluctuations, $\sigma_{seas}(x)$, in the secondary sector are five to ten times larger than those in the primary market. Seasonal fluctuations in the labor market are therefore disproportionately absorbed by those in the secondary segment.

Tertiary sector

Conventional discussions of the DLM focus on labor-market participants and ignore those who are only very loosely, or not attached, to the labor force. Our analysis, instead, covers the whole population, not only the labor force, to capture the increased importance of the participation margin for labor market trends and fluctuations. Consistent with this interpretation, Table 5 shows that only 9 percent of the persons in this tertiary sector participate in the labor market. Those in the tertiary segment that do participate have a relatively high unemployment rate of 20 percent.

Figure 4, Table 3, and Table 6 reveal that the nature of unemployment in the tertiary sector is very different from that in the secondary sector. Unemployment in the tertiary sector is

Table 8: Contribution to aggregates by segment

	Primary	Secondary	Tertiary	Total
Share of population	54.46	13.75	31.79	100.00
Unemployment rate	1.66	4.09	0.88	6.62
Labor-force participation rate	52.91	10.04	2.81	65.77
Employment-to-population ratio	51.83	7.36	2.24	61.42
Flows per capita	0.27	0.44	0.20	0.91

Source: Current Population Survey and authors' calculations.
Notes: Average percentage-point contribution by market segment to labor market aggregates over sample period. Flows per capita are annual flows between E,U, and N per person.

mostly because (re-)entrants into the labor force look for a job for a while before finding one. In the secondary sector job-loss is a much more important reason for unemployment. In terms of its sensitivity to business-cycle and seasonal fluctuations, the tertiary sector falls in between the primary and secondary sectors.

5 The Labor Market Through the Lens of the DLM

Thinking about the labor market as being made up of three distinct segments provides a very useful perspective of what drives labor market fluctuations as well as long run trends. As is implicitly implied in the previous section, particular segments play an outsized role in different aspects of the labor market. In this section we focus on three of the most prominent aspects: Unemployment, participation, and turnover. We show how the secondary sector drives both the long run trends as well as business cycle and seasonal fluctuations in the unemployment rate and turnover. The long-run trend in the labor supply is mostly due to changes in the composition of the population in terms of segment shares.

Unemployment

Even though the secondary market only consists of 14% of the population, the high unemployment rate in this segment means that more than 60% of unemployment in the labor market occurs among those in the secondary sector. This can be seen from Table 8 which reports the contributions of each of the market segments to the average aggregates listed in the column labeled "Total." The second row in the table shows that the secondary sector accounts for 4.1 percentage points of the 6.6 percent average unemployment rate in our sample. Thus, those in the secondary market are four times overrepresented among the unemployed.

This is an important observation for the following reason. A common practice is to use average transition rates between employment, unemployment (and non-participation) for quantitative macroeconomic analyses of the labor market, including for the calculation of the cost of unemployment (e.g. Krusell *et al.*, 2010). Such calculations tend to find that the costs of unemployment are low. This is because average flow rates imply that the average person is not likely to become unemployed and, if they do, they tend to be so for only a short period, because average unemployment outflow rates are high. However, looking through the lens of the DLM we see that these are not the relevant metrics to consider since the unemployment cost of business cycles is disproportionately borne by those in the secondary sector. Compared to the average, they are much more likely to become unemployed and to remain so once they are. A proper quantification of this cost should take this inequality into account and distinguish between the costs for workers in the three labor market tiers.

The differences between the market segments are not only important for the average level of unemployment. The secondary sector also accounts for about half of the fluctuations in the aggregate unemployment rate as can be seen from Table 9. The rows labeled with "$\sigma^2(\Delta x_t)$ " and "$\sigma^2(\Delta_{12} x_t)$" show the contributions of changes in the composition of the population (Share Total) and fluctuations in the labor market aggregates in each segment (Shift) to the monthly and 12-month changes in the aggregates respectively. The second row shows that the secondary market contributes 0.073 to the 0.154 variance of monthly changes in the unemployment rate.

The importance of the secondary sector for unemployment fluctuations reflects that this is the labor market that is most affected by search frictions. It is also the segment that contributes most to fluctuations in these frictions, as captured by shifts in the Beveridge curve. Figure 5 plots the Beveridge curves, using the aggregate vacancy rate measure from Barnichon (2010), for the three labor market segments as well as for the labor market as a whole. As has been widely documented, the aggregate Beveridge curve, plotted in panel (d) of the figure, exhibits a negative slope and has shifted several times over our sample period. Our results, plotted in panels (a) through (c) of Figure 5, indicate that the shifts in the aggregate Beveridge curve are mainly due to changes in match efficiency in the secondary sector. Therefore, explanations for changes in aggregate match efficiency, for example those that emphasize the importance of job-to-job transitions (Eeckhout and Lindenlaub (2019); Cheremukhin and Restrepo-Echavarria (2022); Moscarini and Postel-Vinay (2022)) are only plausible to the extent they are most applicable to workers in the secondary tier.

The secondary segment does not only account for the bulk of short-run fluctuations in the unemployment rate, it also accounts for more than half of the trend decline in the unemployment rate in the US. This decline is well documented (e.g. Shimer, 1999). The origin of this decline is the stark moderation in the incidence of unemployment which declined by more than 50%

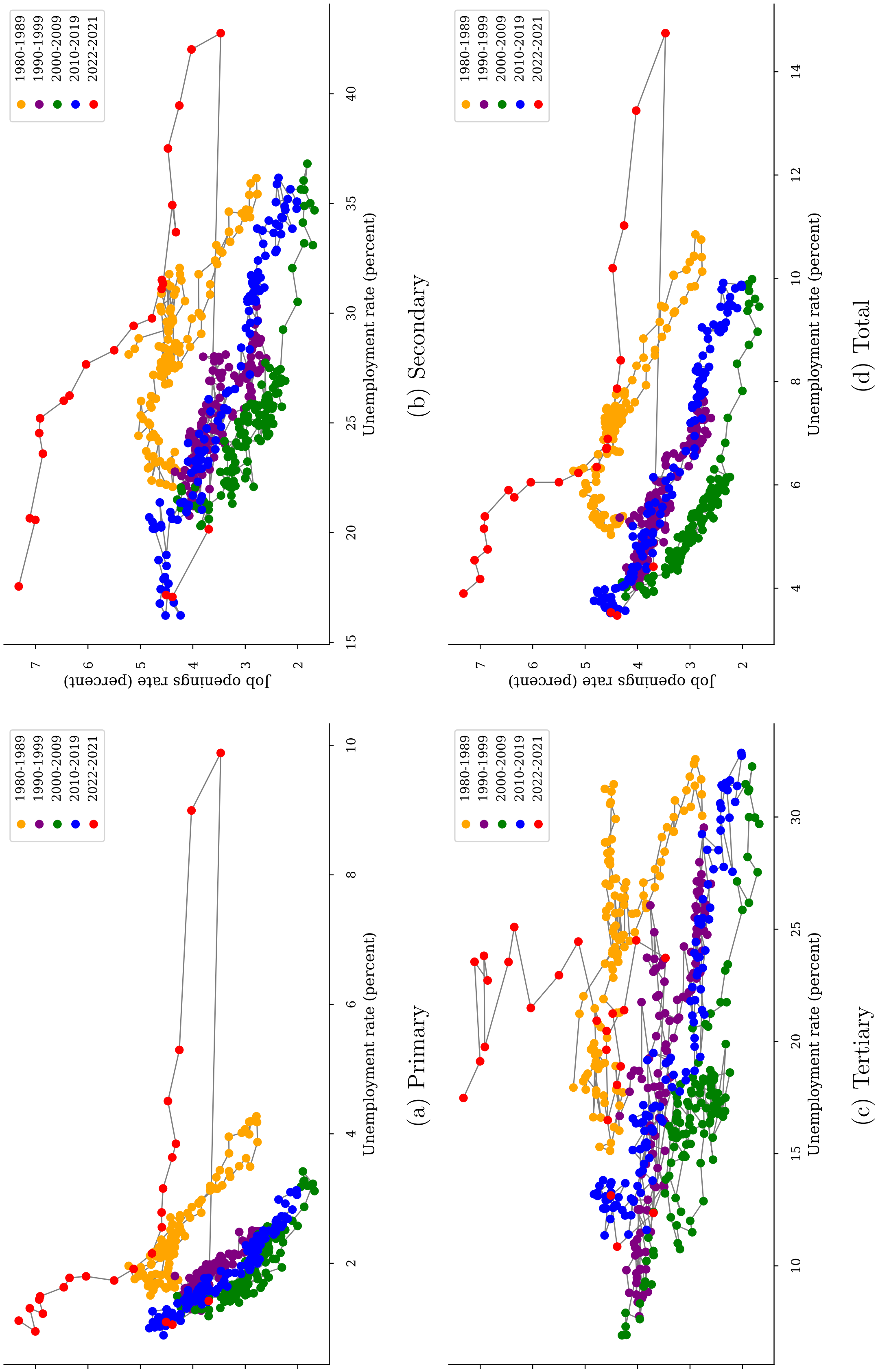


Figure 5: Beveridge curves for three market segments and total.

Source: BLS, Barnichon (2010), and authors' calculations. Monthly observations. Seasonally adjusted.

Table 9: Shift-share analysis of changes in labor market aggregates

		Sum Total	Share Total	Shift Primary	Secondary	Tertiary
Unemployment rate	$\bar{\Delta x_t}$	-0.0929	0.0023	-0.0240	-0.0577	-0.0135
	$\sigma^2(\Delta x_t)$	0.1546	0.0123	0.0516	0.0746	0.0161
	$\sigma^2(\Delta_{12} x_t)$	1.1299	0.1734	0.3628	0.4652	0.1285
Labor-force participation rate	$\bar{\Delta x_t}$	0.0001	0.0373	-0.0061	-0.0219	-0.0092
	$\sigma^2(\Delta x_t)$	0.1602	0.0329	0.0133	0.0574	0.0565
	$\sigma^2(\Delta_{12} x_t)$	0.1636	0.1102	0.0093	0.0207	0.0234
- 1980 - 2000	$\bar{\Delta x_t}$	0.1627	0.2262	-0.0029	-0.0518	-0.0088
	$\sigma^2(\Delta x_t)$	0.2148	0.0439	0.0164	0.0813	0.0731
	$\sigma^2(\Delta_{12} x_t)$	0.1179	0.0617	0.0016	0.0271	0.0275
- 2001 - 2021	$\bar{\Delta x_t}$	-0.1807	-0.1779	-0.0097	0.0156	-0.0087
	$\sigma^2(\Delta x_t)$	0.1006	0.0204	0.0100	0.0316	0.0387
	$\sigma^2(\Delta_{12} x_t)$	0.1582	0.0862	0.0200	0.0287	0.0233
Employment-to-population ratio	$\bar{\Delta x_t}$	0.0594	0.0335	0.0093	0.0219	-0.0053
	$\sigma^2(\Delta x_t)$	0.1802	0.0340	0.0380	0.0689	0.0393
	$\sigma^2(\Delta_{12} x_t)$	0.6837	0.2736	0.1655	0.1773	0.0672
Flows per capita	$\bar{\Delta x_t}$	-0.0060	-0.0004	-0.0005	-0.0037	-0.0014
	$\sigma^2(\Delta x_t)$	0.0081	0.0001	0.0025	0.0032	0.0022
	$\sigma^2(\Delta_{12} x_t)$	0.0022	0.0002	0.0008	0.0007	0.0005

Source: Current Population Survey and authors' calculations.
Notes: Contributions to average changes and the variance of 1-month and 12-month changes. See Appendix A for a derivation of the shift-share decomposition and calculation of contributions to the variance.

from 1980s to 2020s as evident from the job-loss and job destruction rates presented in Davis *et al.* (2010) and Crump *et al.* (2019). Table 9 reports that the unemployment declined on average by 0.093 percentage points per year in the 1981-2021 period. Of this decline, 60% (0.0577 percentage points annually) is due to a decrease in the unemployment rate in the secondary market. Changes in two flow rates are large drivers of this. First of all, a reduction in the secondary-market employment-to-unemployment transition rate from about 10% to 5%. Secondly, the inflow rate into short-term unemployment from non-participation went down from around 20% to 10%.[23] The primary sector, which encompasses almost 85% of employment, accounted for only 25% of the trend decline in unemployment.

To summarize, the 14% of the population that is part of the secondary tier of the labor market accounts for half or more of the level of, fluctuations in, and trend of aggregate unemployment. This observation is important for the discussion of policies that aim to stabilize labor-market fluctuations in the short-run. The focus of most of these policies is to maintain unemployment at or around Friedman's (Friedman, 1968) natural rate of unemployment. Our results imply that it is crucial to pay particularly close attention to those in the secondary tier of the labor who drive most of the movements in the unemployment rate. In light of our

[23]See Figures B.7 and B.10 for the relevant time series.

results, the unemployment insurance system can be thought of as a transfer from those in the primary and tertiary segments to workers in the secondary segment for absorbing a large part of aggregate economic risk over the seasons and business cycle.

Participation and Employment

While the bulk of the unemployed are in the secondary tier of the labor market, this segment accounts for only 15% of participation and 12% of employment as can be seen in Table 8. Most of the non-participants are in the tertiary sector with only 4.3% of participation accounted for by tertiary sector. Labor market production is concentrated in the primary sector with 80% of labor market participants and 84% of employed being part of the primary sector.

The LFPR trended up until the late 1990s and has been trending down since. We examine the drivers of the upward and downward trend separately by dividing our sample into two time periods: 1980-2000 and 2001-2021. For changes in trend participation, changes in the shares of persons in each of the market segments have been very important. The participation rate increased until the turn of the century as the share of the population in the primary sector increased and that in the tertiary sector declined reflecting the rise in female labor force participation as we discuss in detail in Section 7. Since 2000 this trend reversed because the aging of the population resulted in a larger share of individuals in the tertiary sector, with low labor force attachment.

Turnover, dynamics, and turbulence

Underlying the levels of and fluctuations in the unemployment and participation rates in the three market segments are the flows between unemployment, employment, and participation that drive the dynamics of the U.S. labor market. To compare the three market segments, we capture turnover in the labor market using the number of flows per capita. The last line in Table 8 shows that the U.S. labor market largely owes its dynamism to the secondary segment. Figure 6 shows that this segment makes up the majority of EU, UE, UN, and NU flows.

Since the 1990's the U.S. has seen a notable decline in labor market dynamism, first documented by Davis *et al.* (2007). The decline in dynamism is evident in many different labor market statistics such as job creation, job-to-job transitions and declining business formation as well as in flows per capita. Table 9 shows that there was a 0.006 decline in flows per capita per year which implies a decline in flows per capita from about 1.0 in 1980 to 0.8 in 2021. Similar to the trend in the unemployment rate, around 60% of the decline is accounted for by changes in the secondary sector. While there is also some decline in the tertiary sector, it is notable

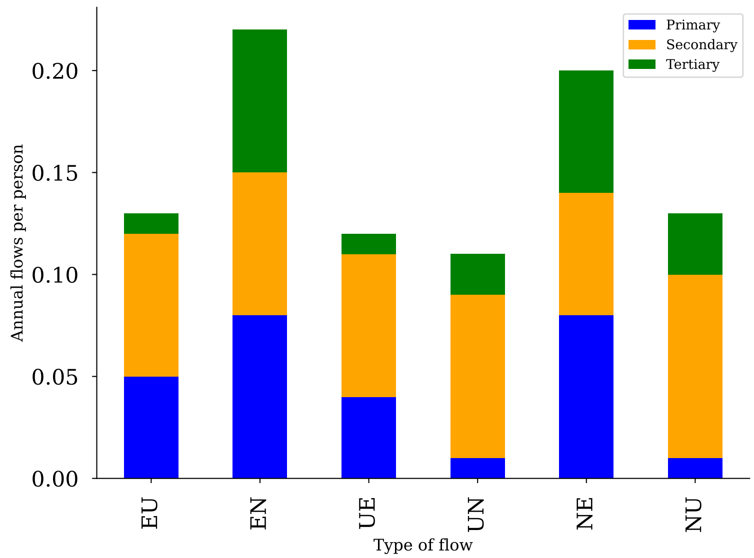


Figure 6: Composition of aggregate flows per person by type of flow and market.

Source: CPS and authors' calculations.
Notes: Averages taken over 1980-2021.

that flows per capita in the primary sector remained largely unchanged at 0.5 over the last 40 years.

6 Robustness and Alternative Model Specifications

The results we presented in the previous three sections are for our benchmark specification which is an HMM with: (*i*) three labor market segments, (*ii*) each with four hidden states (employed, two types of unemployed, and non-participant), (*iii*) time varying transition probabilities between and emission probabilities from these hidden states, and (*iv*) no mobility between the three market segments. Here we explain why this benchmark is our preferred specification and discuss the extent to which our qualitative results are robust to changes in these four main assumptions.

6.1 Model comparison and selection

Our choice of three labor market segments is guided by the DLM with a third sector for those only marginally attached to the labor force. This choice stands in stark contrast with the baseline case of a unified labor market with the three observed states of employment, unemployment, and non-participation for which transitions between these states follow a first-

Table 10: Comparison of model specifications

	segments	states	pars	logL	AIC	BIC
Dual Labor Market (DLM, benchmark)	3	12	84168	-3.41	69.54	70.73
First-Order Markov (FOM)	1	3	17136	-3.80	77.46	77.71
DLM, no tertiary sector	2	8	55944	-3.46	70.59	71.38
DLM, two types of U in secondary	3	10	63000	-3.43	69.94	70.83
DLM, no time-varying parameters	3	12	5823	-3.44	70.04	70.12

Source: Current Population Survey and authors' calculations.
Notes: Total number of observations is 10,178,593 CPS respondents from 1980-2021. Column definitions: *segments*: Number of labor market segments. *states*: Number of hidden states. *pars*: Number of parameters *LogL*: Mean log-likelihood across all individuals in sample. *AIC*: Akaike Information Criterion divided by 1000000. *BIC*: Bayesian Information Criterion divided by 1000000.

order Markov process. In the context of HMMs, this case is known as the First-Order Markov (FOM) model.[24] Our likelihood-based estimation method allows for direct model comparisons using the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). These information criteria indicate that our specification is preferred over the FOM.

This can be seen by comparing the first two rows of Table 10. The table lists the two information criteria as well as the number of labor market segments, hidden states, parameters, the resulting log-likelihood value for the benchmark model and four notable comparison specifications. Models with lower values for the AIC and BIC are preferred over those with higher ones. The rejection of the FOM compared to our benchmark is consistent with the evidence in Kudlyak and Lange (2017) who point out that observed individual-level transitions between E, U, and N are not first-order Markovian. A property that the benchmark model can capture but the FOM baseline, by assumption, cannot.

One might consider it ironic that this paper about the *dual* labor market actually contains a model with three market segments. Of course, this is to capture persistent non-participation, which was ignored in earlier analysis of the DLM but is important when considering the whole working-age population. The third line of Table 10 shows that the inclusion of third market segment is preferred over a "pure" DLM model that does not have a tertiary home production sector. This 2-segment specification lumps the primary and tertiary sectors together because both of them have low turnover rates.

The fourth row of the table illustrates why we include multiple types of unemployed persons in the specification. It shows that the model with only two types of unemployed persons in the secondary tier is rejected compared to our benchmark model. The improved fit for multiple

[24] For example, in the textbook Diamond-Mortensen-Pissarides model transitions between employment and unemployment follow a first-order Markov process. See Shibata (2019) for a further discussion of the usefulness of the FOM model as a baseline.

types of unemployed persons is consistent with a large number of studies that emphasize that they are necessary to match the unemployment duration distribution and the existence of long-term unemployment.[25] Other permutations of the model without multiple types of unemployed in different segments, not reported in the table, all yield higher information criteria than our benchmark specification.

The last row of the table reports the statistics from the benchmark model with constant transition and emission probabilities (constant-probability model). As we shut down time variation in the probabilities the number of parameters reduces substantially compared to our benchmark. The AIC prefers the non-homogeneous benchmark model over the homogeneous model with constant parameters, while the BIC prefers the homogeneous model over the benchmark model. This is because the BIC penalizes an increase in the number of parameters more than the AIC does. So, based on the information criteria, it is hard to distinguish between the benchmark model and the one with constant transition and emission probabilities.

6.2 Benefit of time-variation in the parameters

Even though the model selection criteria are inconclusive, we prefer the non-homogeneous benchmark with time-varying parameters over the model with constant parameters because it fits short-run and business cycle fluctuations in transition probabilities much better. As an example, consider the actual the unemployment-to-employment transition rates and those estimated using the benchmark and homogeneous models plotted in Figure 7. It clearly shows that, even though the model with constant transition and emission probabilities gets the average right, it does not capture the magnitude of both the seasonal as well as the cyclical fluctuations in the UE probability well. Capturing these movements in the EU rate correctly is important because they drive a large part of business cycle variation in the unemployment rate (Elsby *et al.*, 2015b).

The results are similar for other transition probabilities. The model with time-invariant transition and emission probabilities yields very similar average stocks and flow rates as our benchmark specification. However, it does not fit the movements in flow rates well. While it generates some business cycle fluctuations through variations in the shares of the population in each segment over time, i.e. through variation in the $\delta_{l,t}$'s, it underestimates business cycle fluctuations for all aggregates, especially for the unemployment rate as we report in Appendix Table B.1. This comparison emphasizes the importance of allowing for time variation in transition and emission probabilities for business cycle analysis.

[25] For example, Hornstein (2012), Kroft *et al.* (2016), and Ahn and Hamilton (2020a).

In principle, the fit of the homogeneous model can be improved by adding more hidden states and generating more changes in the fitted transition probabilities through shifts between those states. But that would violate our research objective of identifying Macro Heterogeneity using a *parsimonious* representation of individual and aggregate labor market dynamics in terms of a mixture of a *limited* number of hidden first-order Markov processes. Moreover, the more hidden states are added the harder it is to impose economically meaningful identifying restrictions that give each of them a distinct economic interpretation.

7 Reasons for labor market segmentation

At the core of discussions of the DLM is the implicit normative assessment that being part of the secondary tier of the labor market is undesirable. This perceived undesirability comes from the implication of DLM that jobs in the primary tier generally pay high wages, come with benefits, offer potential for job advancement, and provide job security. While jobs in the secondary tier have high turnover, pay low wages, come with limited benefits, offer few career opportunities, and provide little job security. We show that the market segments we identified indeed have these properties. We construct our evidence by matching the estimated individual-level posterior probabilities of market-segment membership with data from the CPS on demographic characteristics, industry and occupation of employment, as well as tenure and earnings.[26]

To better understand the context for this value judgment, it is important to analyze the potential sources of the segmentation of the labor market that we identified. Our ability to link estimated labor market segment membership with all observables in the CPS also allows us to explore reasons for why the dual labor market structure we observe emerged and has persevered. This is important, because early research on the DLM hypothesis was criticized for not coming to an agreement on these reasons.

Many causes of the segmentation have been emphasized by studies on the DLM. They can be broadly categorized into five themes: (*i*) Life-cycle career choices, (*ii*) discrimination, (*iii*) insider-outsider structure due to labor-market institutions and unionization, (*iv*) efficiency wage theory, and (*v*) labor demand fluctuations. These five mechanisms are closely intertwined. For example, the insider-outsider structure of certain parts of the U.S. labor market emerged along racial and gender lines.[27] Saint-Paul (1997) provides a link between efficiency wages,

[26] One can consider this as a verification of overidentifying restrictions.

[27] See Hill (1996) for discussion of discrimination by labor unions. Ashenfelter (1972) provides evidence of racial and gender discrimination by labor unions in the '60's and '70's. His evidence indicates that discrimination by unions was less than that in the labor market overall.

labor adjustment costs, and turnover in the wake of labor-demand fluctuations. We provide a set of facts that shine a light on the relative importance of each of the five main causes for labor-market segmentation considered in the literature.

Life-cycle career choices

Several papers emphasize how individuals learning about their own ability can result in some workers having a sequence of many short employment spells, while others find their comparative advantage and a good match early on in their careers and end up having much longer tenure (Morchio, 2020; Pries, 2004; Pries and Rogerson, 2021). At the core of each of these studies are different permutations of Jovanovic (1979).

Our results show a very distinct life-cycle pattern of labor-market segment membership. Figure 8 shows the share of the population of a certain age that is part of each market segment for six cohorts of the U.S. population covered by the CPS. All six cohorts show the same broad life-cycle patterns: Those younger than 25 are underrepresented in the primary sector and disproportionately work in the secondary sector. In their early twenties there is a gradual transition to the primary sector for most of them. This pattern suggests that, consistent with theories of life-cycle career choices, a large part of employment in the secondary sector is associated with early-career jobs. Table 11 provides some perspective on the importance of young persons for the secondary segment of the market. Those age 16 to 24 make up about a fifth of our sample and account for one third of those in the secondary tier.

However, panel (b) of Figure 8 also shows that the share of the prime-age population in the secondary tier of the labor market is of the same magnitude as that of the overall population. About one in every eight prime-age persons in the U.S. are part of the secondary segment of the labor market. This share has slowly increased across cohorts over time. Table 11 shows that, even though they are underrepresented in the secondary market, prime-age persons still make up the bulk of it.

Naturally, job choice is only one part of early-career decisions. The choice of education is the other part. Young people are likely to have jobs in the secondary segment while they work on their education that provides them access to jobs in the primary sector later on in their career. Table 11 shows that this mechanism is supported in the data. Those with a college education are overrepresented in the primary sector, while those with a high school education make up a disproportionate part of the secondary sector. However, one in every five people in the secondary sector has a college education. So, even though education seems to matter, the bifurcation by education into the primary and secondary tiers is not as stark as one might expect.

Table 11: Composition of market segment by demographic group

Topic	Segment Group	Primary	Secondary	Tertiary	Total
Sex	Male	54.8	49.8	37.7	48.5
	Female	45.2	50.2	62.3	51.5
Race	White	82.7	75.3	81.6	81.3
	Black	11.2	18.1	12.4	12.5
	Other	6.1	6.7	6.0	6.1
Ethnicity	Not hispanic	87.8	83.0	88.9	87.5
	Hispanic	12.2	17.0	11.1	12.5
Age	16-24	16.8	33.7	21.0	20.5
	25-54	68.8	53.1	29.2	53.6
	55 and over	14.4	13.1	49.8	25.9
Education	High school or less	42.0	56.8	63.0	50.6
	Some college	25.1	23.3	19.1	23.0
	College degree or higher	32.8	19.8	17.9	26.5
Citizenship	Born in U.S.	85.4	82.8	86.2	85.3
	Naturalized citizen	6.2	5.3	6.1	6.0
	Not a citizen	8.4	11.9	7.7	8.7
Unionization	No union coverage	85.6	88.3	87.0	86.0
	Covered by union but not a member	1.7	1.4	1.6	1.6
	Member of labor union	12.8	10.3	11.4	12.4

Source: Current Population Survey and authors' calculations.
Notes: *Some college* denotes some college or associates degree. *Born in U.S.* also includes those born abroad to American parents and those born in outlying U.S. areas.

Discrimination

Early work on the DLM Hypothesis focused on discrimination as the reason for segmentation. According to this view, the cause of duality is employer's discrimination against women, young adults, racial minorities, and immigrants (Doeringer and Piore, 1970; Berger *et al.*, 1980; Dickens and Lang, 1985). Consistent with this view, we find racial and ethnic minorities and non-naturalized immigrants are overrepresented in the secondary tier of the labor market compared to the primary tier as well as to the overall population, as summarized in Table 11. While the gender composition of the secondary sector is even, there are notable gender differences in the primary and tertiary tiers. Women make up more than 60 percent of the tertiary sector and but only 45% of the primary sector. These differences largely reflect the lower labor force attachment of women compared to men.

While the evidence in the Table 11 is consistent with discrimination being a source of segmentation, membership of one or more of the aforementioned groups only explains a small fraction of the cross-sectional variation in segment membership. Table 12 provides the results from regressions of the individual-level posterior probabilities of segment membership on a set of demographic characteristics. Specifically, we estimate

$$P_i\left(M\right) = \phi_{t,M} + \boldsymbol{x}_i'\boldsymbol{\beta}_M + \varepsilon_{i,M}, \text{ for } M \in \{P, S, T\}\,. \tag{18}$$

Here, $P_i(M)$ is the estimated posterior probability that individual i is part of segment M, as defined in (15). Segment-specific time dummies, $\phi_{t,M}$, are included to capture time trends in segment membership. The coefficients $\boldsymbol{\beta}_M$ measure the marginal likelihood that a person is part of segment M for the dependent variables in $\boldsymbol{x}$ that cover gender, education, race, and ethnicity. The estimated coefficients, $\hat{\boldsymbol{\beta}}_M$ for the three segments are reported in Table 12.

The estimated coefficients are in line with the results from Table 11. Women, young adults, and racial and ethnic minorities are all less likely to be in the primary sector than their counterparts. The R^2s for the three regressions, however, show that these demographic characteristics explain only a small part of the segmentation of the labor market. For all three regressions the R^2s are smaller than 0.25. The one for the tertiary sector is highest, mainly because of the life-cycle patterns we discussed above being captured by the age variables. Most notably, the R-squared for the secondary-tier regression is only 0.049. Even though we find evidence consistent with discrimination, it only accounts for a small portion of the segmentation of the labor market.

Moreover, the importance of demographic characteristics for segment membership has declined over time. Figure 9 shows time series for the equivalent of the coefficients for gender, race, and ethnicity from Table 12 for 10-year rolling samples.[28] The coefficients for women, Blacks, and Hispanics for the primary segment all increased over time. This shows that the wedge between white men and those of other gender, race, and/or ethnicity, with similar other characteristics, in primary-sector membership declined over time.

As can be seen from panel (a) in the figure, for women this convergence was entirely due to their shift from the tertiary to the primary segment over time. This reflects an increase in female labor force attachment and is consistent with Goldin and Mitchell (2017). However, women remain less likely to belong to the primary sector even though secondary sector probabilities do not vary by gender. This finding showcases the importance of the tertiary sector in capturing some important labor market developments.

Panels (b) and (c) show the effect of race and ethnicity on segment membership. Workers who identify as Black were 10% less likely to be in the primary segment in 1990 controlling for gender, age, and education which has declined to 5% by 2020. For Hispanic workers the convergence was even more stark. Despite this convergence, the shares of each segment remained stable in the last forty years, again suggesting that factors other than discrimination likely have affected the segment membership.

Insider-outsider structure of labor market

Early discussions of duality in the U.S. labor market have highlighted different eras of economic distress during which the forces that result in a DLM emerged. For example, Reich *et al.* (1973) focus on the late Nineteenth Century while Berger *et al.* (1980) claim dualism emerged in response to the legislation and labor movements in the wake of the Great Depression during the 1930's. One view is that unionization resulted in insiders and outsiders in the labor market. Insiders getting access to stable high paid careers with benefits, while outsiders do not.[29]

Table 13 shows that those in the primary sector do have more stable and higher paid jobs using information on usual median weekly earnings and hours from the CPS and tenure information from its Job Tenure Supplement. The top two measures in the table pertain to job stability. Those in the secondary sector switch from job to job twice as frequently as their counterparts in the primary sector. They also have much shorter tenure than those in the primary sector.[30] Those in the primary sector are mostly full-time employed, i.e. 35 hours a

[28] Figures B.11, B.12 and B.13 in the Appendix show the time series for the evolution of all estimated coefficients in Table 12.

[29] The importance of unions and labor market institutions has been at the heart of the analysis of dualism in European labor markets, e.g. Bentolila *et al.* (2019).

[30] One interesting result is the right tail of the tenure distribution for those in the tertiary sector. This suggests that there is a substantial number of respondents in the CPS that are out the labor force for different reasons and then return to their former employer.

Table 12: Regression of segment probabilities on demographic characteristics

	Primary	**Secondary**	**Tertiary**
Female	-0.1205	-0.0051	0.1256
	(-471.03)	(-30.423)	(528.31)
16-24	-0.1151	0.0617	0.0534
	(-267.90)	(217.18)	(133.70)
55 and over	-0.3660	-0.0654	0.4314
	(-1170.1)	(-316.36)	(1483.6)
Less than high school	-0.2278	0.0543	0.1735
	(-532.20)	(191.97)	(435.92)
High school diploma	-0.1233	0.0378	0.0855
	(-300.40)	(139.27)	(224.07)
Some college	-0.0703	0.0274	0.0429
	(-171.71)	(101.05)	(112.83)
Black	-0.0706	0.0615	0.0090
	(-180.16)	(237.60)	(24.799)
Other	-0.0591	0.0175	0.0416
	(-109.86)	(49.124)	(83.233)
Hispanic	-0.0297	0.0391	-0.0094
	(-73.941)	(147.18)	(-25.146)
R-squared	0.1912	0.0489	0.2335

Source: Current Population Survey and authors' calculations.
Notes: Number of observations is 10135696. Time fixed effects are included in all regressions. t-statistics reported in parentheses. Dummies are normalized to represent a prime-age college-educated white non-hispanic male as the baseline.

week or higher, while more than half the employed respondents in the secondary and tertiary sectors report working less than 35 hours a week.

This lower number of hours is reflected in usual weekly earnings. The fourth measure in Table 13 reports the distribution of relative usual weekly earnings measured as the percent deviation of an individuals' usual weekly earnings from the median usual weekly earnings in the year. Half of those employed in the secondary segment make 45 percent less than median earnings. This number is similar for those working that are part of the tertiary sector. The difference in weekly earnings is not all due to the fact that those in the secondary and tertiary sectors tend to work less hours. The last measure in Table 13 shows that median hourly earnings in the primary sector are about 30 percent higher than those in the other two tiers.

These differences between the types of jobs in the primary and secondary sectors in terms of stability and earnings are consistent with the DLM hypothesis. They could reflect that the primary sector is made up of insiders while the secondary sector consists of outsiders. However, we find only limited importance of unionization for the dualism of the U.S. labor market. This can be seen from the part of Table 11 related to "unionization". In our sample, 12.8 percent of those in the primary tier report to be members of a union, while only a slightly lower percentage in the secondary tier, i.e. 10.3, does so. Moreover, our results cover 1980-2021 and indicate that the dualism of the U.S. labor market was just as pronounced in the 1980's as in the 2010's. However, union coverage of the U.S. payroll employed halved during that period. This suggests that unions and labor movements more generally, though they might possibly play a role, are not the most important factor driving dualism in the U.S. labor market.

Efficiency wages

Insider-outsider effects are only one possible explanation for the differences between jobs in the primary and secondary sectors reported in Table 13. Several studies have emphasized that segmentation of the labor market can be the equilibrium outcome in the presence of market imperfections rather than institutionalized in terms of legislation of unionization. In particular, dualism can emerge when workers' effort on jobs in the primary sector is hard to monitor and on those in the secondary sector it is not. This results in an efficiency wage structure in the former and a competitive wage structure in the latter (Bulow and Summers, 1986; Albrecht and Vroman, 1992; Saint-Paul, 1997). The equilibrium outcome is higher job stability and wages in the primary sector than in the secondary one.

Our results provide two pieces of support for this efficiency wage theory of dualism. First of all, occupations with higher shares of workers in the primary sector tend to be high-skilled service occupations where effort is hard to monitor and efficiency-wage considerations are likely

Table 13: Turnover, tenure, hours worked, and earnings

measure	statistic	Primary	Secondary	Tertiary	Total
J2J rate	mean	2.1	4.5	3.3	2.4
Tenure	10th percentile	0.5	0.2	0.2	0.4
	25th percentile	1.9	0.5	0.5	1.5
	median	5.0	1.8	2.0	4.0
	75th percentile	11.0	5.0	7.0	10.0
	90th percentile	20.0	12.0	20.0	20.0
Weekly hours	10th percentile	30	16	10	24
	25th percentile	40	25	20	40
	median	40	37	40	40
	75th percentile	42	40	40	40
	90th percentile	50	44	45	50
Weekly earnings	10th percentile	-57.3	-79.2	-86.1	-65.7
	25th percentile	-31.3	-66.0	-73.2	-39.3
	median	8.3	-44.6	-43.8	0.0
	75th percentile	66.7	-10.0	1.0	59.2
	90th percentile	138.4	47.1	66.6	130.3
Hourly earnings	10th percentile	-47.5	-56.3	-57.1	-50.0
	25th percentile	-29.5	-48.2	-48.5	-34.1
	median	5.9	-32.2	-29.4	0.0
	75th percentile	69.6	8.5	21.2	63.5
	90th percentile	191.5	139.2	177.8	188.4

Source: Current Population Survey and authors' calculations.
Notes: *Tenure* Percent of employed that change employers the next month for those who responded to this question (starting in 1994). *Weekly hours* Usual weekly hours worked on all jobs. *Weekly and Hourly earnings* percent deviation from annual median weekly and hourly earnings.

to play a role in compensation. This can be seen from Panel (a) of Figure 10. The evidence for industries, shown in Panel (b) of the same figure, is not as clear-cut.[31]

The second piece of evidence, in line with that provided by Dickens and Lang (1985), is that there are significant differences in wage dynamics across the labor market segments. Jobs in the primary sector pay both a higher return to schooling as well as to experience. We show this by running a generalized Mincer (1974) regression in which we include segment-specific coefficients. The Mincer regressions are of the form:

$$w_{i,t} = \phi_t + \sum_{M \in \{P,S,T\}} P_i(M)\, \boldsymbol{x}'_{i,t} \boldsymbol{\beta}_M + \varepsilon_{i,M}, \text{ for } M \in \{P, S, T\}. \tag{19}$$

Here, $w_{i,t}$ is log hourly earnings of individual i in month t. The time fixed effect, ϕ_t, for a time trend in wage growth and prices. The vector of regressors, $\boldsymbol{x}_{i,t}$, includes education, experience, and experience squared.[32] The coefficients $\boldsymbol{\beta}_M$ can be interpreted as the Mincer coefficients associated with segment M. Their estimates are what is reported in Table 14.

The table shows that the return to a year of schooling is about 1.3 percentage points higher in the primary sector than in the secondary and tertiary ones. This is true for both men and women. Returns to experience are also higher in the primary sector, by about 1.1 percentage points for a year. of experience. These results imply that lifetime earnings are also likely to be lower for workers in the secondary sector; a finding that resonates with the findings of Karahan *et al.* (2023). Using earnings data from U.S. Social Security Administration, Karahan *et al.* (2023) group workers by their lifetime earnings and examine the role of different factors. They find that low lifetime earning workers switch jobs more often and face substantially high job loss risk while high lifetime earnings workers enjoy a high level of returns to experience with little career interruptions. Borovičková and Macaluso (2023) document similar patterns in career dynamics in the Austrian labor market.

Demand fluctuations

Another potential reason for duality is that it allows workers and firms to organize in a way that insulates jobs that involve match-specific capital from being dissolved in response to negative economic shocks. Piore calls this the endogenous "response to flux and uncertainty" (See Berger *et al.*, 1980, Chapter 2). Saint-Paul (1997) illustrates the same intuition in the first figure of his book in the context of labor-adjustment costs due to efficiency wages.

[31] Efficiency-wage theory is most often analyzed by looking at inter-industry wage differentials (e.g Krueger and Summers, 1988) rather than at the occupational level.

[32] Here, experience is "potential"experience that is calculated as the age of the individual at time t minus their years of education.

Table 14: Mincer regressions separated by market segment

	Total	**Male**	**Female**
Dep. Variable	hourly wage (log)	hourly wage (log)	hourly wage (log)
Years of schooling - Primary	0.0640	0.0582	0.0822
	(681.66)	(458.96)	(622.61)
Years of schooling - Secondary	0.0504	0.0443	0.0685
	(384.98)	(243.90)	(384.48)
Years of schooling - Tertiary	0.0521	0.0442	0.0725
	(266.80)	(153.81)	(287.71)
Experience - Primary	0.0349	0.0407	0.0277
	(612.07)	(508.85)	(364.14)
Experience - Secondary	0.0216	0.0290	0.0181
	(152.77)	(141.96)	(98.481)
Experience - Tertiary	0.0209	0.0322	0.0210
	(78.123)	(71.390)	(65.725)
Experience^2 - Primary	-0.0006	-0.0007	-0.0004
	(-476.85)	(-391.20)	(-277.11)
Experience^2 - Secondary	-0.0003	-0.0004	-0.0002
	(-96.285)	(-91.412)	(-57.794)
Experience^2 - Tertiary	-0.0003	-0.0005	-0.0003
	(-47.955)	(-49.593)	(-44.021)
No. Observations	4051029	1971912	2079117
R-squared	0.2488	0.2828	0.2658
Effects	Time	Time	Time

Source: Current Population Survey and authors' calculations.
Notes: Year fixed effects are included in all regressions. t-statistics reported in parentheses. Experience is defined as age minus years of schooling minus six.

Most discussions of this channel emphasize this mechanism in the context of business cycle fluctuations. However, the non-seasonally-adjusted nature of the individual-level CPS data we use reveals that something similar is true at seasonal frequencies. This points to dualism in the U.S. labor market persisting to organize the division of labor in the face of labor market imperfections in a way that minimizes adjustment costs in response to predictable seasonal as well as unpredictable business cycle fluctuations. Our results likely underestimate the importance of this channel because the CPS data we use is collected at the monthly frequency. It does not include the high-frequency turnover in the labor market.[33]

What remains an open question is what determines why particular workers end up in different labor market segments. The model in Albrecht and Vroman (1992) emphasizes heterogeneity in the value of non-employment as a major determinant of this. To make progress on this question and test this hypothesis, better measures of this value at the individual level would be required.

8 Conclusion

The dynamics of the stocks and flows in the U.S. labor market are well captured by a DLM with a tertiary sector made up of those who participate infrequently. This interpretation provides a parsimonious framework within which many aspects that have puzzled labor- and macroeconomists can be interpreted. The three market segments can be disentangled using an unsupervised machine learning method that involves the estimation of an HMM with identifying inequality constraints on the transition probabilities. These restrictions are what ensures that the hidden states we uncover can be interpreted as making up the primary, secondary, and tertiary labor-market tiers. What emerges is a tale of three totally different sub-markets.

Labor market frictions are basically irrelevant for primary sector workers who make up around 55 percent of the population. These workers are almost always employed and they very rarely experience unemployment. They also seamlessly move from non-participation to employment unlike workers in the secondary and tertiary sectors. The secondary sector, which constitutes 14 percent of the population, exhibits high turnover and high unemployment and absorbs most of the short-run fluctuations in the labor market, at both seasonal and business cycle frequencies. Workers in this sector are six times more likely to move between labor market states than those in the primary tier and are 10 more likely to be unemployed than their primary counterparts. The tertiary sector mostly includes workers who are only loosely

[33]For example, the Starbucks barista who only works peak-demand shifts during workdays from 7.30am through 10.30am and spends the rest of the day studying for her law degree.

attached to the labor market and has a very low employment-to-population ratio. These workers mostly experience unemployment when they enter the labor force from non-participation but do not share the high job-loss rate of secondary workers.

The qualitative properties for the auxiliary models we estimated over the four-month samples of the CPS are very similar to the ones we discussed for our benchmark model. We thus consider our results robust to the assumption of no-mobility between markets that we impose. However, we recognize that estimation of a DLM-HMM using data with longer labor market histories than in the CPS would potentially allow for a better quantification of cross-sector mobility. We view this as a promising research agenda applicable to different data sources, potentially from countries other than the U.S..

Because the total labor market is the sum of these three very different parts, average outcomes, which are often used for to quantitatively discipline macroeconomic models of the labor market, are not reflective of the labor market experiences of anyone in the population. This observation helps put into perspective the difficulties mainstream quantitative models face in capturing key aspects of labor market dynamics such as the *unemployment volatility puzzle* (Hall, 2005; Shimer, 2005), the *slow recovery puzzle* (Cole and Rogerson, 1999) and the difficulty of *simultaneously* accounting for the observed behavior of participation and unemployment (Veracierto, 2008). Our analysis suggests that better quantitative analyses should take into account the Macro Heterogeneity in labor market outcomes we uncovered in this paper.

The combination of the aggregate and individual-level evidence we provide points to dualism in the U.S. labor market being an equilibrium division of labor, under labor market imperfections, that minimizes adjustment costs in response to predictable seasonal as well as unpredictable business cycle fluctuations. These observations resonate with the dualism observed in the European Labor Markets (Bentolila *et al.*, 2019). Interestingly, the shares of employees with fixed-term contracts in France, Netherlands and Portugal are close to our estimates of the share of employment in the secondary segment supporting the view that differences in production technologies, demand fluctuations and adjustments costs could be the drivers of dualism. It remains an open question why seemingly similar workers end up in different segments of the labor market. We believe that differences in the value workers put on leisure or the cost of labor market participation together with differences in effort requirements of different careers —a hypothesis put forth by Albrecht and Vroman (1992)— is a promising explanation.

References

Abowd, John M., and Zellner, Arnold. 1985. Estimating Gross Labor-Force Flows. *Journal of Business and Economic Statistics*, **3**(3), 254–283.

Abraham, Katharine G, Haltiwanger, John C, and Rendell, Lea E. 2020. How Tight is the U.S. Labor Market? *Brookings Papers on Economic Activity*, **2020**(1), 97–165.

Ahn, Hie Joo, and Hamilton, James D. 2020a. Heterogeneity and Unemployment Dynamics. *Journal of Business and Economic Statistics*, **38**(3), 554–569.

Ahn, Hie Joo, and Hamilton, James D. 2020b (June). *Measuring Labor-Force Participation and the Incidence and Duration of Unemployment*. Working Paper 27394. National Bureau of Economic Research.

Albrecht, James W., and Vroman, Susan B. 1992. Dual Labor Markets, Efficiency Wages, and Search. *Journal of Labor Economics*, **10**(4), 438–461.

Allman, Elizabeth S., Matias, Catherine, and Rhodes, John A. 2009. Identifiability of parameters in latent structure models with many observed variables. *Ann. Statist.*, **37**(6A), 3099–3132.

Andersen, Martin, Dahl, Joachim, Liu, Zhang, and Vandenberghe, Lieven. 2011. Interior-point methods for large-scale cone programming. *Optimization for machine learning*, **5583**.

Ashenfelter, Orley. 1972. Racial Discrimination and Trade Unionism. *Journal of Political Economy*, **80**(3), 435–464.

Barnichon, Regis. 2010. Building a composite Help-Wanted Index. *Economics Letters*, **109**(3), 175–178.

Barnichon, Regis, and Nekarda, Christopher J. 2012. The Ins and Outs of Forecasting Unemployment: Using Labor Force Flows to Forecast the Labor Market. *Brookings Papers on Economic Activity*, **43**(2 (Fall)), 83–131.

Baum, Leonard E, Petrie, Ted, Soules, George, and Weiss, Norman. 1970. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *The annals of mathematical statistics*, **41**(1), 164–171.

Baumeister, Christiane, and Hamilton, James D. 2015. Sign Restrictions, Structural Vector Autoregressions, and Useful Prior Information. *Econometrica*, **83**(5), 1963–1999.

Bentolila, Samuel, Dolado, Juan J., and Jimeno, Juan F. 2019 (Jan.). *Dual Labour Markets Revisited*. IZA Discussion Papers 12126. Institute of Labor Economics (IZA).

Berger, S., Berger, P.P.S.S., Piore, M.J., Suzanne, B., and Press, Cambridge University. 1980. *Dualism and Discontinuity in Industrial Societies*. Cambridge University Press.

Blanchard, Olivier, and Diamond, Peter. 1990. The Cyclical Behavior of the Gross Flows of U.S. Workers. *Brookings Papers on Economic Activity*, **1990**(2), 85–155.

Boeschoten, Laura, Filipponi, Danila, and Varriale, Roberta. 2019. Combining Multiple Imputation and Hidden Markov Modeling to Obtain Consistent Estimates of Employment Status. *Journal of Survey Statistics and Methodology*, 01, 1–25.

Borovičková, Katarína, and Macaluso, Claudia. 2023. *Heterogeneous Job Ladders*. Working Paper. Richmond Fed.

Bulow, Jeremy I, and Summers, Lawrence. 1986. A Theory of Dual Labor Markets with Application to Industrial Policy,Discrimination, and Keynesian Unemployment. *Journal of Labor Economics*, **4**(3), 376–414.

Cain, Glen G. 1975. The Challenge of Dual and Radical Theories of the Labor Market to Orthodox Theory. *The American Economic Review*, **65**(2), 16–22.

Cheremukhin, Anton A., and Restrepo-Echavarria, Paulina. 2022 (Sept.). *The Dual Beveridge Curve*. Working Papers 2022-021. Federal Reserve Bank of St. Louis.

Christiano, Lawrence J, Eichenbaum, Martin, and Vigfusson, Robert. 2006 (July). *Assessing Structural VARs*. Working Paper 12353. National Bureau of Economic Research.

Cole, Harold, and Rogerson, Richard. 1999. Can the Mortensen-Pissarides Matching Model Match the Business-Cycle Facts? *International Economic Review*, **40**(4), 933–59.

Cooley, Thomas F., and Prescott, Edward C. 1995. Economic Growth and Business Cycles. *Chap. 1, pages 1–38 of:* Cooley, Thomas F. (ed), *Frontiers of Business Cycle Research*. Princeton, New Jersey: Princeton University Press.

Costain, James, Jimeno, Juan F., and Thomas, Carlos. 2010 (Apr.). *Employment fluctuations in a dual labor market*. Working Papers 1013. Banco de España.

Crump, Richard K., Eusepi, Stefano, Giannoni, Marc, and Şahin, Ayşegül. 2019 (June). *A Unified Approach to Measuring u*. NBER Working Papers 25930. National Bureau of Economic Research, Inc.

Daly, Mary C., Hobijn, Bart, Şahin, Ayşegül, and Valletta, Robert G. 2012. A Search and Matching Approach to Labor Markets: Did the Natural Rate of Unemployment Rise? *Journal of Economic Perspectives*, **26**(3), 3–26.

Davis, Steven J., Haltiwanger, John, Jarmin, Ron, and Miranda, Javier. 2007. Volatility and Dispersion in Business Growth Rates: Publicly Traded versus Privately Held Firms. *Pages 107–180 of: NBER Macroeconomics Annual 2006, Volume 21.* NBER Chapters. National Bureau of Economic Research, Inc.

Davis, Steven J., Faberman, R. Jason, Haltiwanger, John, Jarmin, Ron, and Miranda, Javier. 2010. Business Volatility, Job Destruction, and Unemployment. *American Economic Journal: Macroeconomics*, **2**(2), 259–287.

Dempster, Arthur P, Laird, Nan M, and Rubin, Donald B. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, **39**(1), 1–22.

Dickens, William T., and Lang, Kevin. 1985. A Test of Dual Labor Market Theory. *The American Economic Review*, **75**(4), 792–805.

Doeringer, P.B., and Piore, M.J. 1970. *Internal Labor Markets and Manpower Analysis*. Manpower Administration (Department of Labor).

Eeckhout, Jan, and Lindenlaub, Ilse. 2019. Unemployment Cycles. *American Economic Journal: Macroeconomics*, **11**(4), 175–234.

Elsby, Michael W. L., Hobijn, Bart, Şahin, Ayşegül, and Valletta, Robert G. 2011. The Labor Market in the Great Recession — An Update to September 2011. *Brookings Papers on Economic Activity*, **42**(2 (Fall)), 353–384.

Elsby, Michael W. L., Michaels, Ryan, and Ratner, David. 2015a. The Beveridge Curve: A Survey. *Journal of Economic Literature*, **53**(3), 571–630.

Elsby, Michael W.L., Hobijn, Bart, and Şahin, Ayşegül. 2015b. On the Importance of the Participation Margin for Labor Market Fluctuations. *Journal of Monetary Economics*, **72**(C), 64–82.

Feng, Shuaizhang, and Hu, Yingyao. 2013. Misclassification Errors and the Underestimation of the US Unemployment Rate. *American Economic Review*, **103**(2), 1054–70.

Ferraro, Domenico. 2018. The Asymmetric Cyclical Behavior of the U.S. Labor Market. *Review of Economic Dynamics*, **30**(October), 145–162.

Flood, Sarah, King, Miriam, Rodgers, Renae, Ruggles, Steven, and Warren, J. Robert. 2020. *Integrated Public Use Microdata Series, Current Population Survey, Version 8.0 [dataset]*. University of Minnesota.

Friedman, Milton. 1968. The Role of Monetary Policy. *The American Economic Review*, **58**(1), 1–17.

Goldin, Claudia, and Mitchell, Joshua. 2017. The New Life Cycle of Women's Employment: Disappearing Humps, Sagging Middles, Expanding Tops. *Journal of Economic Perspectives*, **31**(1), 161–82.

Granger, C. W. J., and Morris, M. J. 1976. Time Series Modelling and Interpretation. *Journal of the Royal Statistical Society. Series A (General)*, **139**(2), 246–257.

Gregory, Victoria, Menzio, Guido, and Wiczer, David G. 2021 (Apr.). *The Alpha Beta Gamma of the Labor Market*. NBER Working Papers 28663. National Bureau of Economic Research, Inc.

Guner, Nezih, Ventura, Gustavo, and Xu, Yi. 2008. Macroeconomic implications of size-dependent policies. *Review of Economic Dynamics*, **11**(4), 721–744.

Hall, Robert E. 1982. The Importance of Lifetime Jobs in the U.S. Economy. *The American Economic Review*, **72**(4), 716–724.

Hall, Robert E. 2005. Employment Fluctuations with Equilibrium Wage Stickiness. *American Economic Review*, **95**(1), 50–65.

Hall, Robert E., and Kudlyak, Marianna. 2019 (Feb.). *Job-Finding and Job-Losing: A Comprehensive Model of Heterogeneous Individual Labor-Market Dynamics*. NBER Working Papers 25625. National Bureau of Economic Research, Inc.

Hall, Robert E., and Schulhofer-Wohl, Sam. 2018. Measuring Job-Finding Rates and Matching Efficiency with Heterogeneous Job-Seekers. *American Economic Journal: Macroeconomics*, **10**(1), 1–32.

Harrison, Bennett, and Sum, Andrew. 1979. The Theory of "Dual" or Segmented Labor Markets. *Journal of Economic Issues*, **13**(3), 687–706.

Hill, Herbert. 1996. The Problem of Race in American Labor History. *Reviews in American History*, **24**(2), 189–208.

Hornstein, Andreas. 2012 (11). *Accounting for unemployment: the long and short of it*. Working Paper 12-07. Federal Reserve Bank of Richmond.

Hughes, James P., Guttorp, Peter, and Charles, Stephen P. 1999. A Non-Homogeneous Hidden Markov Model for Precipitation Occurrence. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, **48**(1), 15–30.

Hyatt, Henry, and Spletzer, James. 2016. The Shifting Job Tenure Distribution. *Labour Economics*, **41**(C), 363–377.

Jovanovic, Boyan. 1979. Job Matching and the Theory of Turnover. *Journal of Political Economy*, **87**(5), 972–990.

Karahan, Fatih, Ozkan, Serdar, and Song, Jae. 2023. Anatomy of Lifetime Earnings Inequality: Heterogeneity in Job Ladder Risk vs. Human Capital. *Journal of Political Economy Macroeconomics*.

Kroft, Kory, Lange, Fabian, Notowidigdo, Matthew J., and Katz, Lawrence F. 2016. Long-Term Unemployment and the Great Recession: The Role of Composition, Duration Dependence, and Nonparticipation. *Journal of Labor Economics*, **34**(S1), S7–S54.

Krueger, Alan B., and Summers, Lawrence H. 1988. Efficiency Wages and the Inter-Industry Wage Structure. *Econometrica*, **56**(2), 259–293.

Krusell, Per, Mukoyama, Toshihiko, and Şahin, Ayşegül. 2010. Labour-Market Matching with Precautionary Savings and Aggregate Fluctuations. *The Review of Economic Studies*, **77**(4), 1477–1507.

Krusell, Per, Mukoyama, Toshihiko, Rogerson, Richard, and Şahin, Ayşegül. 2017. Gross Worker Flows over the Business Cycle. *American Economic Review*, **107**(11), 3447–3476.

Kudlyak, Marianna, and Lange, Fabian. 2017 (Sept.). *Measuring Heterogeneity in Job Finding Rates among the Non-Employed Using Labor Force Status Histories*. Working Paper Series 2017-20. Federal Reserve Bank of San Francisco.

Levinson, S. E., Rabiner, L. R., and Sondhi, M. M. 1983. An introduction to the application of the theory of probabilistic functions of a Markov process to automatic speech recognition. *The Bell System Technical Journal*, **62**(4), 1035–1074.

Marston, Stephen T. 1976. Employment Instability and High Unemployment Rates. *Brookings Papers on Economic Activity*, **1976**(1), 169–210.

Maruotti, Antonello. 2011. Mixed Hidden Markov Models for Longitudinal Data: An Overview. *International Statistical Review*, **79**(3), 427–454.

McNabb, Robert, and Psacharopoulos, George. 1981. Further Evidence of the Relevance of the Dual Labor Market Hypothesis for the U. K. *The Journal of Human Resources*, **16**(3), 442–448.

Menzio, Guido, and Shi, Shouyong. 2011. Efficient Search on the Job and the Business Cycle. *Journal of Political Economy*, **119**(3), 468–510.

Mincer, Jacob A. 1974. *The Human Capital Earnings Function*. NBER. Pages 83–96.

Morchio, Iacopo. 2020. Work Histories and Lifetime Unemployment. *International Economic Review*, **61**(1), 321–350.

Mortensen, Dale T., and Pissarides, Christopher A. 1994. Job Creation and Job Destruction in the Theory of Unemployment. *The Review of Economic Studies*, **61**(3), 397–415.

Moscarini, Giuseppe, and Postel-Vinay, Fabien. 2022. *The Job Ladder: Inflation vs. Reallocation*. Mimeo. Yale University.

Mueller, Andreas I, and Spinnewijn, Johannes. 2023 (February). *The Nature of Long-Term Unemployment: Predictability, Heterogeneity and Selection*. Working Paper 30979. National Bureau of Economic Research.

Mui, Preston, and Schoefer, Benjamin. forthcoming. Reservation Raises: The Aggregate Labor Supply Curve at the Extensive Margin. *Review of Economic Studies*.

Otterpohl, Jens R. 2002. Baum-Welch Learning in Discrete Hidden Markov Models with Linear Factorial Constraints. *Pages 1180–1185 of:* Dorronsoro, José R. (ed), *Artificial Neural Networks — ICANN 2002*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Petrie, T. 1969. Probabilistic Functions of Finite State Markov Chains. *The Annals of Mathematical Statistics*, **40**(1), 97 – 115.

Piore, Michael J. 1970. The Dual Labor Market: Theory and Implications. *Pages 55–59 of:* Beer, Samuel H., and Barringer, Richard E. (eds), *The State and the Poor*.

Pissarides, Christopher. 1985. Short-run Equilibrium Dynamics of Unemployment Vacancies, and Real Wages. *American Economic Review*, **75**(4), 676–90.

Powell, Jerome H. 2021. *Getting Back to a Strong Labor Market.*

Pries, Michael J. 2004. Persistence of Employment Fluctuations: A Model of Recurring Job Loss. *Review of Economic Studies*, **71**(1), 193–215.

Pries, Michael J, and Rogerson, Richard. 2021. Declining Worker Turnover: The Role of Short Duration Employment Spells. *American Economic Journal: Macroeconomics.*

Reich, Michael, Gordon, David M., and Edwards, Richard C. 1973. A Theory of Labor Market Segmentation. *The American Economic Review*, **63**(2), 359–365.

Saint-Paul, Gilles. 1997. *Dual Labor Markets: A Macroeconomic Perspective.* MIT Press Books, vol. 1, no. 0262193760. The MIT Press.

Sedláček, Petr. 2016. The Aggregate Matching Function and Job Search from Employment and Out of the Labor Force. *Review of Economic Dynamics*, **21**(July), 16–28.

Shibata, Ippei. 2019 (Dec.). *Labor Market Dynamics: A Hidden Markov Approach.* IMF Working Papers 2019/282. International Monetary Fund.

Shimer, Robert. 1999. Why is the U.S. Unemployment Rate So Much Lower? *Pages 11–74 of: NBER Macroeconomics Annual 1998, volume 13.* NBER Chapters. National Bureau of Economic Research, Inc.

Shimer, Robert. 2005. The Cyclical Behavior of Equilibrium Unemployment and Vacancies. *American Economic Review*, **95**(1), 25–49.

Shimer, Robert. 2012. Reassessing the ins and outs of unemployment. *Review of Economic Dynamics*, **15**(2), 127 – 148.

Stock, James H., and Watson, Mark W. 2001. Vector Autoregressions. *Journal of Economic Perspectives*, **15**(4), 101–115.

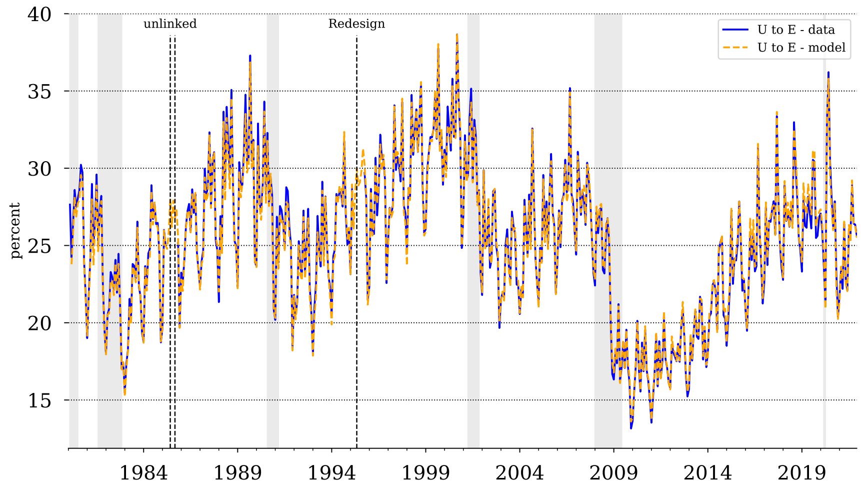
van den Berg, Gerard J., and van Ours, Jan C. 1996. Unemployment Dynamics and Duration Dependence. *Journal of Labor Economics*, **14**(1), 100–125.

Veracierto, Marcelo. 2008. On the Cyclical Behavior of Employment, Unemployment and Labor Force Participation. *Journal of Monetary Economics*, **55**(6), 1143–1157.

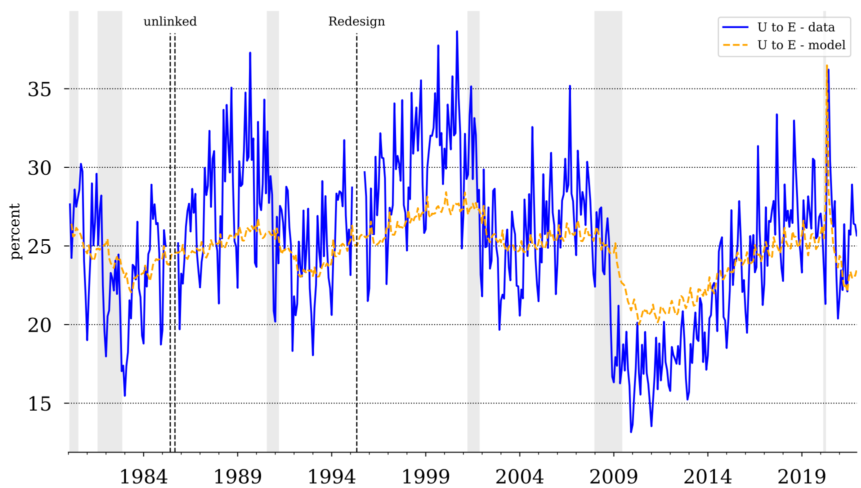
Wachter, Michael L. 1974. Primary and Secondary Labor Markets: A Critique of the Dual Approach. *Brookings Papers on Economic Activity*, **5**(3), 637–694.

Welch, Lloyd R. 2003. Hidden Markov Models and the Baum-Welch Algorithm. *IEEE Information Theory Society Newsletter*, **53**(4), 1,10–23.

Yellen, Janet. 2014. *Labor Market Dynamics and Monetary Policy: a speech at the Federal Reserve Bank of Kansas City Economic Symposium, Jackson Hole, Wyoming, August 22, 2014.* Tech. rept. Board of Governors of the Federal Reserve System (US).



(a) Time-varying transition and emission probabilities (benchmark)

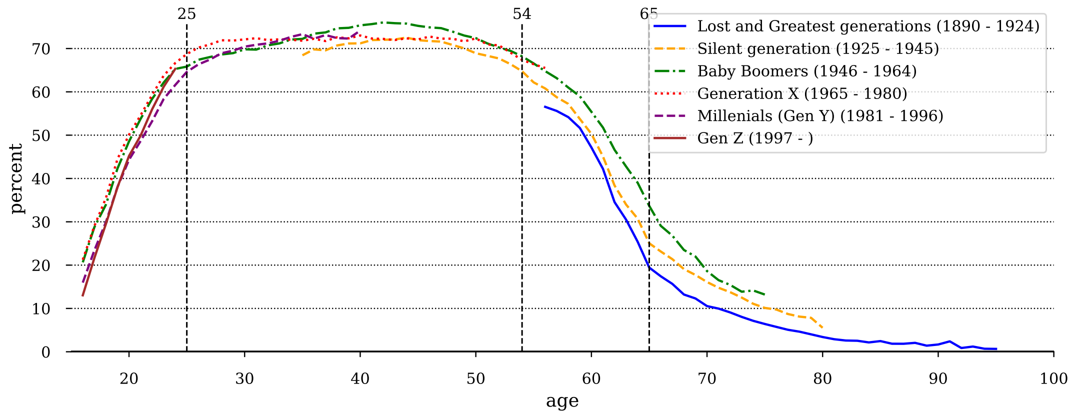


(b) Constant transition and emission probabilities

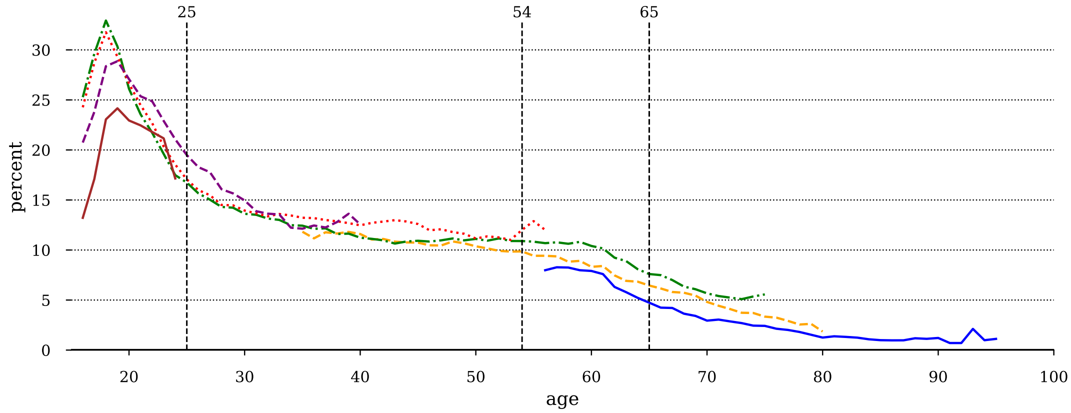
Figure 7: Monthly actual and estimated job-finding rates (EU probability).

Source: CPS and authors' calculations.

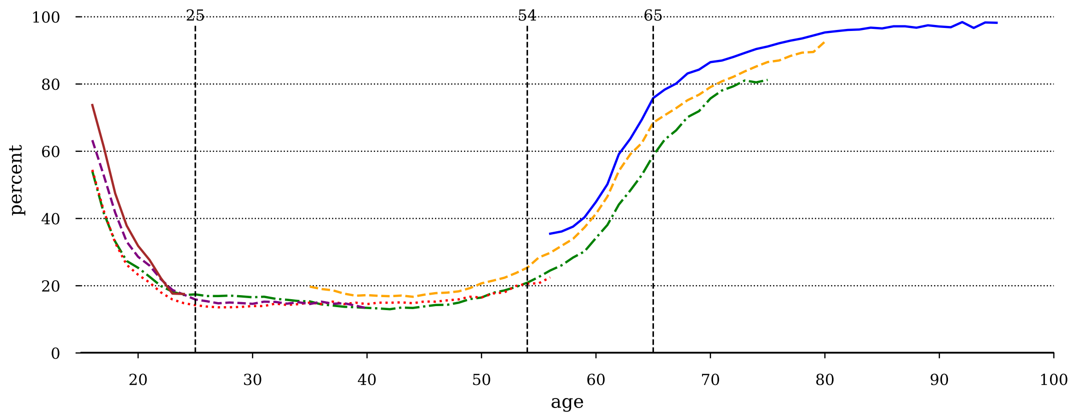
Notes: Model with constant transition and emission probabilities allows for a change in the emission probabilities in 1994 to take into account the 1994 CPS redesign.



(a) Primary sector



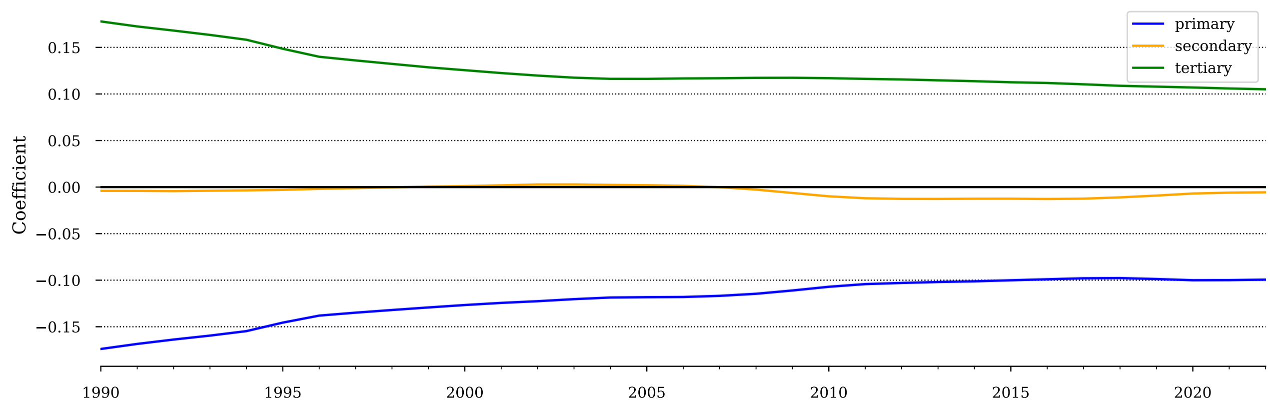
(b) Secondary sector



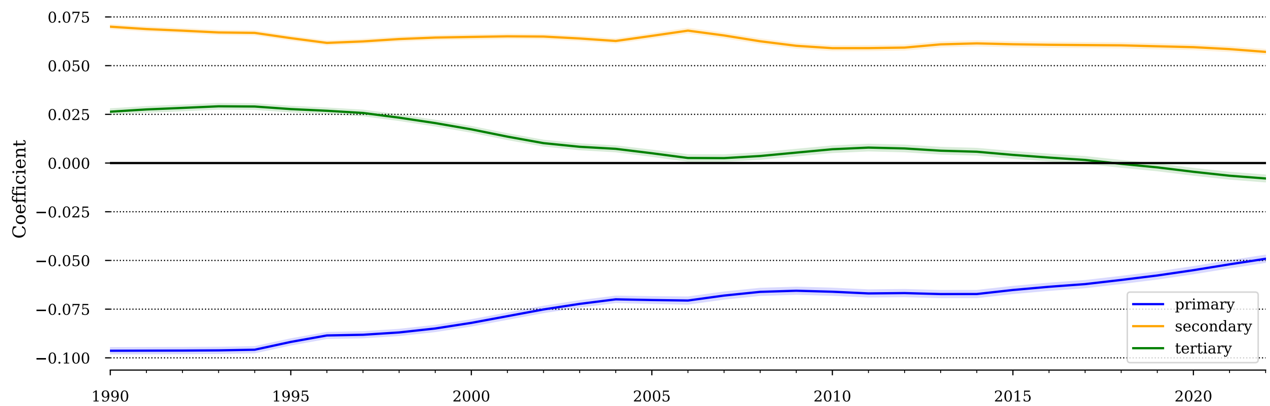
(c) Tertiary sector

Figure 8: Segment share by cohort as a function of age

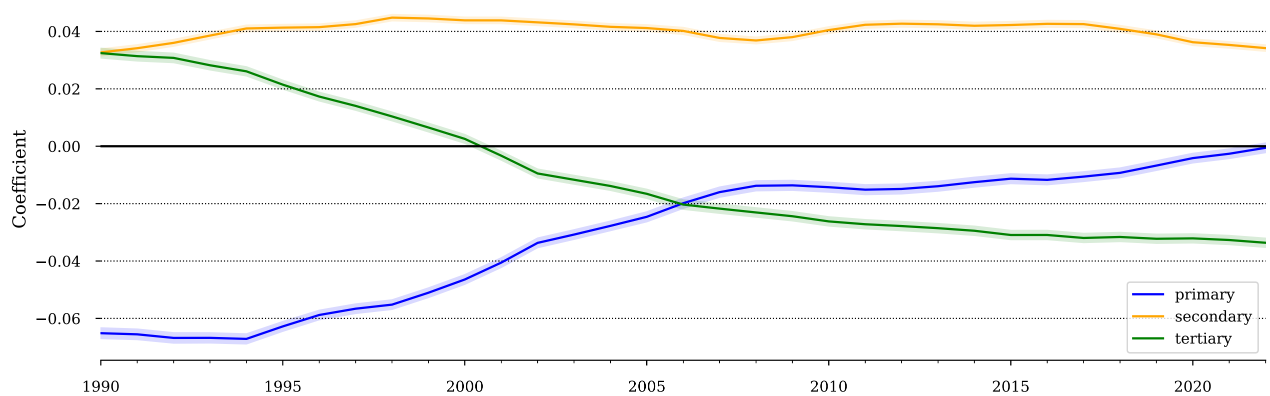
Source: CPS and authors' calculations.



(a) Sex: Female



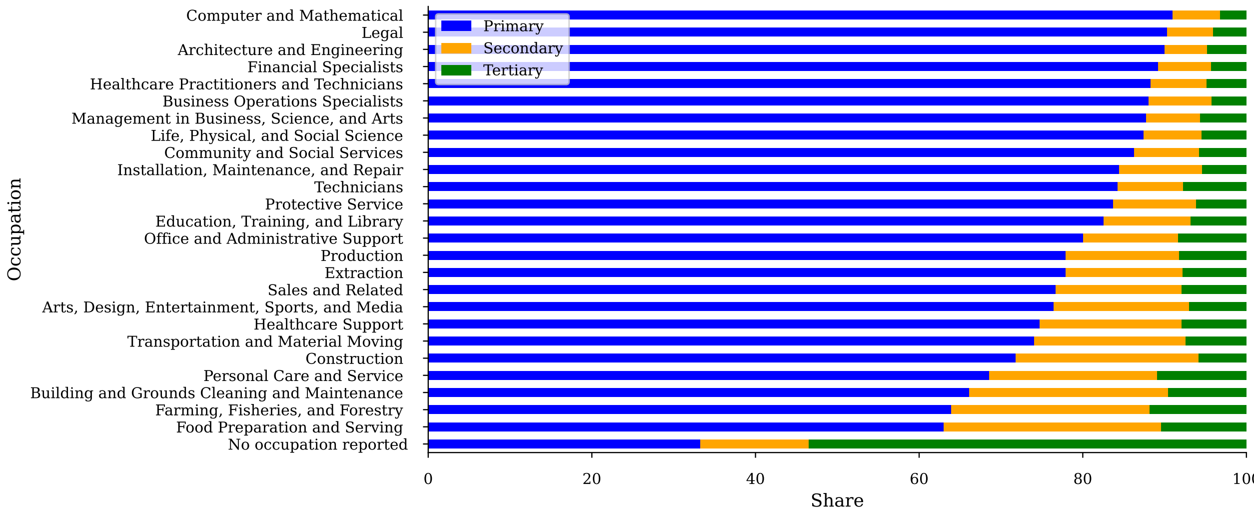
(b) Race: Black



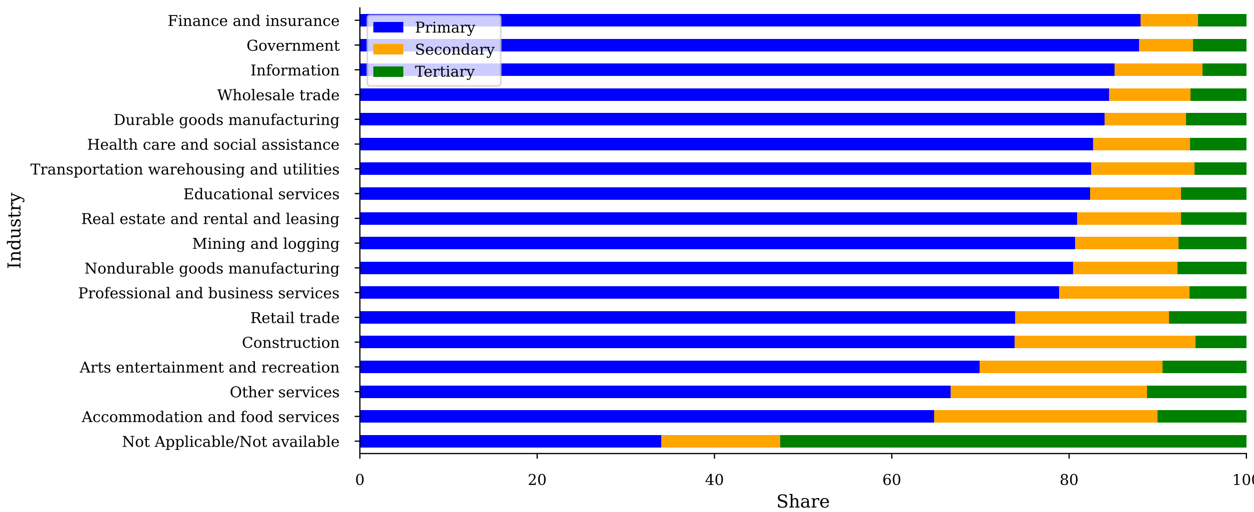
(c) Ethnicity: Hispanic

Figure 9: Evolution of regression coefficients for posterior probabilities (sex, race and ethnicity).

Source: CPS and authors' calculations.
Notes: Regression coefficients for annual 10-year rolling samples for same specification as in Table 12.



(a) Occupation



(b) Industry

Figure 10: Segment distribution by industry and occupation

Source: CPS and authors' calculations.

Notes: Industries are based 2-digit NAICS codes and occupations on 2-digit 2010 SOC codes.

A Mathematical and computational details

E-step: Conditional expectation of the complete-data log-likelihood

The conditional expectation of the complete-data log likelihood function can be derived by considering the expectations of $u_{i,t,l}$ and $v_{i,t,l,l'}$. For the first one, we obtain

$$\begin{aligned}\hat{u}_{i,t,l} &= E\left[u_{i,t,l} \mid \boldsymbol{x}_i;\boldsymbol{\theta}\right] = P\left(\ell_{i,t} = l \mid \boldsymbol{x}_i;\boldsymbol{\theta}\right) \\ &= \frac{P\left(\ell_{i,t} = l \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right)}{P\left(\boldsymbol{x}_i;\boldsymbol{\theta}\right)} = \frac{P\left(\ell_{i,t} = l \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right)}{\sum_{l'\in L} P\left(\ell_{i,t} = l' \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right)}.\end{aligned} \tag{20}$$

Similarly

$$\begin{aligned}\hat{v}_{i,t,l,l'} &= E\left[v_{i,t,l,l'} \mid \boldsymbol{x}_i;\boldsymbol{\theta}\right] = P\left(\ell_{i,t-1} = l \cap \ell_{i,t} = l' \mid \boldsymbol{x}_i;\boldsymbol{\theta}\right) \\ &= \frac{P\left(\ell_{i,t-1} = l \cap \ell_{i,t} = l' \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right)}{P\left(\boldsymbol{x}_i;\boldsymbol{\theta}\right)} \\ &= \frac{P\left(\ell_{i,t-1} = l \cap \ell_{i,t} = l' \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right)}{\sum_{h'\in L}\sum_{h\in L} P\left(\ell_{i,t-1} = h \cap \ell_{i,t} = h' \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right)}.\end{aligned} \tag{21}$$

Here, we can express

$$\begin{aligned}P\left(\ell_{i,t_i+k} = l \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right) &= P\left(x_{i,t_i},..,x_{i,t_i+k} \cap \ell_{i,t_i+k} = l\right) \\ &\quad \times P\left(x_{i,t_i+k+1},..,x_{i,t_i+15} \mid \ell_{i,t_i+k} = l\right) \\ &= \alpha_{i,k}\left(l\right)\beta_{i,k}\left(l\right),\end{aligned} \tag{22}$$

where $\alpha_{i,k}\left(l\right)$ is as defined in the main text and

$$\beta_{i,k}\left(l\right) = P\left(x_{i,t_i+k+1},..,x_{i,t_i+15} \mid \ell_{i,t_i+k} = l\right). \tag{23}$$

Moreover

$$\begin{aligned}P\left(\ell_{i,t_i+k-1} = l \cap \ell_{i,t_i+k} = l' \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right) &= P\left(x_{i,t_i},..,x_{i,t_i+k-1} \cap \ell_{i,t_i+k-1} = l\right) \\ &\quad \times q_{t_i+k,l,l'}\omega_{x_{i,t_i+k},l',t_+k} \\ &\quad \times P\left(x_{i,t_i+k+1},..,x_{i,t_i+15} \mid \ell_{i,t_i+k} = l'\right) \\ &= \alpha_{i,k-1}\left(l\right) q_{t_i+k,l,l'}\omega_{x_{i,t_i+k},l',t_+k}\beta_{i,k}\left(l'\right)\end{aligned} \tag{24}$$

This yields that

$$\hat{u}_{i,t_i+k,l} = \frac{\alpha_{i,k}(l)\,\beta_{i,k}(l)}{\sum_{l'\in L}\alpha_{i,k}(l')\,\beta_{i,k}(l')} \tag{25}$$

and

$$\hat{v}_{i,t_i+k,l,l'} = \frac{\alpha_{i,k-1}(l)\,q_{t_i+k,l,l'}\omega_{x_{i,t_i+k},l',t+k}\beta_{i,k}(l')}{\sum_{h'\in L}\sum_{h\in L}\alpha_{i,k-1}(l)\,q_{t_i+k,h,h'}\omega_{x_{i,t_i+k},h',t+k}\beta_{i,k}(l')}. \tag{26}$$

Just like $\alpha_{i,k}(l)$, $\beta_{i,k}(l)$ can be calculated using a recursion.

$$\begin{aligned}
\beta_{i,15}(l) &= 1, \text{ and} && (27)\\
\beta_{i,k}(l) &= P\left(x_{i,t_i+k+1},..,x_{i,t_i+15} \mid \ell_{i,t_i+k}=l\right) && (28)\\
&= \sum_{l'} q_{t_i+k+1,l,l'}\beta_{i,k+1}(l') && (29)\\
&\quad\times\left[(1-\eta_{i,t_i+k+1})+\eta_{i,t_i+k+1}\omega_{x_{i,t_i+k+1},l',t_i+k+1}\right], \text{ for } k=0,\ldots,14 && (30)
\end{aligned}$$

is the backward recursion that is part of the Forward-Backward method (BW).

Property of posterior probabilities

Let $\{\mathcal{P},\mathcal{S},\mathcal{T}\}$ be the sets of hidden labor market states that are part of the primary and secondary tiers respectively. If there is no mobility between these tiers then it must be the case that for $\mathcal{M}\in\{\mathcal{P},\mathcal{S},\mathcal{T}\}$:

$$\begin{aligned}
P\left(\boldsymbol{x}_i\cap\ell_{i,t}\in\mathcal{M}\right) &= \sum_{l\in\mathcal{T}} P\left(\boldsymbol{x}_i\cap\ell_{i,t}=l\right) && (31)\\
&= \sum_{l\in\mathcal{M}}\sum_{l'\in\mathcal{M}} P\left(\ell_{i,t}=l \mid \ell_{i,t-1}=l'\right)P\left(\boldsymbol{x}_i\cap\ell_{i,t-1}=l'\right) && (32)\\
&= \sum_{l'\in\mathcal{M}}\sum_{l\in\mathcal{M}} P\left(\ell_{i,t}=l \mid \ell_{i,t-1}=l'\right)P\left(\boldsymbol{x}_i\cap\ell_{i,t-1}=l'\right) && (33)\\
&= \sum_{l'\in\mathcal{M}} P\left(\boldsymbol{x}_i\cap\ell_{i,t-1}=l'\right) && (34)\\
&= P\left(\boldsymbol{x}_i\cap\ell_{i,t-1}\in\mathcal{M}\right) && (35)
\end{aligned}$$

Thus, the posterior probability that a person is in a particular segment of the labor market is constant over time when there is no mobility across the labor market tiers.

M-step: Updated parameter estimates

In the M-step, the parameters, δ_{l,t_i}, $q_{t_i+k,l,l'}$, and $\omega_{x_{i,t_i+k},l,t_i+k}$, are chosen to maximize

$$\begin{aligned}\ln \mathcal{L} &= \sum_{i=1}^{n} w_i \left\{ \sum_{l\in L} \hat{u}_{i,t_i,l} \ln \delta_{l,t_i} + \sum_{k=1}^{15} \sum_{l'\in L} \sum_{l\in L} \hat{v}_{i,t_i+k,l,l'} \ln q_{t_i+k,l,l'} \right. \\ &\quad \left. + \sum_{k=0}^{15} \eta_{i,t_i+k} \sum_{l\in L} \hat{u}_{i,t_i+k,l} \ln \omega_{x_{i,t_i+k},l,t_i+k} \right\}.\end{aligned} \tag{36}$$

subject to the adding-up constraints

$$\sum_{l} \delta_{l,t} = 1, \text{ for } t = 1, \ldots, T \tag{37}$$

$$\sum_{l}^{\prime} q_{t,l,l'} = 1, \text{ for } t = 1, \ldots, T \text{ and } l \in L, \text{ and} \tag{38}$$

$$\sum_{x\in X} \omega_{x_{i,t_i+k},l,t_i+k} = 1, \text{ for } t = 1, \ldots, T \text{ and } l \in L \tag{39}$$

as well as the additional (in-)equality restrictions we described in Subsection 3.2.

Without the additional identifying (in-)equality constraints, the above maximization problem has a closed-form solution derived in Baum *et al.* (1970); Welch (2003). The implementation of the BW with parameter constraints has been studied extensively (most notably Levinson *et al.*, 1983; Otterpohl, 2002). Under some types of constraints the M-step yields closed-form solutions. But that is not the case for our application. Instead, we rely on numerical methods to maximize the expected complete-data likelihood.

We exploit that the identifying restrictions we impose have two important properties. The first is that they are all contemporaneous in that they impose restrictions on parameters at the same point in time. The second is that they are separated between transition probabilities, $q_{t,l,l'}$, and emission probabilities, $\omega_{x,l,t}$.

This property simplifies the M-step to $3T$ convex maximization problems. To see how this works, define the set $N(t)$ as the individuals i who are respondents in period t. Then we can write

$$\ln \mathcal{L} = \sum_{T}^{t=1} \sum_{i\in N(t)} w_i \left\{ \sum_{l\in L} \hat{u}_{i,t,l} \ln \delta_{l,t} + \sum_{l\in L} \sum_{l'\in L} \hat{v}_{i,t,l,l'} \ln q_{t,l,l'} + \sum_{l\in L} \hat{u}_{i,t,l} \ln \omega_{x_{i,},l,t} \right\}.$$

Then, for each month t the M-step involves three maximization problems. The first is to maximize

$$\sum_{l \in L} \hat{u}_{i,t,l} \ln \delta_{l,t}, \tag{40}$$

with respect to the unconditional probabilities (stocks), $\{\delta_{l,t}\}_{l \in L}$, subject to the adding-up constraint (37). This is a well-defined convex problem that solves for the Weighted Analytic Center that can be solved using the algorithm from Andersen *et al.* (2011).

The other two problems also involve solving for the Weighted Analytic Center but subject to more constraints. The transition probabilities in month t, $\{q_{l,l',t}\}_{(l,l') \in L \times L}$, in the M-step maximize

$$\sum_{l \in L} \sum_{l' \in L} \hat{v}_{i,t,l,l'} \ln q_{t,l,l'}, \tag{41}$$

subject to (38) and the identifying inequality constraints introduced in Subsection 3.2. This, again can be solved using the algorithm from Andersen *et al.* (2011). The same is true for the emission probabilities, $\{\omega_{x,l,t}\}_{(x,l) \in X \times L}$, which maximize

$$\sum_{l \in L} \hat{u}_{i,t,l} \ln \omega_{x_{i,},l,t}, \tag{42}$$

subject to (39).

Details of calculations for Table 9

Table 9 provides a decomposition of the average change, $\Delta \bar{x}_t$, the variance of the month-to-month changes, $\sigma^2 (x_t)$, and the variance of 12-month changes, $\sigma^2 (\Delta_{12} x_t)$, in four aggregate variables: (i) the unemployment rate, (ii) LFPR, (iii) EPOP, and (iv) flows per capita.

Each of these four aggregates can be written as a share-weighted, with shares $s_{M,t}$, sum of the segment-specific aggregates, $x_{M,t}$. For the unemployment rate the shares are the shares of the labor force and for the other three aggregates they are the segment-share of the population. Thus, all the aggregates in Table 9 can be written as

$$x_t = \sum_{M \in \{P,S,T\}} s_{M,t} x_{M,t}. \tag{43}$$

This allows us to use shift-share analysis to write the change in x_t as

$$\Delta x_t = \left[\sum_{M \in \{P,S,T\}} \bar{x}_{M,t} \Delta s_{M,t} \right] + \sum_{M \in \{P,S,T\}} \left[\bar{s}_{M,t} \Delta x_{M,t} \right], \tag{44}$$

where

$$\bar{x}_{M,t} = \frac{1}{2}\left(x_{M,t} + x_{M,t-1}\right), \text{ and } \bar{s}_{M,t} = \frac{1}{2}\left(s_{M,t} + s_{M,t-1}\right) \tag{45}$$

The first term of (44) is the contribution of the changes in the composition due to changes in the shares of the market segments. This is the part that corresponds to the second column of Table 9. The last three columns correspond to each segment's term in the second part of (44). It captures the impact of the shift in the sector-specific aggregate on the overall aggregate. The first column of the table corresponds to the left-hand side of (44). The lines in Table 9 for $\bar{\Delta x}_t$ report the sample averages of the terms in (44) across all months in the sample.

For the decomposition of the variance we use that the additive decomposition of Δx_t in (44) implies that

$$\text{Var}\left(\Delta x_t\right) = \text{Cov}\left(\Delta x_t, \left[\sum_{M\in\{P,S,T\}} \bar{x}_{M,t}\Delta s_{M,t}\right]\right) + \sum_{M\in\{P,S,T\}} \text{Cov}\left(\Delta x_t, \left[\bar{s}_{M,t}\Delta x_{M,t}\right]\right) \tag{46}$$

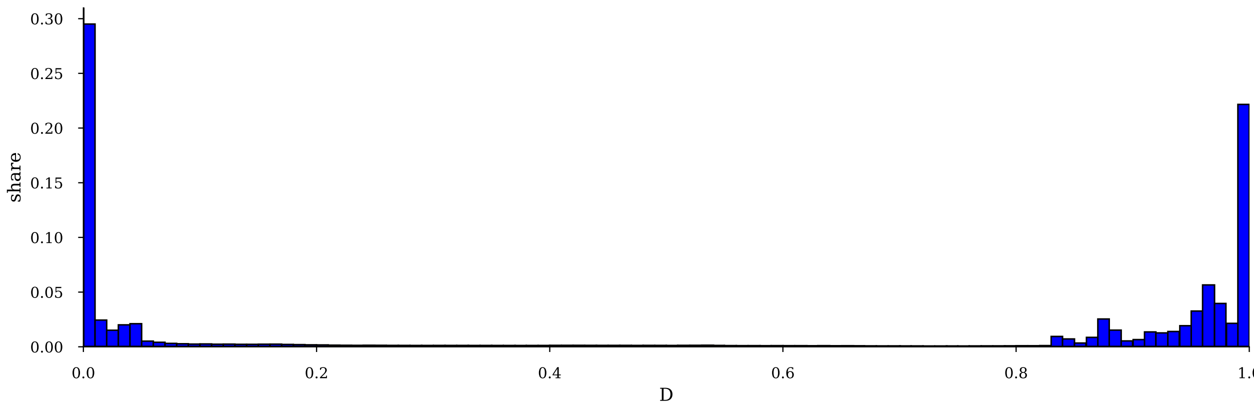
Thus, the variance can also be split into the contributions of the changes in the shares as well as the shifts in segment-specific measures. This is what is reported in the lines for $\sigma^2\left(\Delta x_t\right)$ in the table. The lines in the table for $\sigma^2\left(\Delta_{12} x_t\right)$ are calculated in a similar way, except the difference is now taken between variables at times t and $t-12$.

B Additional empirical results

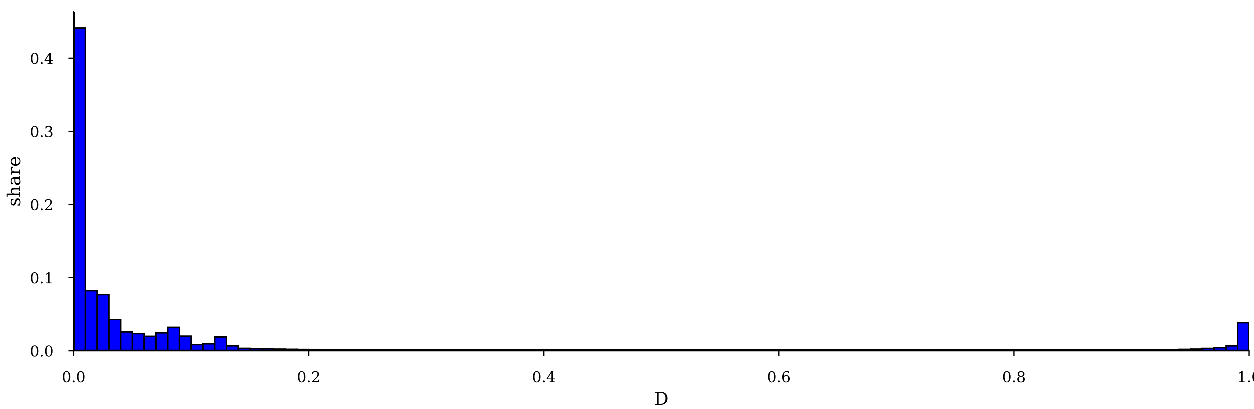
Table B.1: Volatility by labor market segment, $\sigma(x)$, for the benchmark model and the restricted model with time-invariant transition and emissions probabilities

	Primary		Secondary		Tertiary	
measure	Benchmark model	No time variation	Benchmark model	No time variation	Benchmark model	No time variation
Unemployment rate	0.52	0.47	2.58	1.39	2.48	1.09
Labor-force participation rate	0.20	0.11	1.10	0.75	0.34	0.22
Employment-to-population ratio	0.62	0.50	1.99	1.00	0.37	0.24
Flows per capita	0.06	0.03	0.13	0.05	0.02	0.01

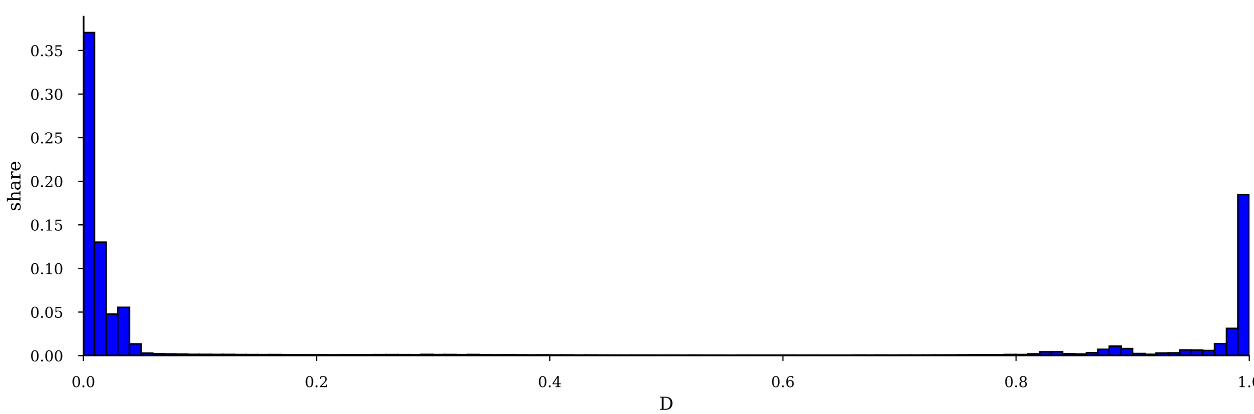
Source: Current Population Survey and authors' calculations.
$\sigma(x)$ is the standard deviation of HP-filtered cyclical gap from quarterly seasonally adjusted data for our benchmark and without time variation in transition and emissions probabilities. HP-filter applied with smoothing parameter of 1600.



(a) Primary

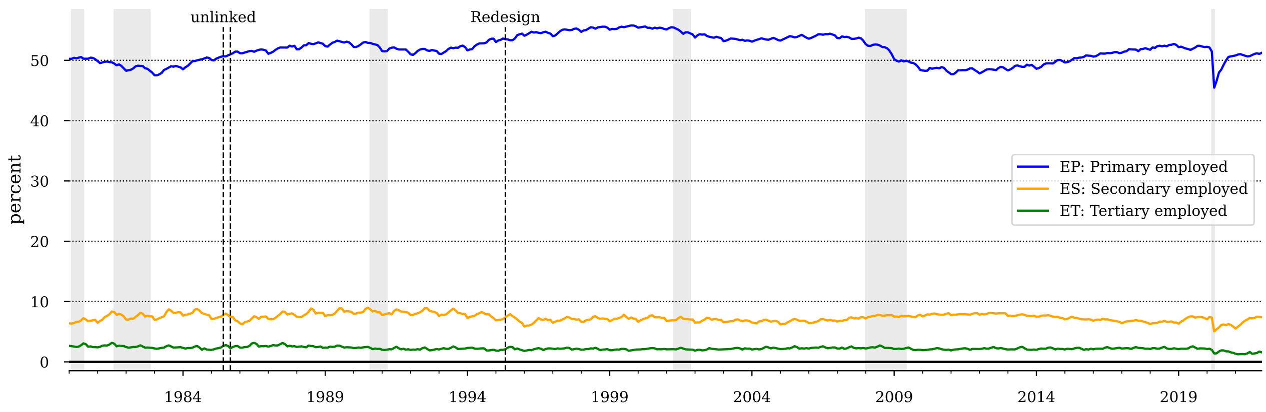


(b) Secondary

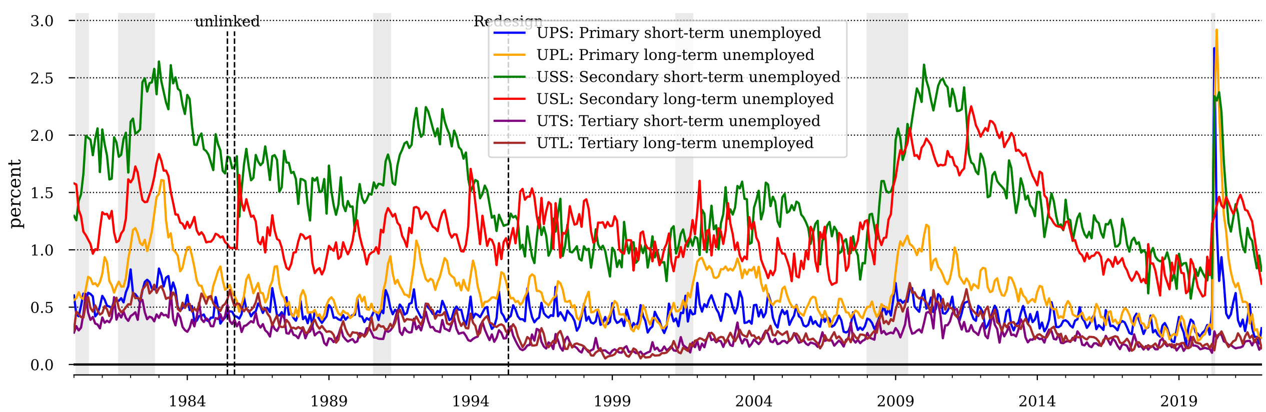


(c) Tertiary

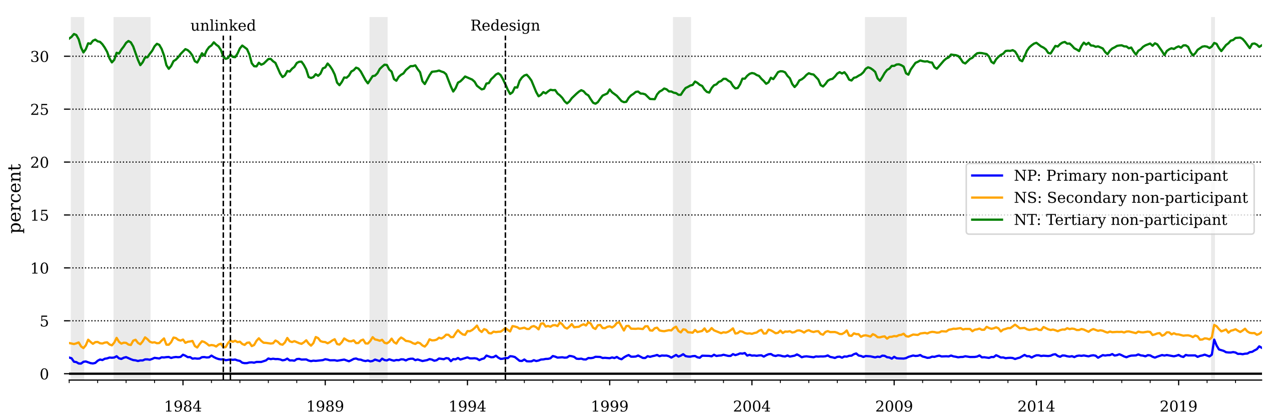
Figure B.1: Distribution of posterior probabilities by market segment (1980-2021)



(a) Employed



(b) Unemployed

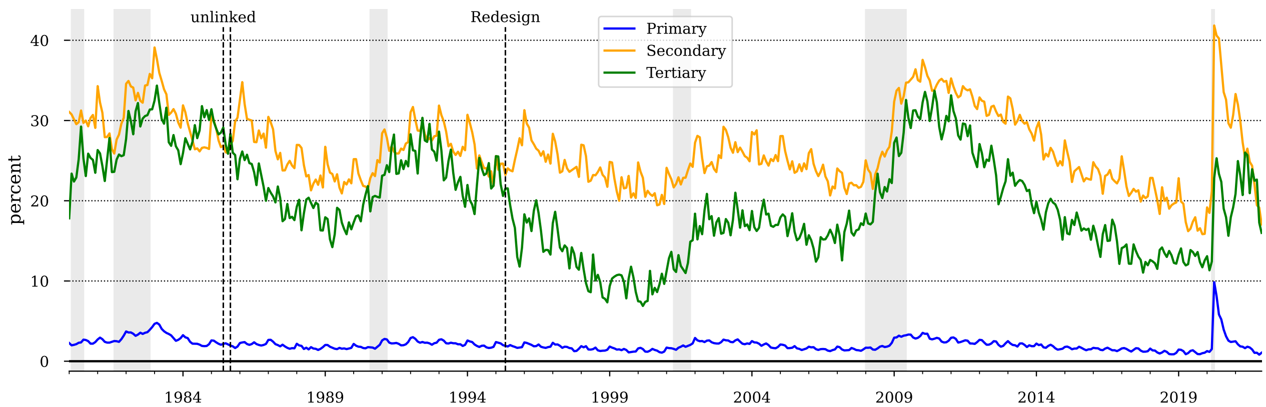


(c) Non-participant

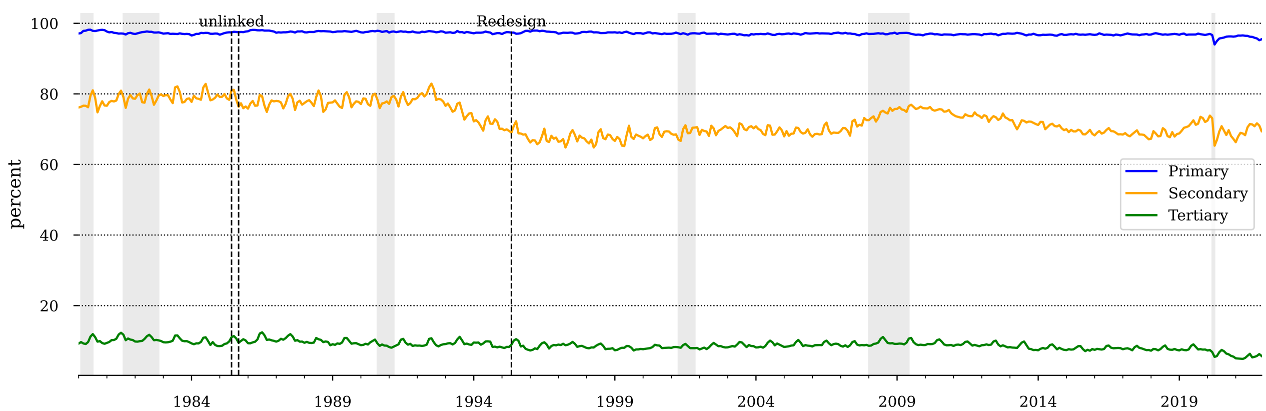
Figure B.2: Estimated shares of population in hidden states.

Source: CPS and authors' calculations.

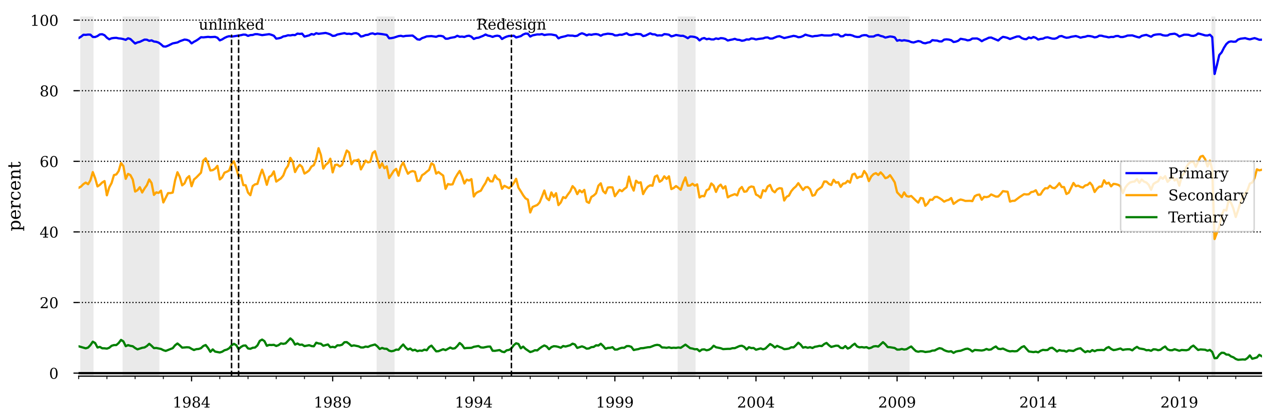
Notes: "unlinked" is period where household identifiers in CPS are scrambled and 4-8-4 respondent histories are incomplete and "redesign"is the 1994 for which respondent histories are incomplete and the questions used to construct the emissions are changed.



(a) Unemployment rates



(b) Labor force participation rates

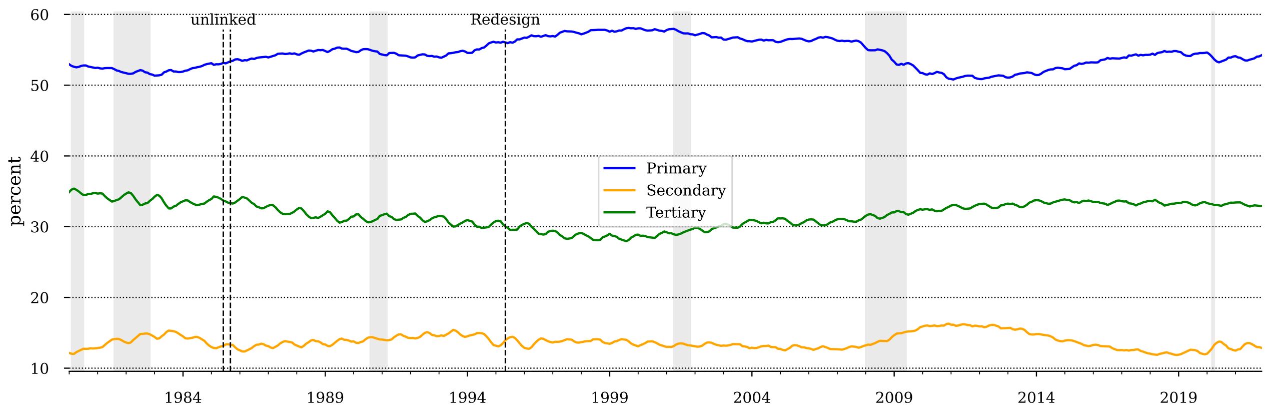


(c) Employment-population ratios

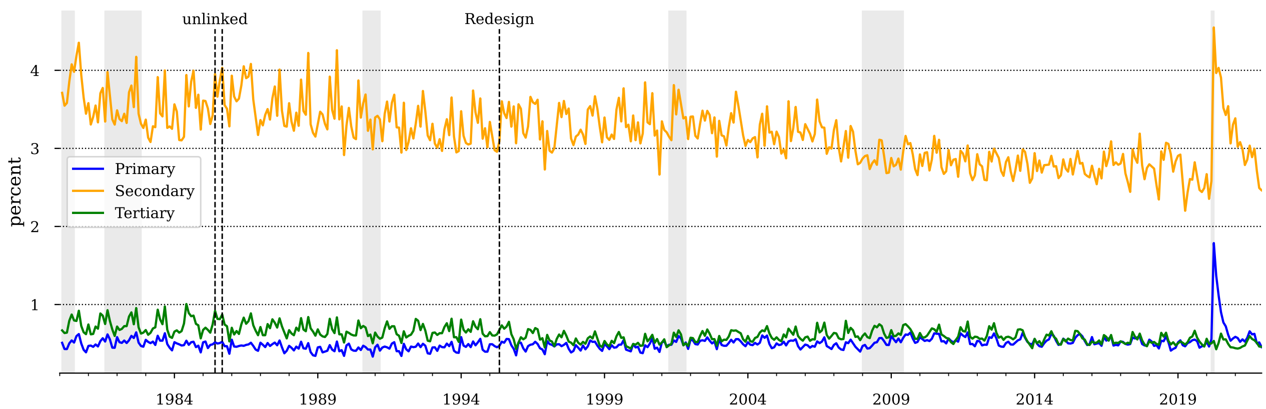
Figure B.3: Labor market statistics by tier.

Source: CPS and authors' calculations.

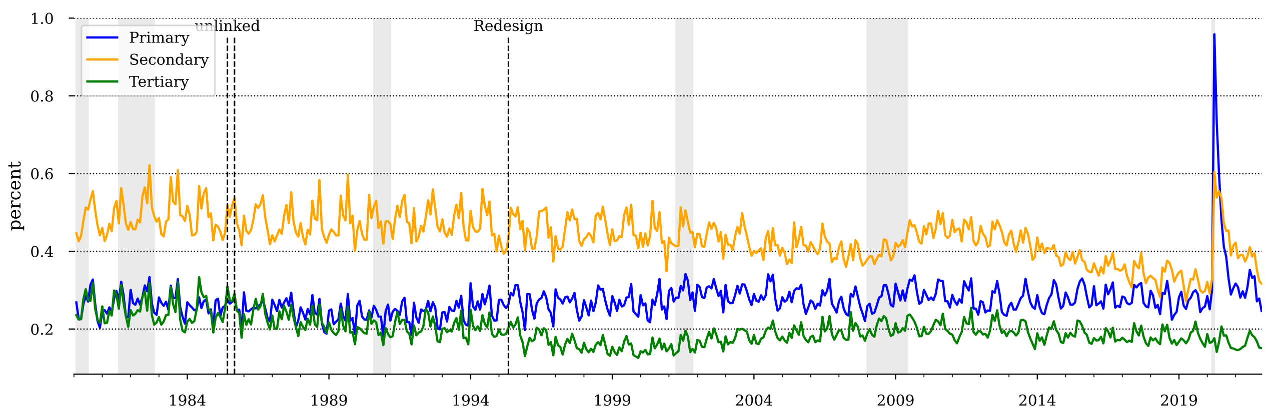
Notes: "unlinked" is period where household identifiers in CPS are scrambled and 4-8-4 respondent histories are incomplete and "redesign"is the 1994 for which respondent histories are incomplete and the questions used to construct the emissions are changed.



(a) Share of population



(b) Annualized monthly flows per person in segment



(c) Annualized monthly flows per capita

Figure B.4: Shares of population and flows per person by labor market segment.

Source: CPS and authors' calculations.

Notes: "unlinked" is period where household identifiers in CPS are scrambled and 4-8-4 respondent histories are incomplete and "redesign"is the 1994 for which respondent histories are incomplete and the questions used to construct the emissions are changed.

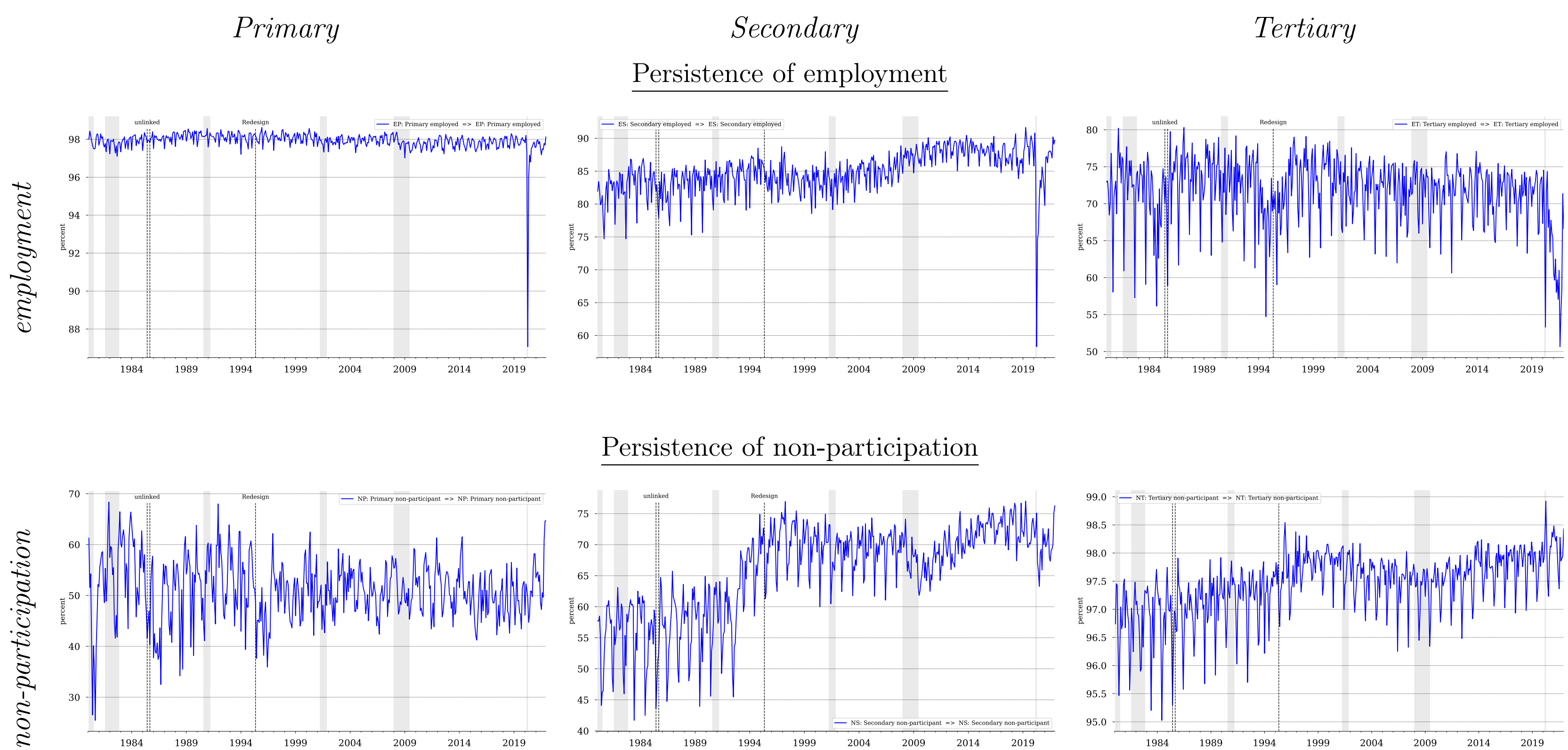


Figure B.5: Persistence of employment and non-participation

Source: CPS and authors' calculations. Monthly observations. Not seasonally adjusted. *Note:* Transition probabilities from lont-term to short-term unemployment are restricted to be zero.

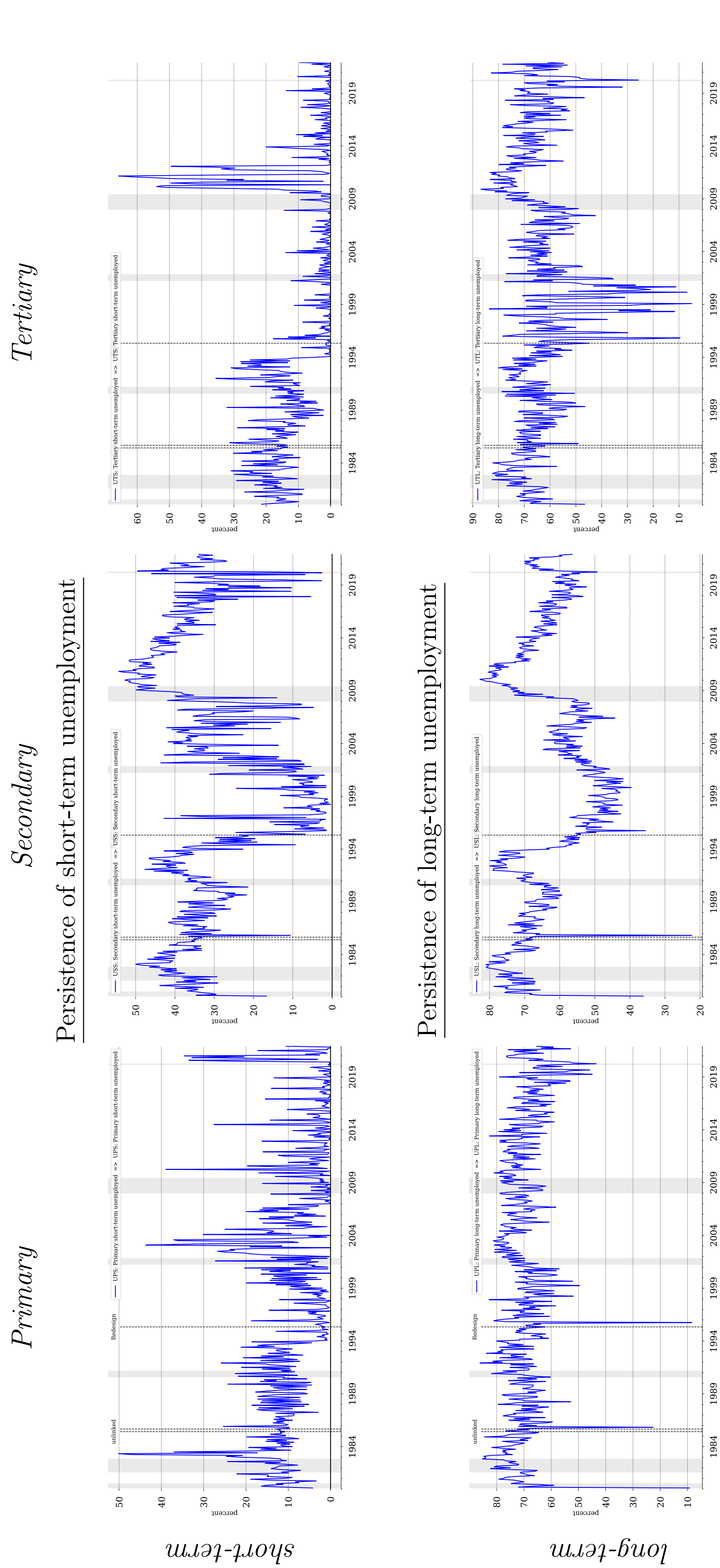


Figure B.6: Persistence of short- and long-term unemployment

Source: CPS and authors' calculations. Monthly observations. Not seasonally adjusted. *Note:* Transition probabilities from lont-term to short-term unemployment are restricted to be zero.

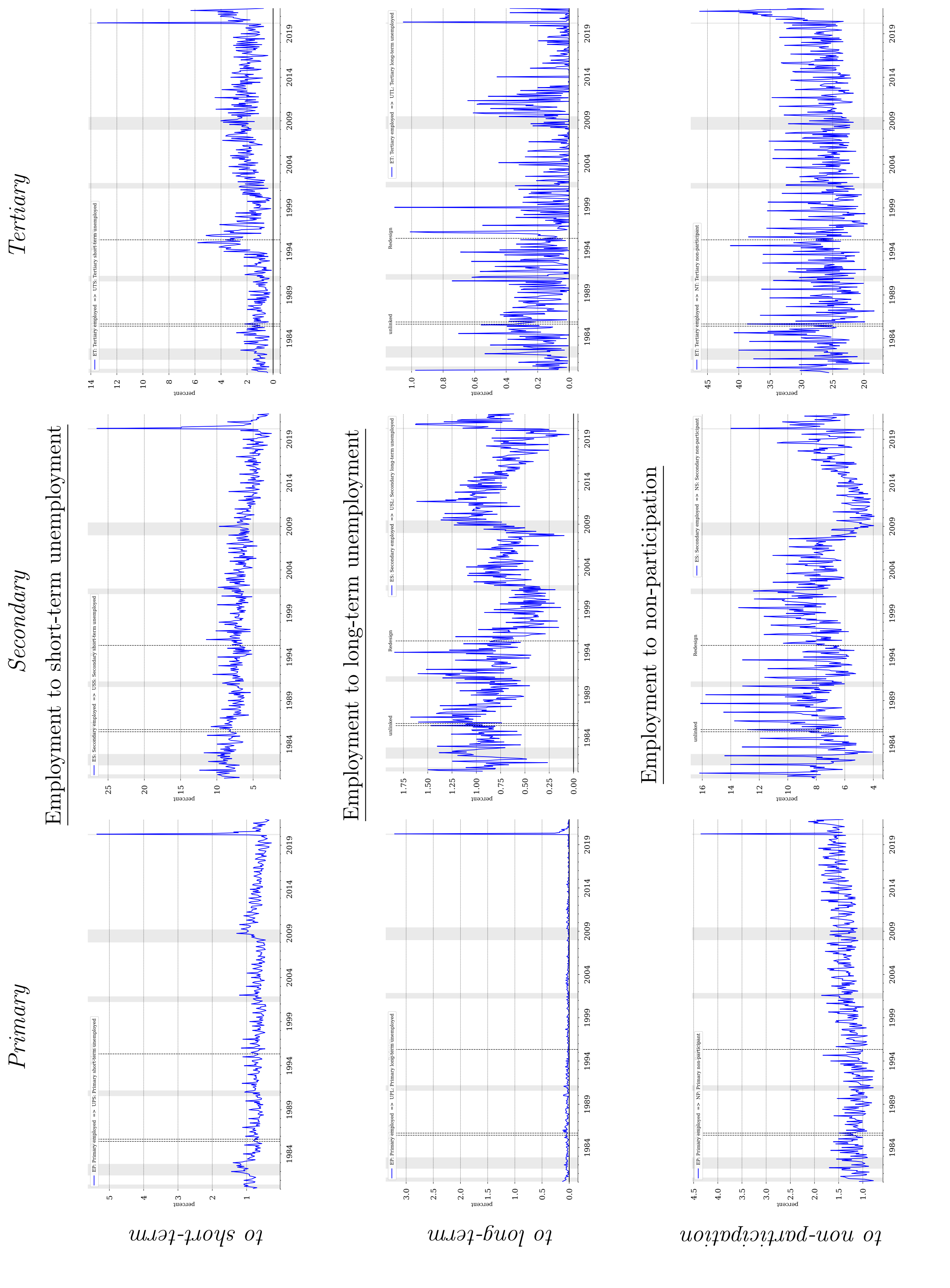


Figure B.7: Estimated transition probabilities from employment

Source: CPS and authors' calculations. Monthly observations. Not seasonally adjusted.

Primary *Secondary* *Tertiary*

Short-term unemployment to employment

to employment

Short-term unemployment to long-term unemployment

to long-term

Short-term unemployment to non-participation

to non-participation

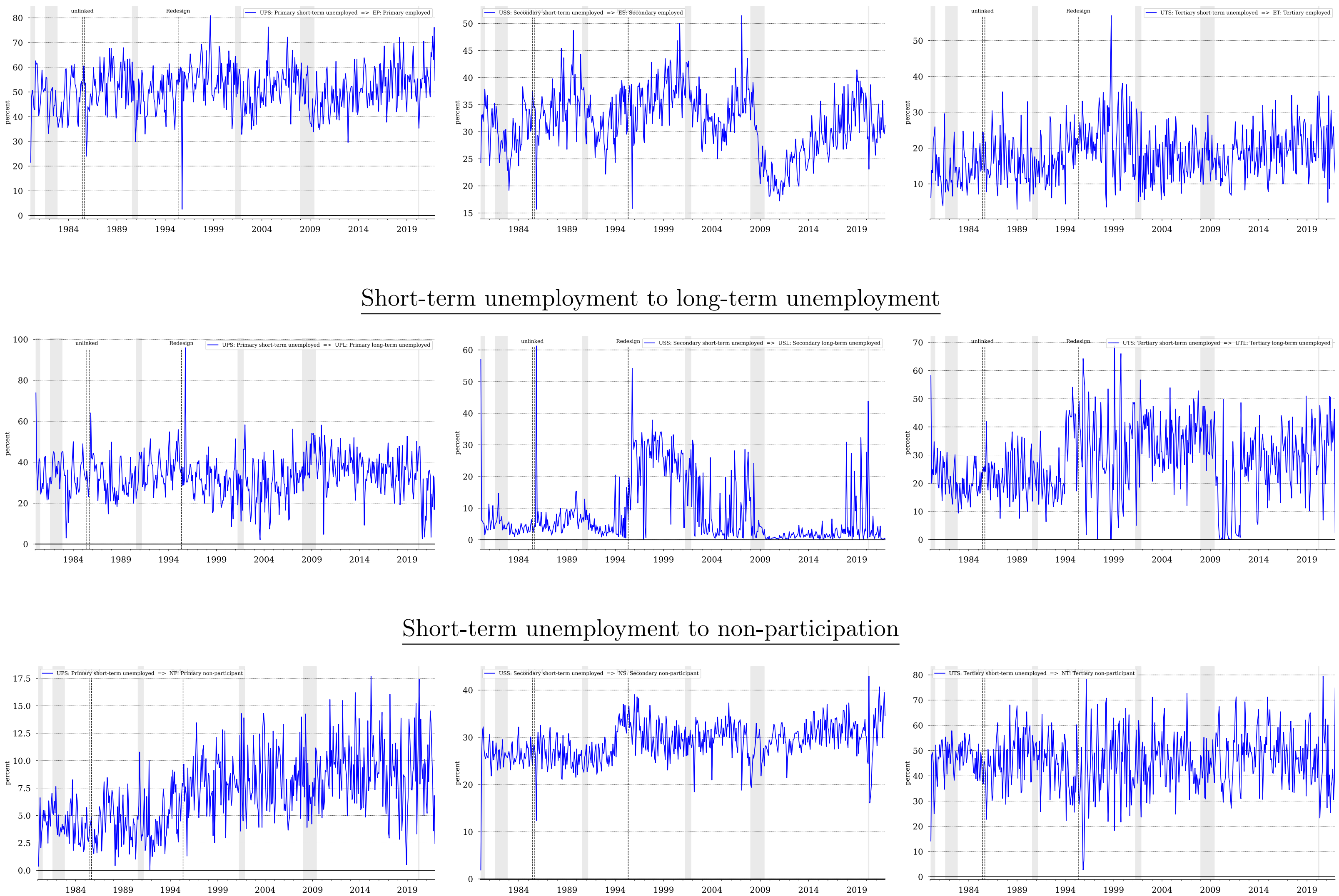


Figure B.8: Estimated transition probabilities from short-term unemployment

Source: CPS and authors' calculations. Monthly observations. Not seasonally adjusted.

Primary *Secondary* *Tertiary*

Long-term unemployment to employment

Long-term unemployment to non-participation

to employment

to non-participation

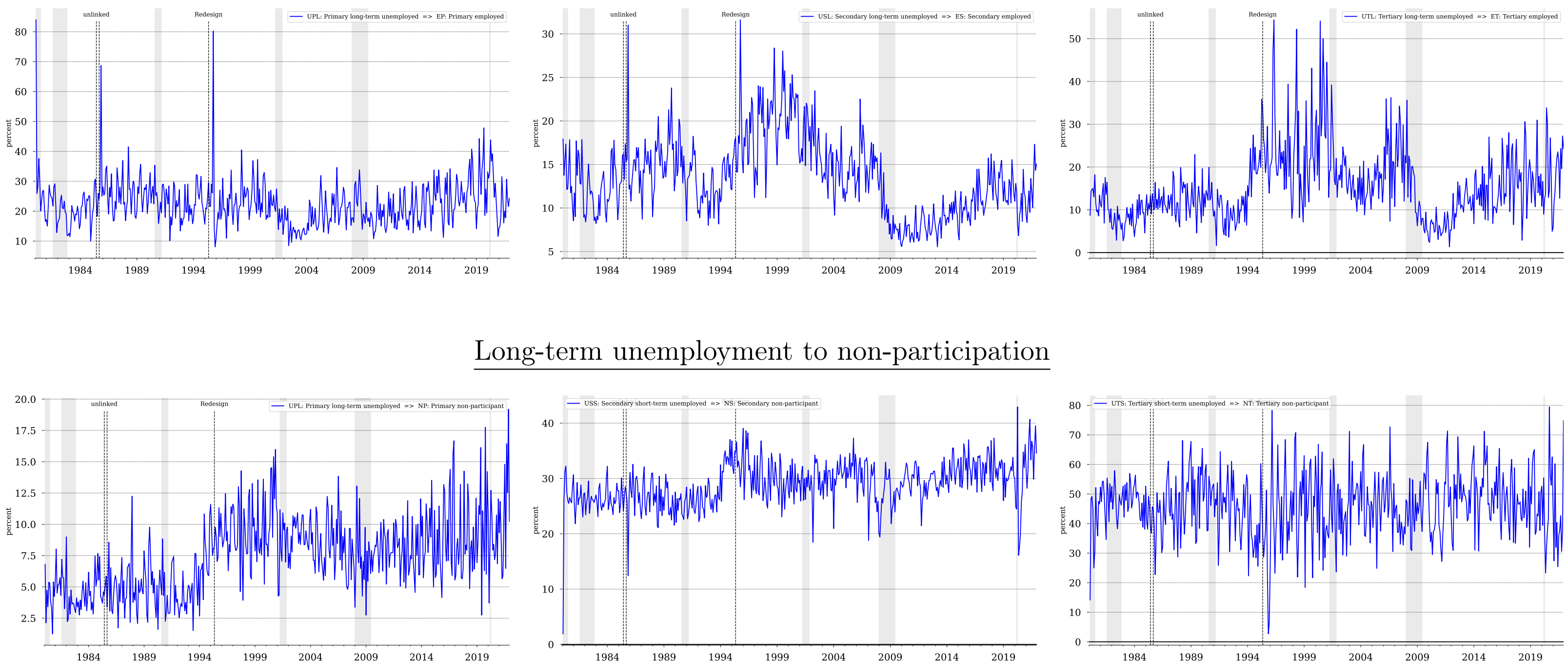


Figure B.9: Estimated transition probabilities from long-term unemployment

Source: CPS and authors' calculations. Monthly observations. Not seasonally adjusted. *Note:* Transition probabilities from lont-term to short-term unemployment are restricted to be zero.

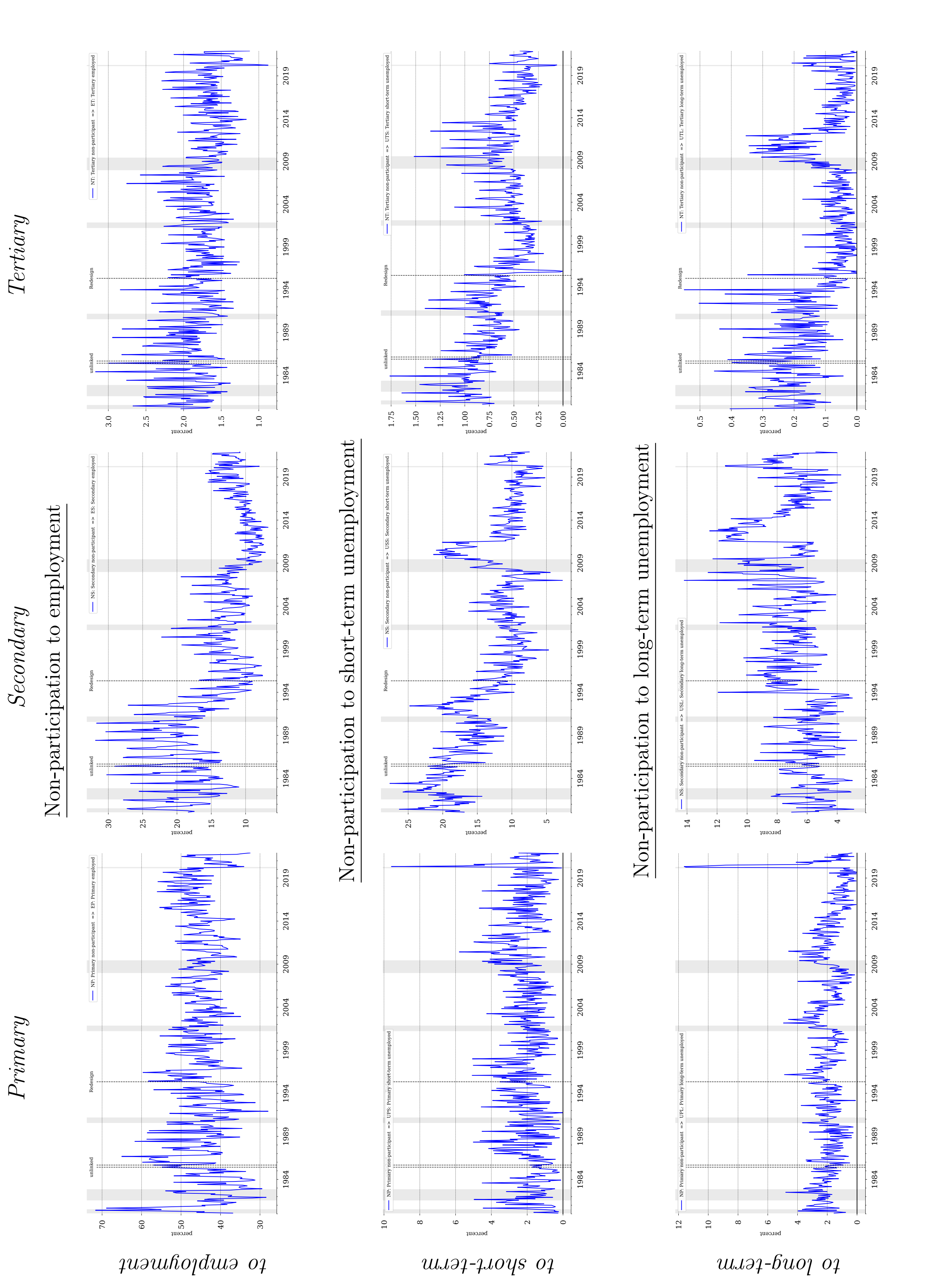
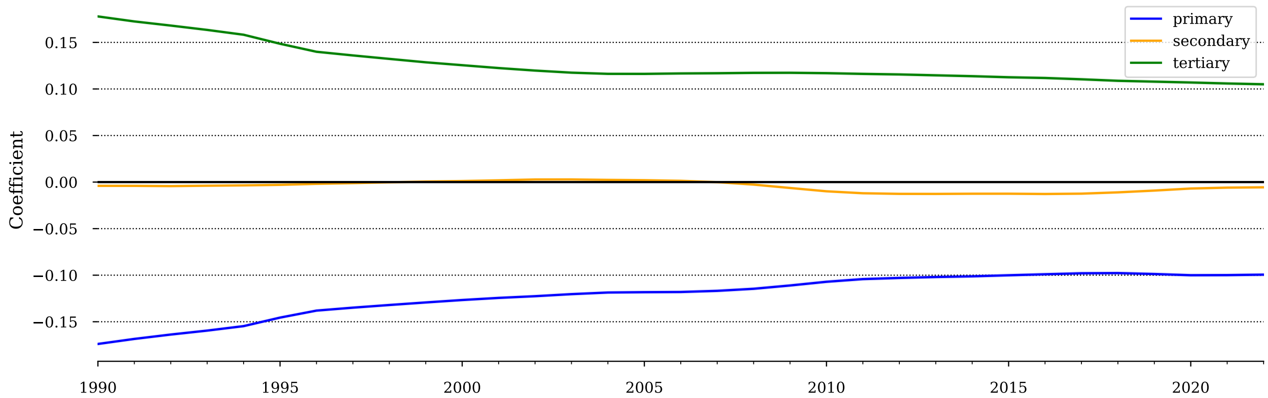
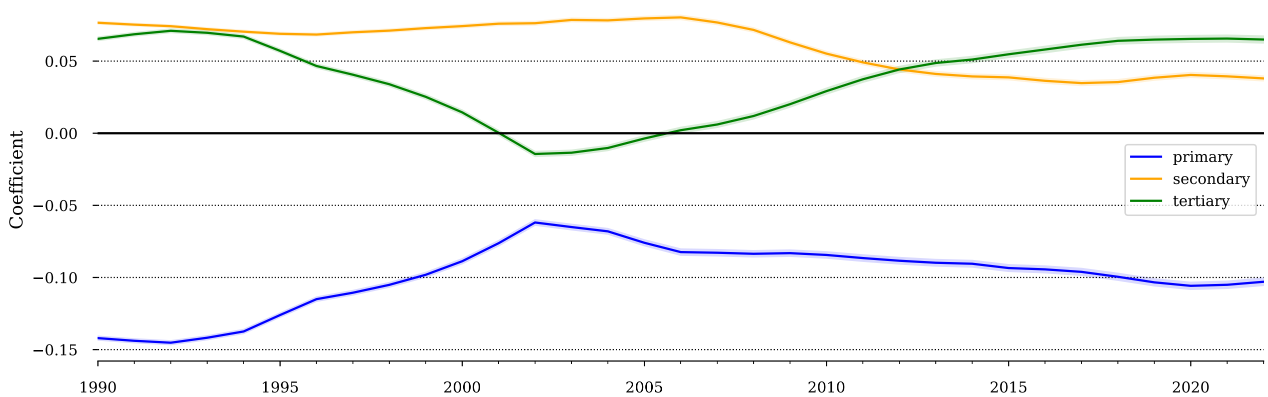


Figure B.10: Estimated transition probabilities from non-participation

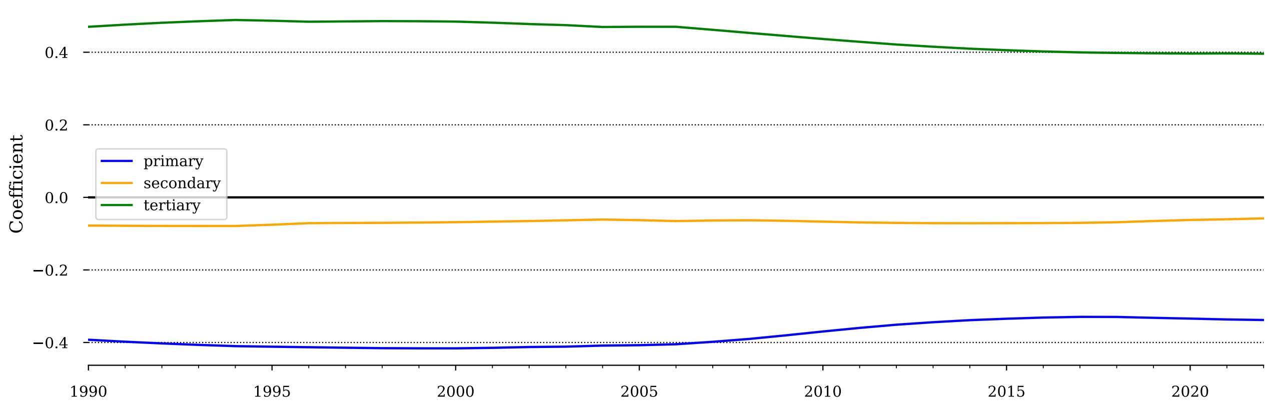
Source: CPS and authors' calculations. Monthly observations. Not seasonally adjusted.



(a) Sex: Female



(b) Age: 16-24 years old

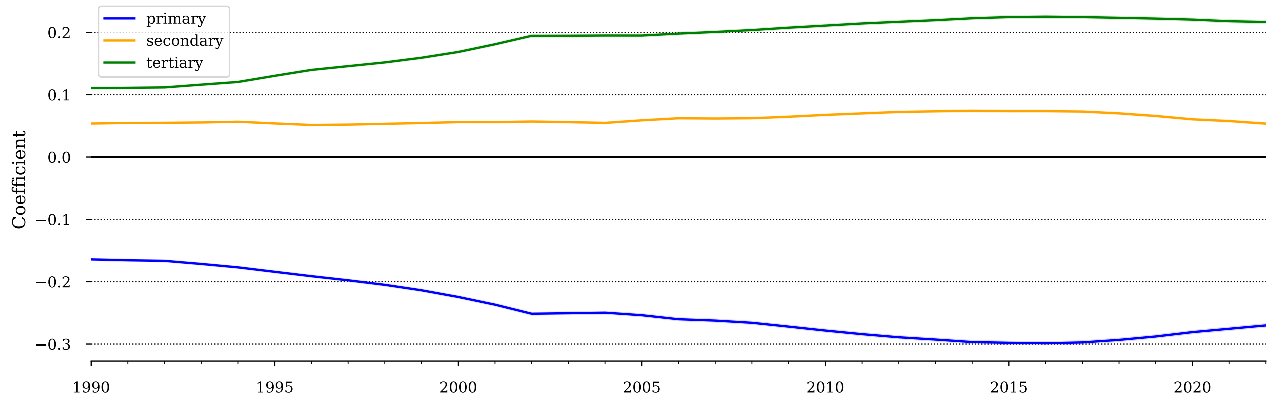


(c) Age: 55 and over

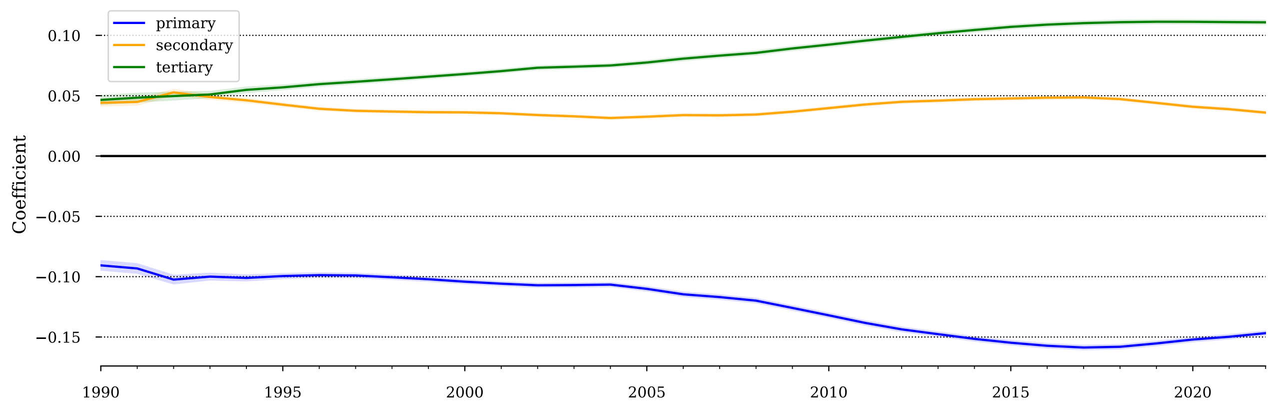
Figure B.11: Evolution of regression coefficients for posterior probabilities (sex and age).

Source: CPS and authors' calculations.

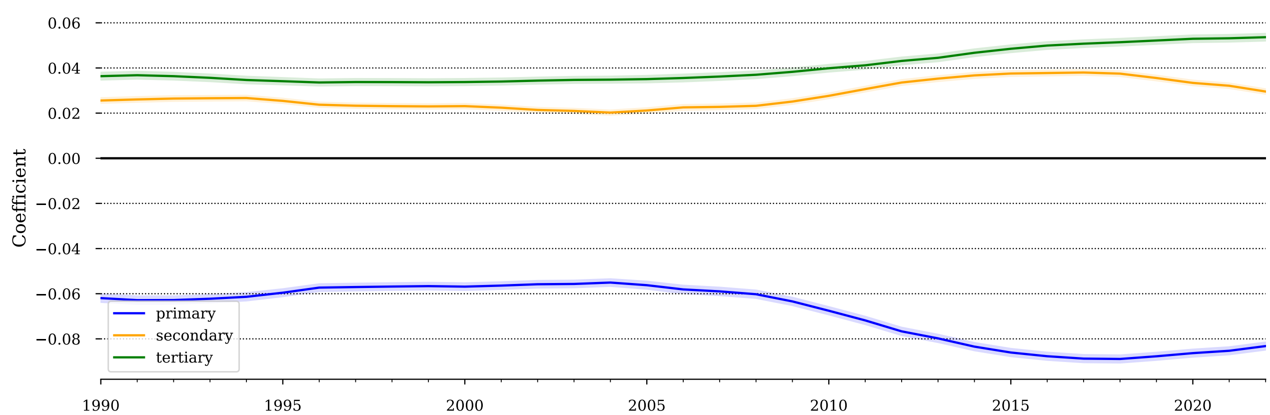
Notes: Regression coefficients for annual 10-year rolling samples for same specification as in Table 12.



(a) Education: Less than high school



(b) Education: High-school graduate

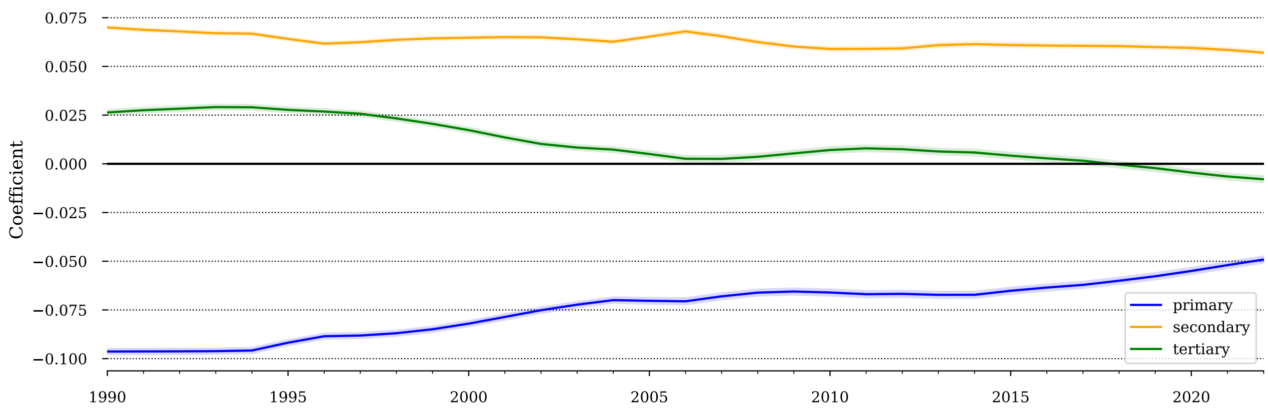


(c) Education: Some college

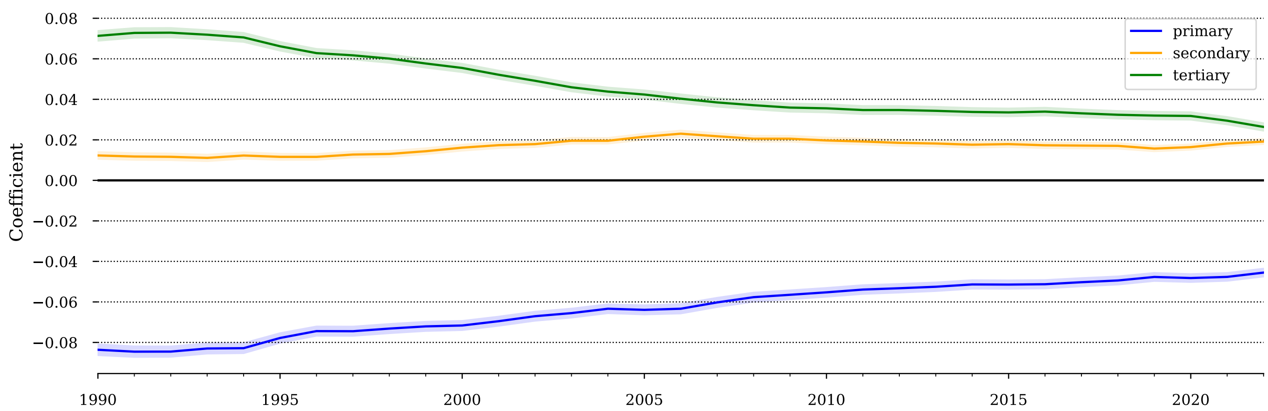
Figure B.12: Evolution of regression coefficients for posterior probabilities (education).

Source: CPS and authors' calculations.

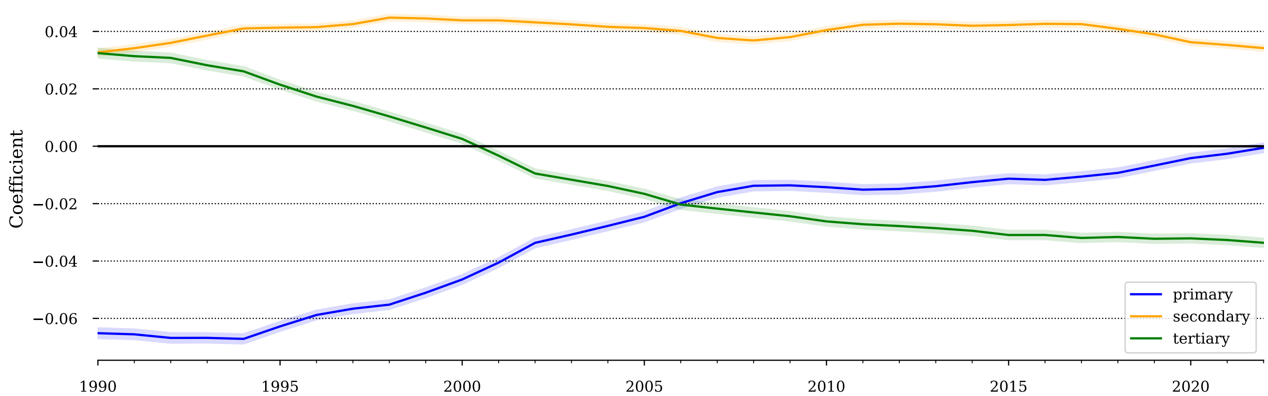
Notes: Regression coefficients for annual 10-year rolling samples for same specification as in Table 12.



(a) Race: Black



(b) Race: Other (neither black nor white)



(c) Ethnicity: Hispanic

Figure B.13: Evolution of regression coefficients for posterior probabilities (race and ethnicity).

Source: CPS and authors' calculations.

Notes: Regression coefficients for annual 10-year rolling samples for same specification as in Table 12.