The Dual U.S. Labor Market Uncovered

Hie Joo Ahn

Federal Reserve Board

Bart Hobijn

Arizona State University
and Federal Reserve Bank of San Francisco

Ayşegül Şahin

University of Texas at Austin and NBER*

Draft: March 23, 2022

Abstract

Aggregate U.S. labor market dynamics are well approximated by a dual labor market supplemented with a third home-production segment. We estimate a Hidden Markov Model with (in-)equality restrictions, a machine-learning method, to uncover this structure in which the different market segments are identified through inequality constraints on labor market transition probabilities. This method yields time series of stocks and flows for the three labor market segments for 1980-2021. Primary sector workers, who make up around 55 percent of the population, are almost always employed and rarely experience unemployment. The secondary sector, which constitutes only 14 percent of the population absorbs most of the short-run fluctuations in the labor market, both at seasonal and business cycle frequencies. Workers in this segment experience 6 times higher turnover rates than those in the primary tier and are 10 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. While we find that young workers, racial minorities, and workers with lower educational attainment are more likely to belong to the secondary sector, the bulk of labor market segment variation across individuals cannot be explained by observables. Our findings imply that aggregate stabilization policies, such as monetary policy, predominantly works through the small but turbulent secondary market.

JEL classification codes: J6, J20.

Keywords: Dual labor markets, Hidden Markov Models, machine learning.

*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 or of any of the institutions that the authors are affiliated with.

1 Introduction

We show that U.S. labor market dynamics at the macro and individual levels 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 the dual labor market structure of the U.S. labor market by estimating a Hidden Markov Model (HMM) with inequality restrictions using labor market 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 labor market 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, like size-dependent policies (Guner *et al.* , 2008), unionization (Berger *et al.* , 1980), and tiered contracts (Bentolila *et al.* , 2019), 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 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 even in the absence of such institutional arrangements and structures. 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 much harder.[2] Moreover, the growing importance of the participation margin in

[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

labor market fluctuations requires taking a clear stand on the home production sector.

We revive the DLM Hypothesis for the U.S. labor market by combining sophisticated machine learning techniques with rich panel data on individual-level labor market histories. Specifically, we introduce a Hidden Markov Model (HMM) with inequality restrictions that classifies the more than 10 million respondents in the Current Population Survey (CPS) from 1980-2021 into three labor market segments. The first two correspond to the primary and secondary tiers that are at the core of DLM theory. We also identify a tertiary sector that is made up of those who only infrequently participate in the labor market who sometimes experience unemployment and labor market turnover.

The methodology we employ is inherently a clustering problem that is solvable with a machine learning algorithm in the spirit of Hall and Kudlyak (2019), Gregory *et al.* (2021), Shibata (2019). We build on this small, but growing, literature and extend the methodology along four key dimensions. First, we use identifying restrictions for the hidden states that give them a direct economic interpretation. These restrictions are informed by the key aspects of dual labor market theory and help bridge the gap between economic theory and unsupervised machine learning.[3] Second, we estimate *time series* of the stocks and flows for each of the hidden states, in contrast to the models of Hall and Kudlyak (2019), Gregory *et al.* (2021) and Shibata (2019). Therefore, we can analyze seasonality, cyclicality and trends in the estimates comprehensively that allow us to study both their long-run trends and business cycle properties. Third, we use detailed labor force status data in the CPS to inform our hidden states.[4] Fourth, we provide a tool that bridges the gap between individual and aggregate dynamics in the labor market. Most models set to fit the aggregates underestimate persistence of employment and unemployment at the individual level since they inherently assume a first-order Markovian process for labor market transitions. Put differently, the assumption of lack of history dependence on labor market transitions fails to match the rich individual heterogeneity in observed labor market histories. Our methodology provides a tool to reconcile this important inconsistency.

Our estimation approach yields two sets of rich results: The first is the time-series estimates of the stocks of and flows between employment, unemployment, and non-participation in each of the three labor-market segments we distinguish. The second is the set of individual-level posterior probabilities that each respondent in the CPS from 1980-2021 is in the respective

and Psacharopoulos, 1981).

[3]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.

[4]While more refined labor force states such as part-time for economic reasons, discouraged or marginally attached states have been the focal point of recent discussions, they have not been systematically introduced into a unified framework.

tiers of the labor market. A stark picture of the U.S labor market emerges from these results.

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 6 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 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 nonparticipation 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 macroeconomic models of the labor market, are not at all reflective of individual labor market experiences and outcomes.

The individual-level probabilities of being in each segment of each of the 10 million individuals in the CPS in the last 40 years allow us to examine the role of observed characteristics of workers in their labor segment assignments. We find that young workers, racial minorities, and workers with lower educational attainment are more likely to belong to the secondary sector. While enriching models with heterogeneity in observed characteristics, such as age, gender, and education, can partially alleviate the conflict between individual and aggregate labor market dynamics, our individual-level findings show that these only account for about a fifth of the cross-individual variation in membership of the three labor market tiers. Moreover, the explanatory power of observables has been declining over time and the bulk of workers' outcomes depend on their labor segment designation.

A useful way to assess the role of each segment in the aggregate economy is to quantify the contribution of each segment to aggregate labor market statistics and to the ongoing trend changes in the economy. We show that the primary market accounts for more than 80% of employment and participation but its contribution to the unemployment rate much smaller—only about a quarter of aggregate unemployment rate. What is probably the most striking finding is that, the secondary market whose employment share is only 13.6% accounts for 60% of unemployment in the economy. Labor market dynamism as measured by flows per capita is also highly uneven with the secondary market accounting for half of the turnover in the economy. Moreover, the two notable trends, the trend decline in the unemployment rate and the decline in labor market dynamism are mostly accounted by changes in the secondary sector.

The rest of this paper is structured as follows. In the next section we discuss the details of our methodology in the context of the literature and explain how the DLM provides a way to think about many dimensions of micro and macro heterogeneity at the same time. Next, 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 overall labor market and we quantify the importance of each of the three segments for the trends and cycles in commonly analyzed aggregates. In Section 5 analyze the individual-level evidence. We discuss implications of our estimates for theory and policy in Section 6 and conclude in the next.

2 Related Literature

The classification of individuals in labor force surveys into employed, unemployed, and non-participants is the common way of summarizing the underlying macro heterogeneity in the economy. While it captures some very important differences in workers' labor market outcomes, it remains too coarse to characterize many different aspects of individual and aggregate labor market outcomes. Various recent studies have emphasized the importance of different subcategories of persons within the three labor market states reported in the CPS for individual and aggregate outcomes. These include heterogeneity among the unemployed that accounts for duration distribution of unemployment (van den Berg and van Ours, 1996; Hornstein, 2012; Ahn and Hamilton, 2020a; Kroft *et al.* , 2016; Elsby *et al.* , 2015), heterogeneity in the type of jobs for the employed to account for worker turnover and the tenure distributions (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; Sedlacek, 2016). 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 categories of employment, unemployment, and non-participation. We refer to these subcategories as "Macro Heterogeneity".

Relatedly, commonly used models with search frictions (e.g. Mortensen and Pissarides, 1994; Shimer, 2005) 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 the multi-period transition probabilities between labor market states and are likely to provide biased estimates of the costs of unemployment and fail to match the persistence and asymmetric dynamics of the unemployment rate over the business cycle. Since most macroeconomic models of the labor market imply this Markovian property for transitions between different states, it is important to find a representation of the U.S. labor market in terms of hidden states between which the flows have this property.

The identification of Macro Heterogeneity, by definition, involves classifying individuals at each point in time into untagged hidden labor market states. The method we use is a Hidden Markov Model which is a statistical tool that estimates latent states and their dynamics from data of categorical sequences. It is a model-based clustering method that classifies subjects into a handful of distinct groups based upon their latent dynamic behavior characterized by a Markovian process. This type of classification problem is a form of machine learning that is often referred to as *unsupervised* since the algorithm does not use prior information about who belongs to which group.

This methodology has various advantages for the identification of the dual structure of the U.S. labor market. First, the identification of duality structure is informed by the persistence and turbulence in transitions between labor force states that form the basis for latent dynamics. Second, the method produces estimates of three objects that are crucial for empirical and theoretical analyses of the labor market—the fraction of the population in each market segment; transition probabilities within each segment; individual-level probability of belonging to each segment.[5]

The closest related paper to ours is Shibata (2019) who also estimate an HMM. However, there are various notable differences. Our HMM with (in-)equality restrictions is specifically designed to recover the duality structure. The identifying restrictions are guided by the Dual Labor Market Hypothesis posited in Doeringer and Piore (1970). For our analysis, we focus on two aspects that distinguish the three market segments. The first is differences in turnover rates. Employment stability is higher in the primary tier than in the secondary and tertiary tiers. Non-participation in the tertiary market is more persistent than in the primary and secondary tiers. The second is limited mobility of workers between the market segments. The

[5] A distance-based clustering method such as K-means clustering as in Gregory *et al.* (2021) is not directly applicable since there is no natural concept of distance in categorical data.

former is implemented with inequality restrictions on transitions between hidden states and the latter is implemented with zero restrictions on transitions between the markets. Another important distinction is economic interpretability. Typically the estimated hidden states from HMMs do not have direct economic interpretations. We do not allow the misclassification of labor force status by imposing zero restrictions to the emission block of HMM, so that the estimates are directly interpretable in the context of observed labor force status and the DLM theory. Moreover, our HMM is a *dynamic* model, which produces the time-varying estimates of aggregate and individual-level statistics. As a result, we can analyze seasonality, cyclicality and trends in the estimates comprehensively.

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 models (SVAR) as in Stock and Watson (2001), Christiano *et al.* (2006), and Baumeister and Hamilton (2015). However, it has not been yet been applied in labor economics since its implementation is numerically challenging. 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. Just like the BW algorithm, our method is an application of the Expectation-Maximization (EM) algorithm (Dempster *et al.* , 1977). The difference is that in our method the M-step involves the numerical maximization of the expectation of the complete-data likelihood function with respect to the identifying (in-)equality restrictions we impose. We show that this maximization is manageable because it can be split up into a set of well-behaved convex maximization problems for which efficient numerical methods are available.

3 Identification of a Dual Labor Market in an HMM

Though HMMs have been used in several empirical studies of U.S. labor market data,[6] the specific application to uncover the segments of the DLM using inequality constraint as identifying assumptions is new in this paper. To explain our contribution, we first describe the structure of the HMM we estimate. We then discuss the identifying assumptions. Finally, we describe how we deal with these restrictions in the estimation of the model and how it does not only yield parameter estimates but also posterior probabilities that each of the survey respondents in the CPS are in each of the market segments.

[6] For example, Feng and Hu (2013), Boeschoten *et al.* (2020), and Shibata (2019).

3.1 Structure of the DLM 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 (Flood *et al.* , 2020) we use for that purpose. We describe our benchmark specification here and discuss why we chose this specification in Subsection 4.1, where we cover model selection.

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). Here $M \in \{P, S, T\}$ denotes the market segment. The 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 aggregate 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. 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

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

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) for example. 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.[7]

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).

In addition to the emissions listed in Table 2, there are also missing observations. The 4-8-4 panel structure of the CPS is such that for each individual i who enters the sample in period t_i we have observations for $t = t_i, \ldots, t_i + 15$. At least 8 of these observations, and possibly more, are missing.

To summarize, we have a panel of incomplete observed 16-month long labor market histories

[7] 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.

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 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 Identification

Without any further restrictions on the parameters the model would be observationally equivalent for any permutation of the hidden states in the set L. This is known as "label swapping" (Allman *et al.* , 2009).[8] To avoid this and to make each element of L correspond to a particular economically meaningful hidden state, we introduce three types of restrictions on the parameters of the HMM we estimate.

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. 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 tertiary market is the one in which people go through persistent spells of non-participation. We use the following restrictions

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

to make sure our estimated model has this property.

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

[8] One approach, taken by Shibata (2019), is to estimate the unrestricted model and then label the hidden states ex-post based on the unrestricted parameter estimates. This however, is not practical in our case with time-varying parameters, since the meaning of each of the states can change from period to period in that case, i.e. the unrestricted version of our model suffers from "temporal label swapping".

matches their labels, we assume that long-term unemployment is more persistent than short-term unemployment. That is

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

In addition, we impose that persons can only flow from short- to long-term employment 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 above four constraints 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.

No classification errors: Zero restrictions on emission probabilities

Our goal is 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 Bureau of Labor Statistics (BLS). With that in mind, we follow the BLS and assume that respondents correctly report their labor market status of employment, unemployment, and non-participation. Thus, we do not allow for classification errors.[9]

If respondents always correctly report their labor market status (employed, unemployed, non-participant), then this implies that 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.

[9] There is an extensive literature on such classification errors (Abowd and Zellner, 1985; Blanchard *et al.*, 1990; Feng and Hu, 2013; Elsby *et al.*, 2015; Ahn and Hamilton, 2020b) and a large degree of disagreement about their importance. This is beyond the scope of our analysis in this paper.

Random missing values

In addition to the (in-)equality restrictions, we also 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 she or he is in. This way of treating the missing values means that no information is gleaned from whether an observation is missing or not.[10] This assumption is, by definition, true for the 8-month reporting gap in the CPS during which respondents drop out of the sample. So, during that gap the probability of a missing emission is one no matter what the hidden state.

3.3 Estimation

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, $L(\boldsymbol{\theta})$, 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.[11]

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

It is the joint probability of the observed data from t_i through t_i+k *and* individual i being in

[10] 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. Though they do not use an explicit HMM.

[11] 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.

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 could quickly become 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

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

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

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

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.

Here, we 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 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

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

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

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

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.

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

Thus, the BW algorithm does not only yield a set of parameter estimates. For these estimates it also provides posterior probabilities of the stocks for each of the individuals in the data.

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

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

Because we impose the restriction that individuals cannot flow from one market segment to another, this probability is constant over time.[12]

Our estimation procedure thus yields two additional variables for each respondent in the CPS that reflect the posterior probabilites that she or he is part of the primary or secondary segment of the labor market.[13]

[12]See Appendix A for a proof.

[13]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.

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 solved easily using the numerical method introduced in Andersen *et al.* (2011).[14]

Thus, even though our specification has a large number of parameters and constraints, we are able 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.

4 Results

In this section we present the estimated stocks and flows, as well as emission probabilities, for the benchmark specification described in Section 3. We first discuss how our choice of benchmark specification was guided by several model selection criteria. We then show how each of the three market segments are characterized by very different outcomes in terms of unemployment and participation rates, employment-to-population (EPOP) ratios, as well as turnover. In the final two subsections we aggregate the segment-specific results and show what each of the tiers contribute to averages, trends, and fluctuations in aggregate labor market statistics.

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

4.1 Model selection

We estimate our model for all respondents, $i = 1, \ldots, n$, in the CPS from 1980-2021. The resulting sample size is $n = 10,271,333$ individual 4-8-4 labor market histories. Our benchmark specification has two main characteristics: (i) It has three labor-market segments, and (ii) each segment has two types of unemployed persons. Our choice of this model as our benchmark is based on a comparison of the model with several alternative specifications, all estimated using the 29 emissions listed in Table 2.

Table 3 lists the relevant statistics for the benchmark model and the three most notable alternative models we compared it to.[15] The table lists the number of labor market segments, hidden states, parameters, the resulting log-likelihood value, as well as two information criteria for each model. Models with lower values for the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are preferred over those with higher ones.

The top line contains the statistics for the benchmark model for which we present our estimates in the rest of this section. As discussed in the previous section, it consists of three labor market tiers with 12 hidden states. The total number of parameters is 90,216, which is 179 for each of the 504 months in the sample.

The second line in the table provides a baseline for comparison. It is the First-Order Markov (FOM) model in which there is one labor market with the three observed states of employment, unemployment, and non-participation for which transitions between these states are assumed to follow a first-order Markov process.[16] The information criteria show that this baseline is clearly rejected compared to our benchmark of the DLM with tertiary sector. This 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 baseline, by assumption, cannot.

The third line shows the results for a specification of 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, as we discuss in more detail below. The AIC and BIC both show that the 2-segment specification is not preferred compared to the benchmark.

The final 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

[15] In principle, there are many different permutations of specifications possible. We focus on the most relevant here.

[16] For example, in the Mortensen and Pissarides (1994) 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.

secondary tier is rejected compared to our benchmark model. The improved fit for multiple 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.[17]

4.2 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. These stark differences between the market tiers are clear from their average outcomes over time.[18] These averages are listed in Table 4, which provides them for the main labor market aggregates for each segment as well as for the labor market as a whole.

The last row of the table reports a new labor market aggregate that we introduce to measure the degree of dynamism. It measures the average annual flows per capita between the observed labor market states of employment (E), unemployment (U), and non-participation (N). It is a single summary statistic of the incidence of turnover that is comparable across segments.[19]

The first line of the table shows the relative size of the segments. The majority of the population, 54 percent on average, is part of the primary market. The secondary market is the smallest and consists of 14 percent of the population. The remaining one-third is in the tertiary sector.

Primary sector

The primary sector is characterized by a low unemployment rate, high labor force participation rate (LFPR), and, consequently, a high 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 1 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 almost irrelevant for primary sector workers who constitute more than half of the population. These workers are almost always employed and they 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 5 for the three sectors.

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

[18]We focus on these averages here and provide the underlying estimated time series in Figures B.1 through B.4 in Appendix B.

[19]The estimated flow rates, $q_{l,l',t}$, on which this measure is based are shown in Figures B.5 through B.10.

They also seamlessly move from non-participation to employment and vice versa, unlike workers in the secondary and tertiary sectors for whom these transitions often tend to involve a period of unemployment. In many ways one can think of the labor market experiences of these primary-sector workers as being captured well by those in standard Real Business Cycle (RBC) models (Cooley and Prescott, 1995).

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 1 shows, flow rates in the secondary sector are elevated for all six types of flows.

So, 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 of the outcomes of workers in the secondary sector, search models (like Mortensen and Pissarides, 1994) are most applicable to this segment of the labor market.

Tertiary sector

Only 9 percent of the persons in the tertiary sector participate in the labor market. Those that do have a high unemployment rate of 20 percent. This is mostly caused by those entering the labor force looking for a job. This is different than for the secondary tier, in which job-loss is an important reason for unemployment.

Business cycle fluctuations

So far, we have emphasized the marked differences between the average outcomes of the market segments. In addition to these long-run differences, they also have very distinct business cycle properties. This can be seen from Table 6. It shows the standard deviation of HP-filtered data for each labor market segment. The most striking thing that jumps out from this table is that the secondary market exhibits more volatility for all aggregates we consider. Both the unemployment and labor force participation rates are more than 5 times as volatile in the secondary sector as in the primary sector. The tertiary sector exhibits notable unemployment volatility but the labor force participation rate varies much less than in the secondary segment.

Table 6 also reports the correlation of the cyclical component of each labor market indicator with the cyclical component of GDP. Similar to the aggregate, the unemployment rate is strongly countercyclical in all segments. Despite being very low on average, the primary sector's unemployment rate has the highest negative correlation with GDP growth. Unlike unemployment, the labor force participation rate displays differential cyclical comovement across segments: it is procyclical in the primary sector, countercyclical in the secondary, and almost acyclical in the tertiary. Consequently, the employment-to-population ratio in the primary sector show the most pronounced positive comovement with output.

4.3 Contributions of segments to aggregates

Now that we have established that the U.S. labor market is made up of three very different segments, the next question is how these three distinct parts add up to the whole, i.e. to the labor market aggregates published by the BLS? Here, we answer this question by looking at the contributions of the three market tiers to the averages of and fluctuations in these aggregates.

Table 7 shows these contributions. For example, if we consider the first element of the third row, the primary segment accounts for 53 percentage points of the 66 percent average LFPR in the U.S. over our sample period. The secondary sector, which covers 14 percent of the population, only accounts for about a sixth of the labor supply. The most striking result in the table is that the secondary sector accounts for more than 4 percentage points out of the 6.6 percent average unemployment rate. Thus, one-sixth of the labor force accounts for 60 percent of unemployment. This is the same 14 percent of the population that accounts for 47 percent of gross flows in the economy.

The importance of the secondary sector for the dynamism of the U.S. labor market is illustrated in Figure 2. It shows that the share of the secondary sector is the highest for flows between unemployment (U) and non-participation (N). This reflects that the secondary tier is made up of workers at the margin of the labor market. The primary and secondary segments account for almost all job-loss and job-finding in the economy. Interestingly, flows between employment and non-participation are relatively more evenly distributed across different segments.

The secondary sector does not only disproportionately contribute to the aggregate unemployment rate. Fluctuations in the segment also carry an outsized weight in terms of fluctuations in the labor market aggregates. This can be seen from Table 8. The rows labeled with "$\sigma^2 (\Delta x_t)$

"and "$\sigma^2(\Delta_1 2x_t)$"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. For example, the second row shows that the secondary market contributes 0.073 to the 0.154 variance of monthly changes in the unemployment rate. We report the statistics for both the monthly and 12-month changes, because the latter include seasonal fluctuations.

The shift part in the secondary sector contributes well beyond the 14 percent population weight of the segment to fluctuations in all labor market aggregates considered. The share contributed is higher for monthly changes than for 12-month changes. Thus, the secondary tier absorbs more than its share of seasonal fluctuations and annual fluctuations in the labor market.[20]

4.4 Contributions of segments to long-run aggregate trends

Over the past four decades the U.S. labor market has seen some profound trends. Our analysis provides a unique perspective on the sources of these trends.

The trend decline in the unemployment rate in the US is well known. The origin of this decline is the stark moderation in the incidence of unemployment which declined by more than 50% from 1980s to 2020s as evident by the job-loss and job destruction rates as shown by Davis *et al.* (2010) and Crump *et al.* (2019). This is mostly due to the decline in the employment-to-unemployment transition rates in the secondary market. This rate declined from about 10% to 5% in the last 40 years. Moreover, the inflow rate into unemployment from nonparticipation went down from around 20% to 10% in the secondary segment.[21]

The U.S. economy is known to be one of the most fluid labor markets in the world. However, it has been experiencing a notable decline in labor market dynamism as 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. Similar to other trends it is accounted by mostly by the changes in the secondary sector. While there is also some decline in the tertiary sector, it is notable that flows per capita in the primary sector remained largely unchanged at 0.5 in the primary sector over the last 40 years.

Finally, the labor force participation rate has been trending down since its peak in late 1990s. This decline is the result of a shift of the population from the primary to the tertiary sector over time consistent with the aging of the baby boom cohort.

[20] This can also be seen from Figures B.3 and B.4, that plot the time series for the sector-specific labor market aggregates.

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

5 Individual-level evidence

In addition to the segment-level estimates presented in the previous section, our method also yields posterior probabilities that each of the CPS respondents is part of one of the market tiers. In this section we use these estimated posterior probabilities to analyze the composition of the three segments and show how the DLM representation enhances our understanding of individual-level evidence in the CPS.

5.1 Reliability of market segmentation at the individual level

Contrary to other machine-learning methods,[22] our method does not classify each respondent into one of the segments. Instead, it provides an estimate of the posterior probability an individual, i, is part of a particular tier, M. This probability, $P_i(M)$, is given in (15).

Suppose that whether an individual belongs to a particular segment is highly uncertain. We then would observe that the individual's posterior probabilities are almost equal across the three segments, i.e. $P_i(M) \approx 1/3$. On the contrary, if an individual belongs to the primary segment with certainty, their primary probability individual's primary probability would be one and the secondary and tertiary probabilities zero.

For most respondents, the latter case is close to what we observe from the distribution of posterior probabilities by market. This is shown in Figure 3 that shows histograms of estimated values of $P_i(M)$ for $M \in \{P, S, T\}$ for the more than 10 million respondents in the CPS. This observation indicates that ex-post uncertainty about the classification is low.

This means that the incomplete 16-month labor market histories of the 29 emissions, listed in Table 2, that we use for the estimation of the HMM provide enough information to reliably classify individuals across the market segments. Table 9 shows what emissions are being used in the model to distinguish between members of the three market tiers. The estimated emission probabilities are very different across the columns associated with the different market segments.[23]

The employed in the primary segment, EP, report to be mainly full-time employed, EX. Those in the secondary market, ES, are often part-time employed for economic reasons, EPE. Those in the tertiary sector report that they are absent from work for other reasons more often than their counterparts. This difference in the type of employment reflects the difference in employment stability between the primary and secondary sectors that is at the heart of the DLM Hypothesis.

[22] Like k-means clustering (Gregory *et al.* , 2021).

[23] Petrie (1969) shows that an identified HMM has distinct columns in the emission probability matrix.

The different nature of unemployment across the market segments is revealed by the emissions. The unemployed in the primary sector, columns UPS and UPL, report to mostly have short durations of unemployment and are often on temporary layoff, UTL, or have completed a temporary job, UTJ. The majority of unemployed in the secondary sector, columns USS and USL, are job losers for reasons other than being on temporary layoff or because a temporary job ended, or they are job seekers who did not indicate job loss as their reason for unemployment. The reported duration of unemployment in the secondary sector is higher than in the primary one. The unemployed in the tertiary sector, columns UTS and UTL, tend to come from non-participation and thus report to be part of UX. Their reported unemployment duration is in between those in the primary and secondary sectors.

Non-participants in the primary sector, column NP, mostly report that they are either not searching but would like a job, NNS, or that they do not want a job, $NDNW$. The questions that identify discouraged workers, NDW, and marginal attachment to the labor force, NMA, that were introduced as part of the 1994 redesign of the CPS (Polivka and Miller, 1998), help the model infer whether non-participants are in the secondary sector, column NS. Thus, the secondary sector includes individuals that report to be more loosely attached to the labor force than the primary sector. Non-participants in the tertiary sector, column NT, are least attached and almost all of them report they do not want a job, $NDNW$.

5.2 Segment composition by observables

One of the things emphasized by many discussions of the DLM is that the secondary market is disproportionately made up of minorities, young, and less-educated workers (e.g. Doeringer and Piore, 1970; Piore, 1970; Berger *et al.* , 1980; Dickens and Lang, 1985). Our analysis does not use any observable demographic characteristics for the classification of respondents into the three segments. We do, however, have individual posterior probabilities that link the market segments and observable characteristics in the CPS.

Table 10 reports the distribution of major demographic groups across the market segments in the "within-group" columns. It also shows the composition of each segment across types of demographic groups in the "within-market" columns. It confirms the earlier discussions in the literature.

Women are most overrepresented in the tertiary market. Black workers make up one and a half times as large a share of the secondary market than of the whole population. For young workers, aged 16 to 24, this ratio is even higher. The secondary market is also disproportionately made up of those with a high-school degree or less.

But, one has to be careful about typifying the secondary segment as the one that consists of these four groups. This is not the case at all. We show this using the following regression:

$$p_i^M = c_0^M + c_x^M X_i + D_t^M + \epsilon_i^M \qquad \text{for} \quad M \in \{P, S, T\}, \tag{16}$$

where p_i^M is the posterior probability that individual i belongs to market segment M, X_i denotes a set of observable characteristics of individual i, D_t^M is a dummy for the calendar month in period t in which the individual enters the CPS, and ϵ_i^M is the residual.

The coefficients and R^2 are reported in Table 11. The signs of the coefficients in the regression are all in line with the composition Table 10. The R^2 shows that the observable demographics explain, at maximum, 20 percent of the cross-individual variation in membership of the different segments.[24] Note that the R^2 for the secondary market is by far the lowest. Demographic characteristics explain only about 5 percent of the variation in market membership.

The results in Tables 10 and 11 cover our whole sample from 1980-2021. During that period, however, the U.S. labor market went through important changes in terms of the labor supply of women, an increase in the skill premium as well as in polarization. These changes are reflected in the evolution of the regression coefficients in Table 11 over time.

Figure 4 reports changes in the coefficients of rolling regressions, with a window of 10 years, by market segment. The result reveals substantial variation in the interaction between demographic characteristics and market segmentation. It shows that the largest changes in these coefficients reflects trends in the primary and tertiary sectors. The coefficients for the secondary segment are relatively constant.

The first trend is that women become more likely to be in primary market but less likely to be in tertiary market, as shown by the rising coefficients of primary probability and the falling coefficients of tertiary probability on being female (top left and right panels). Notably, the declining coefficient for the tertiary probability slows after 2000, consistent with the convergence of male and female labor force attachment during this period (Albanesi and Şahin, 2018). We find that the increase in tertiary probability of men is mostly concentrated among men with lower educational attainment consistent with the findings of Heathcote *et al.* (2020).

The second trend is the changes in the age coefficients in the primary and tertiary segments. These changes are associated with the aging of population and the increased college enrollment of young individuals. Specifically, the coefficients of primary probability on being 55 and over

[24] This is an overestimate, because it includes the part of the variance explained by the calendar-month dummies. The addition of the dummies does not materially affect the R^2.

decline, while those on tertiary probability rise through the mid 2000s. From the mid 2000s, however, the coefficients of primary probability rebounds and those of tertiary probability drops again, which is likely driven by the delayed retirement of old workers. Meanwhile, young individuals aged 16-24 become less likely to be in primary segment, but more likely to be in tertiary market. This observation is consistent with the increased college enrollment which slows the labor-force entrance of young workers.

Last, the association between race and and market segmentation has also evolved since 1980. The coefficients on the primary probability for Black and Hispanic workers show an upward trend, making them, everything else equal, about 3 percent more likely to be in the primary sector now than in 1980.

5.3 Individual labor market histories

Throughout, we have emphasized how our DLM representation of the labor market provides an overarching framework within which to the various dimensions of Macro Heterogeneity pointed to as necessary to explain aggregate labor market dynamics. The aggregate dynamics, of course, are the sum of individual-level labor-market histories.

These individual-level histories are markedly non-Markovian in that past labor market statuses beyond that of the previous month significantly help predicts current status and transitions (Kudlyak and Lange, 2017). This observation contradicts the assumptions of most search models of the labor market. These models are based on the same assumption as the FOM model we discussed in the model specification section.

In our DLM representation such histories are a mixture of the different histories associated with the three market segments and, as a result, are non-Markovian. This results in a much better fit of the 12-month transition probabilities in the data. This can be seen from Table 12, which shows the 1-month and 12-month transition probabilities observed in the data, implied by our DLM model, and those implied by the First-Order Markov model.[25]

The table shows how the DLM fits most of the 12-month transition probabilities much better than the FOM model. This includes the persistence of employment, flows from employment to non-participation and vice versa, the persistence of non-participation, as well as flows from unemployment to non-participation.[26] The reason that the DLM model falls short of perfectly matching these is that some non-Markovian aspects of the data, like classification errors and seasonality, are not fully captured by the dynamics of the twelve hidden states in the model.

[25]The underlying time series for Table 12 are plotted in Figure B.11

[26]The results in this table are similar to those in Shibata (2019), who emphasizes the ability of HMMs to better match individual-level histories than the FOM.

6 Implications for theory and policy

Our analysis provides a clear narrative of the duality of the U.S. labor market. It documents the ex-post segmentation of the labor market and is not meant to reveal the reasons for the emergence of this segmentation. However, a better theory of these reasons is needed to improve our understanding of the DLM structure and its importance for labor market trends and fluctuations.[27] Though our analysis does not focus on the reasons for segmentation, our results do provide several useful pointers towards the sources of duality in the U.S. 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. 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.

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 resolved in response to negative economic shocks.[28] 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. The secondary segment of the labor market absorbs the bulk of the seasonal fluctuations.

This points to dualism in the U.S. labor market persisting to organize the division of labor 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 because the CPS data we use is collected at the monthly frequency. It does not include the high-frequency turnover in the labor market, like, 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.

The large differences in the cyclical sensitivity of the labor market tiers have important implications for the assessment of the costs of business cycles and unemployment as well as how

[27] In this sense, our results here are subject to similar criticism as the early research on dualism, as summarized by Cain (1975) and Wachter (1974).

[28] Piore calls this the endogenous "response to flux and uncertainty" (See Berger *et al.* , 1980, Chapter 2) and Saint-Paul (1997) illustrates the same intuition in the first figure in his book.

to think about the role and optimal design of unemployment insurance. Most studies of the costs of unemployment base their estimates on average incidence rates of unemployment across the population (e.g. Krusell *et al.* , 2010). However, our analysis suggests that these are not the relevant metrics to look at. The unemployment cost of business cycles is disproportionately borne by those in the secondary sector. A proper quantification of this cost should take this inequality into account and distinguish between the costs for workers in different labor market tiers. This insight also reveals that 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.

So far, we have discussed positive statements about the sources of labor market segmentation. A lot of normative questions related to the emergence of dualism, which were pointed out from the beginning of the discussion of the DLM Hypothesis (Doeringer and Piore, 1970), still remain and are pertinent for many policy discussions. For example, what is the reason workers select into the different market segments? If workers are ex-ante identical (as in Saint-Paul, 1997, Chapter 3, for example), then there is no reason for ex-post compensation of workers in the secondary sector for their incurrence of a disproportionate amount of seasonal and business cycle risks. If, instead, persons end up in the secondary sector due to unequal access to education, racial or gender discrimination, or because of a lack of information about the opportunities in the primary sector, then this raises possibilities for policies to eliminate these barriers and enhance long-run labor market outcomes.

The evidence we provide in this paper is also relevant for the discussion of policies that aim to stabilize labor-market fluctuations in the short-run. The focus of many of these policies is to maintain unemployment at or around Friedman's (Friedman, 1968) natural rate of unemployment. Because of the different degrees of business-cycle sensitivity across the market segments in our DLM representation, it is important for the implementation of such policies to identify who is in these segments and pay particularly close attention to those in the tier that is most cyclically sensitive. That tier is the secondary market.

7 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 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 nonparticipation but do not share the high job-loss rate of secondary workers.

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. Better quantitative analyses should take into account the Macro Heterogeneity in labor market outcomes we uncovered in this paper.

The HMM-based methodology that we introduced for the measurement of the stocks and flows in the three labor market segments can potentially be extended to account for job-to-job transitions,[29] be implemented in real-time to provide up-to-date estimates of the U.S. DLM for researchers and policy makers, and be used for cross-country analyses. The resulting posterior probability estimates for each respondent provide an additional variable that allows researchers to analyze the difference in outcomes across the segments we have identified for the whole breadth of measures reported in the CPS and its supplements. In short, our methodology and estimates expand the possible empirical analyses of the DLM Hypothesis to a broad set of issues, variables, and countries.

[29] Such an extension has to take into account the limitations of the source data described in Fujita *et al.* (2020).

References

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

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.

Albanesi, Stefania, and Şahin, Ayşegül. 2018. The Gender Unemployment Gap. *Review of Economic Dynamics*, **30**(October), 47–67.

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**.

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 Jean, Diamond, Peter, Hall, Robert E., and Murphy, Kevin. 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. 2020. Combining Multiple Imputation and Hidden Markov Modeling to Obtain Consistent Estimates of Employment Status. *Journal of Survey Statistics and Methodology*, 01. smz052.

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.

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

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.

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).

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., Hobijn, Bart, and Şahin, Ayşegül. 2015. 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.

Fujita, Shigeru, Moscarini, Giuseppe, and Postel-Vinay, Fabien. 2020 (July). *Measuring Employer-to-Employer Reallocation*. Working Paper 27525. National Bureau of Economic Research.

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., 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.

Heathcote, Jonathan, Perri, Fabrizio, and Violante, Giovanni L. 2020. The rise of US earnings inequality: Does the cycle drive the trend? *Review of Economic Dynamics*, **37**, S181–S204. The twenty-fifth anniversary of “Frontiers of Business Cycle Research".

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

Hyatt, Henry, and Spletzer, James. 2016. The shifting job tenure distribution. *Labour Economics*, **41**(C), 363–377.

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.

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.

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.

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.

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.

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*.

Polivka, Anne E., and Miller, Stephen M. 1998. The CPS after the Redesign: Refocusing the Economic Lens. *Pages 249–289 of: Labor Statistics Measurement Issues*. NBER Chapters. National Bureau of Economic Research, Inc.

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.

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

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.

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.

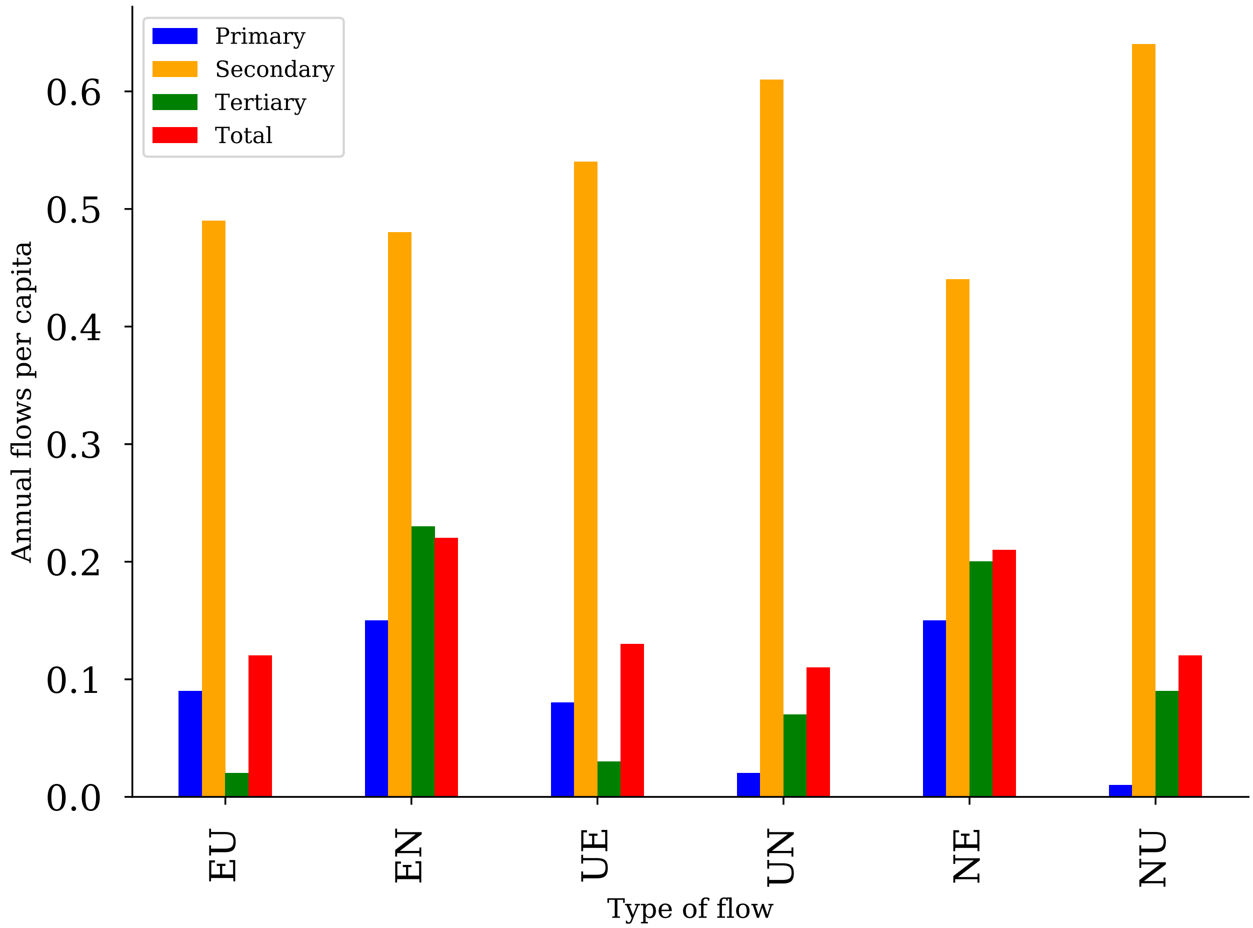


Figure 1: Average annual flows per capita by origin and destination.

Source: CPS and authors' calculations.

Notes: Averages taken over 1980-2021.

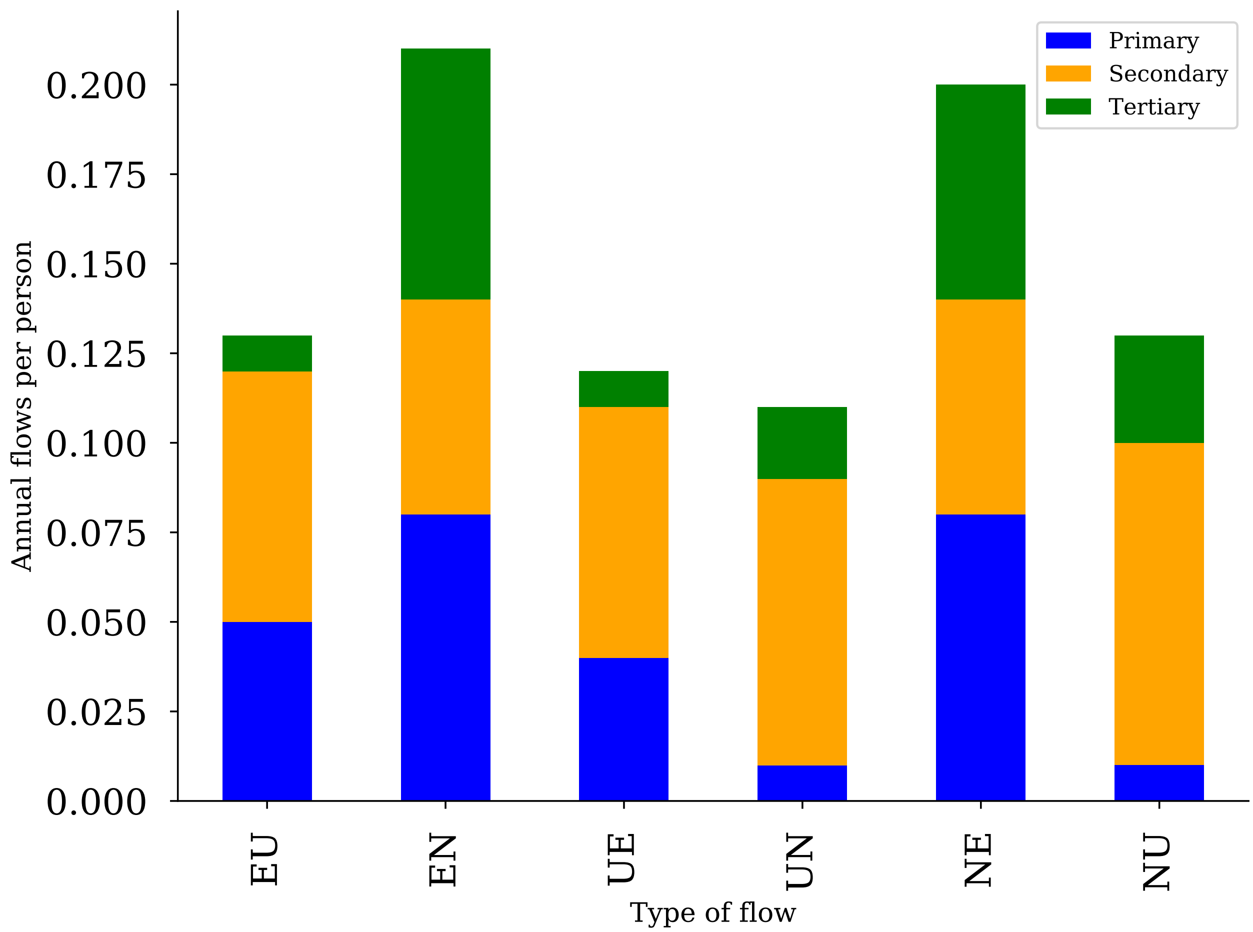


Figure 2: Composition of aggregate flows per person by type of flow and market.

Source: CPS and authors' calculations.

Notes: Averages taken over 1980-2021.

Figure 3: Distribution of posterior probabilities by market segment (1980-2021)

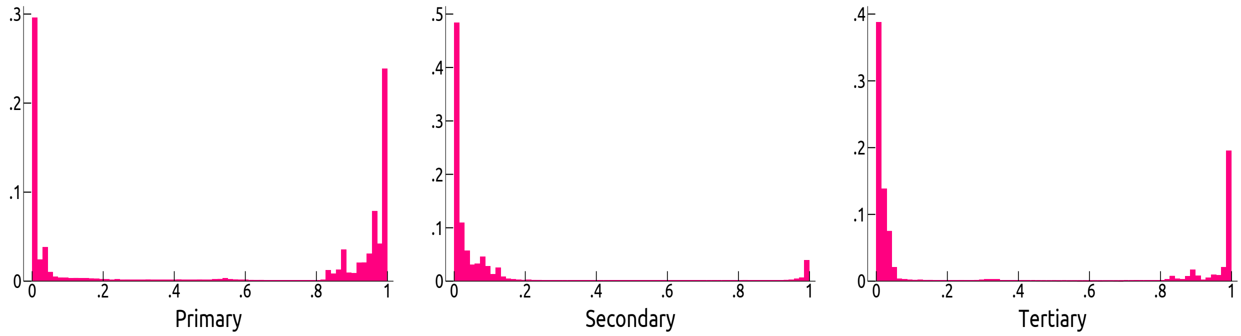


Figure 4: Coefficients from rolling regressions

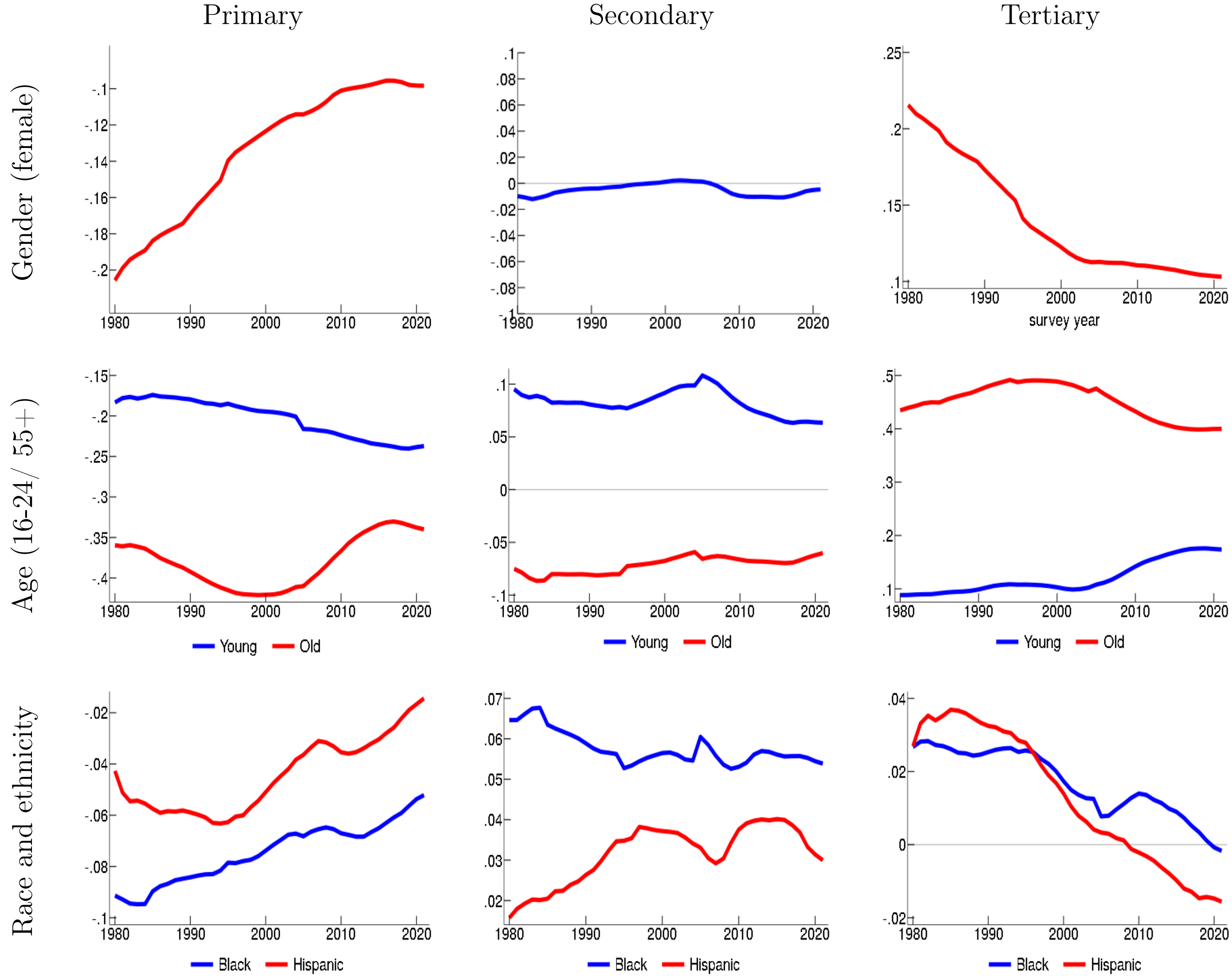


Table 1: Hidden states in HMM

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

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

Table 3: Comparison of model specifications

	segments	states	pars	logL	AIC	BIC
Dual Labor Market (benchmark)	3	12	90216	-3.41	70.18	71.46
First-Order Markov (FOM)	1	3	18648	-3.80	78.18	78.44
Dual Labor Market without tertiary sector	2	8	59976	-3.46	71.24	72.09
DLM, only two types of U in secondary	3	10	68040	-3.43	70.58	71.54

Source: Current Population Survey and authors' calculations.
Notes: Total number of observations is 10,271,333 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.

Table 4: Labor market aggregates by segment

	Primary	Secondary	Tertiary	Total
Share of population	54.39	13.55	32.05	100.00
Unemployment rate	2.05	26.52	20.01	6.62
Labor-force participation rate	97.24	73.13	9.21	65.77
Employment-to-population ratio	95.24	53.66	7.33	61.42
Flows per capita	0.50	3.19	0.63	0.91

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.

Table 5: 1- and 12-month transition probabilities in different market segments

segment	to freq from	E 1-m	US 1-m	UL 1-m	N 1-m	E 12-m	US 12-m	UL 12-m	N 12-m
Primary	E	97.91	0.73	0.04	1.32	95.16	0.82	1.21	2.80
	US	51.12	7.35	34.34	7.19	94.13	0.81	2.05	3.01
	UL	23.34	0.00	69.23	7.43	93.31	0.81	2.70	3.18
	N	46.26	2.15	1.96	49.62	94.91	0.82	1.40	2.87
Secondary	E	85.00	6.79	0.81	7.40	54.85	10.58	8.10	26.46
	US	31.88	31.17	7.75	29.19	53.52	10.62	8.63	27.23
	UL	13.36	0.00	63.62	23.03	51.93	10.52	9.68	27.87
	N	14.12	13.46	6.98	65.44	52.76	10.63	8.94	27.67
Tertiary	E	72.14	1.88	0.15	25.84	8.57	0.85	1.10	89.48
	US	18.72	9.50	26.96	44.82	8.17	0.84	1.40	89.60
	UL	15.04	0.00	64.24	20.71	8.93	0.84	1.95	88.29
	N	1.82	0.66	0.14	97.38	6.85	0.82	1.01	91.33

Source: Current Population Survey and authors' calculations.
Notes: Average 1-month and 12-month transition probabilities between hidden states.

Table 6: Business cycle statistics by labor market segment

measure	statistic	Primary	Secondary	Tertiary
Unemployment rate	$\sigma(x)$	0.52	2.61	2.54
	$\rho(x_t, x_{t-1})$	0.70	0.78	0.82
	$\rho(x_t, Y_t)$	-0.75	-0.63	-0.50
Labor-force participation rate	$\sigma(x)$	0.19	1.12	0.33
	$\rho(x_t, x_{t-1})$	0.60	0.81	0.65
	$\rho(x_t, Y_t)$	0.29	-0.25	-0.11
Employment-to-population ratio	$\sigma(x)$	0.61	2.02	0.37
	$\rho(x_t, x_{t-1})$	0.66	0.73	0.72
	$\rho(x_t, Y_t)$	0.70	0.49	0.26
Flows per capita	$\sigma(x)$	0.06	0.14	0.02
	$\rho(x_t, x_{t-1})$	0.51	0.54	0.38
	$\rho(x_t, Y_t)$	-0.61	-0.20	-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.

Table 7: Contribution to aggregates by segment

	Primary	Secondary	Tertiary	Total
Share of population	54.39	13.55	32.05	100.00
Unemployment rate	1.64	4.05	0.92	6.62
Labor-force participation rate	52.89	9.92	2.95	65.77
Employment-to-population ratio	51.81	7.26	2.35	61.42
Flows per capita	0.27	0.43	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.

Table 8: 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	0.0037	-0.0234	-0.0570	-0.0162
	$\sigma^2(\Delta x_t)$	0.1544	0.0124	0.0521	0.0730	0.0169
	$\sigma^2(\Delta_{12} x_t)$	1.1269	0.1736	0.3517	0.4644	0.1371
Labor-force participation rate	$\bar{\Delta} x_t$	0.0000	0.0405	-0.0048	-0.0257	-0.0099
	$\sigma^2(\Delta x_t)$	0.1596	0.0333	0.0121	0.0548	0.0594
	$\sigma^2(\Delta_{12} x_t)$	0.1623	0.1089	0.0084	0.0201	0.0250
Employment-to-population ratio	$\bar{\Delta} x_t$	0.0594	0.0355	0.0101	0.0189	-0.0050
	$\sigma^2(\Delta x_t)$	0.1797	0.0345	0.0372	0.0664	0.0415
	$\sigma^2(\Delta_{12} x_t)$	0.6848	0.2711	0.1622	0.1787	0.0728
Flows per capita	$\bar{\Delta} x_t$	-0.0060	-0.0003	-0.0005	-0.0037	-0.0016
	$\sigma^2(\Delta x_t)$	0.0080	0.0001	0.0025	0.0031	0.0023
	$\sigma^2(\Delta_{12} x_t)$	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.

Table 9: Average emission probabilities

State Emission	EP	ES	ET	UPS	UPL	USS	USL	UTS	UTL	NP	NS	NT
EX	98.7	67.9	93.4	-	-	-	-	-	-	-	-	-
EPE	1.0	29.7	2.3	-	-	-	-	-	-	-	-	-
ENW	0.3	2.4	4.4	-	-	-	-	-	-	-	-	-
UTL5	-	-	-	29.8	0.7	7.9	0.2	4.0	0.2	-	-	-
UTL14	-	-	-	4.0	14.1	4.1	0.3	0.8	0.8	-	-	-
UTL26	-	-	-	0.3	6.8	1.0	0.6	0.4	0.3	-	-	-
UTLLT	-	-	-	0.4	3.7	0.6	1.9	0.5	0.3	-	-	-
UTJ5	-	-	-	9.7	0.4	5.2	0.2	1.0	0.1	-	-	-
UTJ14	-	-	-	0.6	4.0	2.6	1.7	0.3	0.4	-	-	-
UTJ26	-	-	-	0.1	1.7	0.5	1.7	0.1	0.2	-	-	-
UTJLT	-	-	-	0.4	0.3	0.2	5.1	0.5	0.1	-	-	-
UJL5	-	-	-	32.1	2.9	9.7	1.2	1.6	0.2	-	-	-
UJL14	-	-	-	1.3	33.6	6.8	4.3	0.8	0.5	-	-	-
UJL26	-	-	-	0.2	19.3	1.1	6.5	0.5	0.3	-	-	-
UJLLT	-	-	-	1.0	9.6	0.4	29.8	1.7	0.3	-	-	-
UX5	-	-	-	16.5	0.5	33.8	1.5	71.3	6.9	-	-	-
UX14	-	-	-	2.0	1.8	18.9	9.1	8.3	52.4	-	-	-
UX26	-	-	-	0.5	0.5	3.8	7.1	0.7	23.6	-	-	-
UXLT	-	-	-	1.2	0.2	3.4	28.9	7.5	13.6	-	-	-
NTJDW	-	-	-	-	-	-	-	-	-	0.0	0.1	0.0
NTJMA	-	-	-	-	-	-	-	-	-	0.0	0.1	0.0
NTJNA	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
NTJNS	-	-	-	-	-	-	-	-	-	0.3	0.6	0.0
NTJDNW	-	-	-	-	-	-	-	-	-	1.2	0.9	0.2
NDW	-	-	-	-	-	-	-	-	-	0.9	3.6	0.1
NMA	-	-	-	-	-	-	-	-	-	1.4	7.0	0.1
NNA	-	-	-	-	-	-	-	-	-	0.4	1.9	0.1
NNS	-	-	-	-	-	-	-	-	-	10.2	19.1	1.3
NDNW	-	-	-	-	-	-	-	-	-	85.6	66.8	98.3

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

Table 10: Composition by market segment and by demographic group

	Within-group composition			Within-market composition			
	Primary	Secondary	Tertiary	Primary	Secondary	Tertiary	Total
Total	54.4	13.6	32.0	100	100	100	100
Men	63.6	12.3	24.1	55.0	50.1	36.4	48.5
Women	49.0	11.5	39.5	45.0	49.9	63.6	51.5
White	57.1	10.9	32.0	42.7	38.4	42.0	42.0
Black	50.0	18.3	31.8	5.8	9.9	6.4	6.5
Other	56.1	11.9	32.0	51.5	51.7	51.6	51.6
16-24	46.4	20.9	32.6	17.1	36.2	21.2	20.7
25-54	72.0	11.6	16.3	69.0	52.4	27.4	53.7
55+	30.5	5.3	64.2	13.9	11.4	51.5	25.6
High-school	46.5	14.3	39.2	44.9	64.7	66.3	54.1
Some college	63.7	10.7	25.6	29.7	23.5	21.0	26.2
College	72.3	7.1	20.6	25.4	11.8	12.7	19.7

Note: The category *High-school* denotes high-school graduation or less, and *Some college* denotes some college or associate degree.

Table 11: Coefficient estimates

	Primary	Secondary	Tertiary
Female	-0.13^{**}	-0.0051^{**}	0.13^{**}
	(0.00026)	(0.00017)	(0.00024)
Age 16-24	-0.21^{**}	0.079^{**}	0.13^{**}
	(0.00035)	(0.00023)	(0.00032)
Age 55+	-0.37^{**}	-0.069^{**}	0.44^{**}
	(0.00032)	(0.00021)	(0.00030)
High-school graduation or less	-0.19^{**}	0.054^{**}	0.14^{**}
	(0.00036)	(0.00024)	(0.00034)
Some college or associate degree	-0.067^{**}	0.022^{**}	0.045^{**}
	(0.00040)	(0.00026)	(0.00037)
Black	-0.064^{**}	0.058^{**}	0.0060^{**}
	(0.00040)	(0.00026)	(0.00037)
Hispanic	-0.031^{**}	0.033^{**}	-0.0018^{**}
	(0.00050)	(0.00033)	(0.00046)
R^2	0.198	0.052	0.239

Note: Coefficients on time dummies and a constant are not reported. R^2 is essentially the same as adjusted R^2.

Table 12: Multi-month transition probabilities in data and models

spec periods flow	data 1-month	model 1-month	FOM 1-month	data 12-month	model 12-month	FOM 12-month
E to E	95.61	95.40	95.73	90.09	87.07	72.20
E to N	2.95	2.97	2.74	7.26	8.91	23.97
E to U	1.43	1.63	1.53	2.65	4.02	3.83
N to E	4.55	5.01	4.46	10.76	15.54	40.66
N to N	92.96	92.00	92.88	86.92	80.76	54.56
N to U	2.49	2.99	2.66	2.32	3.70	4.78
U to E	25.26	25.30	24.39	49.99	56.89	56.64
U to N	22.69	22.48	21.18	27.90	29.86	38.98
U to U	52.06	52.22	54.43	22.11	13.25	4.38

Source: Current Population Survey and authors' calculations.
Notes: Average 1-month and 12-month transition probabilities in data, Dual Labor Market model, and First-Order Markov model.

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

$$\begin{aligned}\hat{u}_{i,t,l} &= E\left[u_{i,t,l} \mid \boldsymbol{x}_i;\boldsymbol{\theta}\right] = P\left(\ell_{i,t} = l \mid \boldsymbol{x}_i;\boldsymbol{\theta}\right) \\ &= \frac{P\left(\ell_{i,t} = l \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right)}{P\left(\boldsymbol{x}_i;\boldsymbol{\theta}\right)} = \frac{P\left(\ell_{i,t} = l \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right)}{\sum_{l'\in L} P\left(\ell_{i,t} = l' \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right)}.\end{aligned} \tag{17}$$

Similarly

$$\begin{aligned}\hat{v}_{i,t,l,l'} &= E\left[v_{i,t,l,l'} \mid \boldsymbol{x}_i;\boldsymbol{\theta}\right] = P\left(\ell_{i,t-1} = l \cap \ell_{i,t} = l' \mid \boldsymbol{x}_i;\boldsymbol{\theta}\right) \\ &= \frac{P\left(\ell_{i,t-1} = l \cap \ell_{i,t} = l' \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right)}{P\left(\boldsymbol{x}_i;\boldsymbol{\theta}\right)} \\ &= \frac{P\left(\ell_{i,t-1} = l \cap \ell_{i,t} = l' \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right)}{\sum_{h'\in L}\sum_{h\in L} P\left(\ell_{i,t-1} = h \cap \ell_{i,t} = h' \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right)}.\end{aligned} \tag{18}$$

Here, we can express

$$\begin{aligned}P\left(\ell_{i,t_i+k} = l \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right) &= P\left(x_{i,t_i},..,x_{i,t_i+k} \cap \ell_{i,t_i+k} = l\right) \\ &\quad \times P\left(x_{i,t_i+k+1},..,x_{i,t_i+15} \mid \ell_{i,t_i+k} = l\right) \\ &= \alpha_{i,k}\left(l\right)\beta_{i,k}\left(l\right),\end{aligned} \tag{19}$$

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

$$\beta_{i,k}\left(l\right) = P\left(x_{i,t_i+k+1},..,x_{i,t_i+15} \mid \ell_{i,t_i+k} = l\right). \tag{20}$$

Moreover

$$\begin{aligned}P\left(\ell_{i,t_i+k-1} = l \cap \ell_{i,t_i+k} = l' \cap \boldsymbol{x}_i;\boldsymbol{\theta}\right) &= P\left(x_{i,t_i},..,x_{i,t_i+k-1} \cap \ell_{i,t_i+k-1} = l\right) \\ &\quad \times q_{t_i+k,l,l'}\omega_{x_{i,t_i+k},l',t_+k} \\ &\quad \times P\left(x_{i,t_i+k+1},..,x_{i,t_i+15} \mid \ell_{i,t_i+k} = l'\right) \\ &= \alpha_{i,k-1}\left(l\right)q_{t_i+k,l,l'}\omega_{x_{i,t_i+k},l',t_+k}\beta_{i,k}\left(l'\right)\end{aligned} \tag{21}$$

This yields that

$$\hat{u}_{i,t_i+k,l} = \frac{\alpha_{i,k}(l)\,\beta_{i,k}(l)}{\sum_{l'\in L}\alpha_{i,k}(l')\,\beta_{i,k}(l')} \tag{22}$$

and

$$\hat{v}_{i,t_i+k,l,l'} = \frac{\alpha_{i,k-1}(l)\,q_{t_i+k,l,l'}\omega_{x_{i,t_i+k},l',t_+k}\beta_{i,k}(l')}{\sum_{h'\in L}\sum_{h\in L}\alpha_{i,k-1}(l)\,q_{t_i+k,h,h'}\omega_{x_{i,t_i+k},h',t_+k}\beta_{i,k}(l')}. \tag{23}$$

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

$$\begin{aligned}
\beta_{i,15}(l) &= 1\text{, and} && (24)\\
\beta_{i,k}(l) &= P\left(x_{i,t_i+k+1},..,x_{i,t_i+15} \mid \ell_{i,t_i+k}=l\right) && (25)\\
&= \sum_{l'} q_{t_i+k+1,l,l'}\beta_{i,k+1}(l') && (26)\\
&\quad\times\left[(1-\eta_{i,t_i+k+1})+\eta_{i,t_i+k+1}\omega_{x_{i,t_i+k+1},l',t_i+k+1}\right]\text{, for } k=0,\ldots,14 && (27)
\end{aligned}$$

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 $\mathcal{M}\in\{\mathcal{P},\mathcal{S},\mathcal{T}\}$:

$$\begin{aligned}
P\left(\boldsymbol{x}_i\cap\ell_{i,t}\in\mathcal{M}\right) &= \sum_{l\in\mathcal{T}} P\left(\boldsymbol{x}_i\cap\ell_{i,t}=l\right) && (28)\\
&= \sum_{l\in\mathcal{M}}\sum_{l'\in\mathcal{M}} P\left(\ell_{i,t}=l\mid\ell_{i,t-1}=l'\right)P\left(\boldsymbol{x}_i\cap\ell_{i,t-1}=l'\right) && (29)\\
&= \sum_{l'\in\mathcal{M}}\sum_{l\in\mathcal{M}} P\left(\ell_{i,t}=l\mid\ell_{i,t-1}=l'\right)P\left(\boldsymbol{x}_i\cap\ell_{i,t-1}=l'\right) && (30)\\
&= \sum_{l'\in\mathcal{M}} P\left(\boldsymbol{x}_i\cap\ell_{i,t-1}=l'\right) && (31)\\
&= P\left(\boldsymbol{x}_i\cap\ell_{i,t-1}\in\mathcal{M}\right) && (32)
\end{aligned}$$

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

$$\begin{aligned}\ln \mathcal{L} &= \sum_{i=1}^{n} w_i \left\{ \sum_{l\in L} \hat{u}_{i,t_i,l} \ln \delta_{l,t_i} + \sum_{k=1}^{15} \sum_{l'\in L} \sum_{l\in L} \hat{v}_{i,t_i+k,l,l'} \ln q_{t_i+k,l,l'} \right. \\ &\quad \left. + \sum_{k=0}^{15} \eta_{i,t_i+k} \sum_{l\in L} \hat{u}_{i,t_i+k,l} \ln \omega_{x_{i,t_i+k},l,t_i+k} \right\}. \end{aligned} \tag{33}$$

subject to the adding-up constraints

$$\sum_{l} \delta_{l,t} = 1, \text{ for } t = 1, \ldots, T \tag{34}$$

$$\sum_{l}^{\prime} q_{t,l,l'} = 1, \text{ for } t = 1, \ldots, T \text{ and } l \in L, \text{ and} \tag{35}$$

$$\sum_{x\in X} \omega_{x_{i,t_i+k},l,t_i+k} = 1, \text{ for } t = 1, \ldots, T \text{ and } l \in L \tag{36}$$

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

$$\sum_{l \in L} \hat{u}_{i,t,l} \ln \delta_{l,t}, \tag{37}$$

with respect to the unconditional probabilities (stocks), $\{\delta_{l,t}\}_{l \in L}$, subject to the adding-up constraint (34). 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

$$\sum_{l \in L} \sum_{l' \in L} \hat{v}_{i,t,l,l'} \ln q_{t,l,l'}, \tag{38}$$

subject to (35) 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

$$\sum_{l \in L} \hat{u}_{i,t,l} \ln \omega_{x_{i,},l,t}, \tag{39}$$

subject to (36).

B Additional empirical results

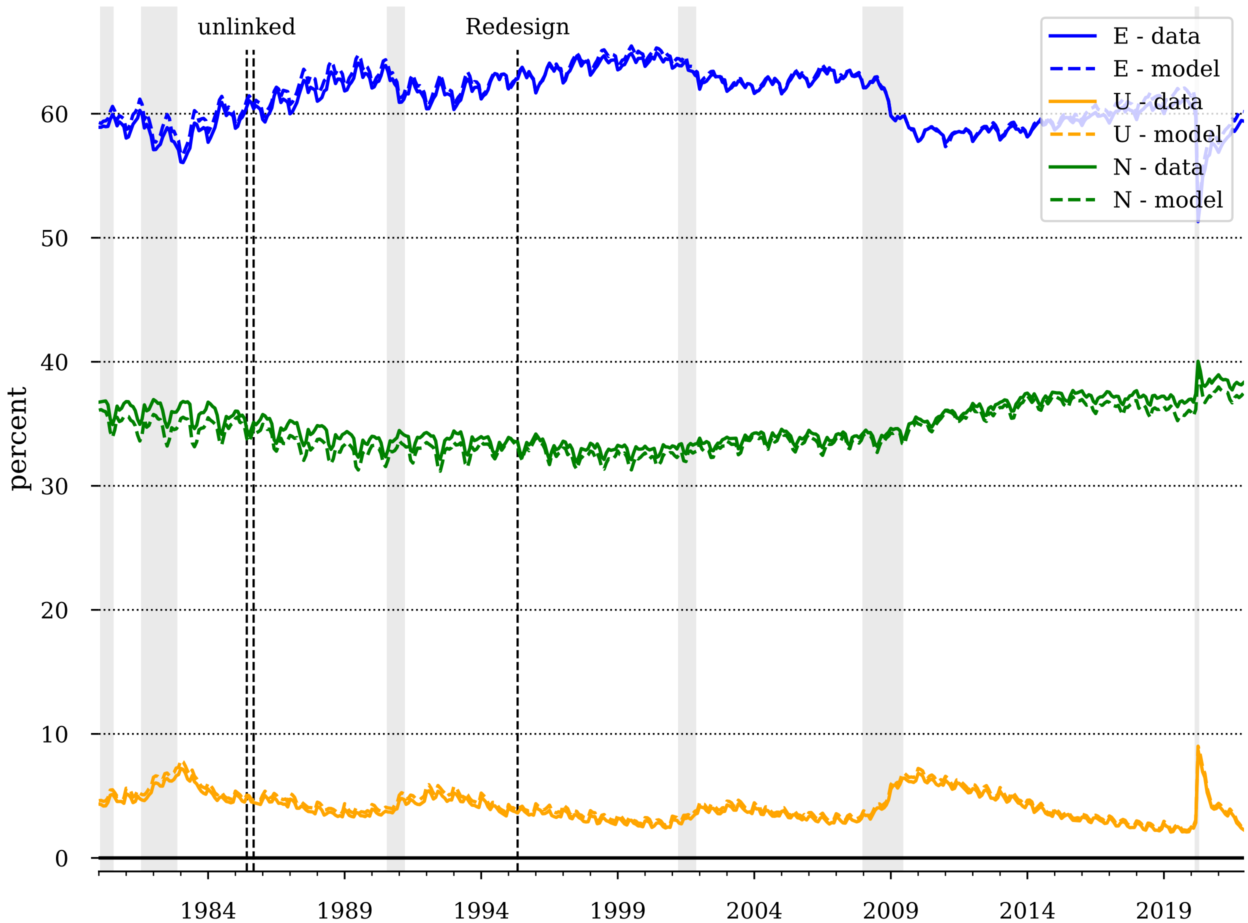
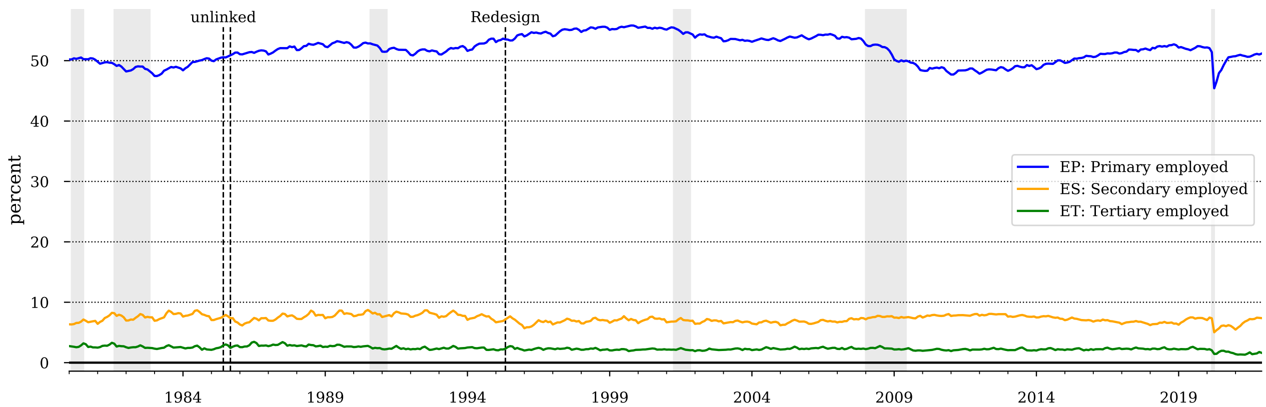
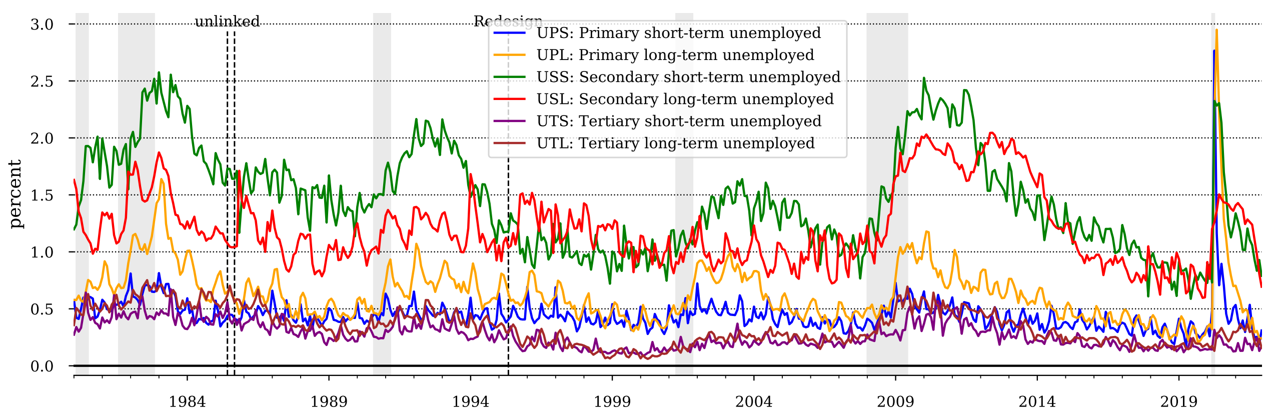


Figure B.1: Actual and fitted shares of population in E, U, and N.

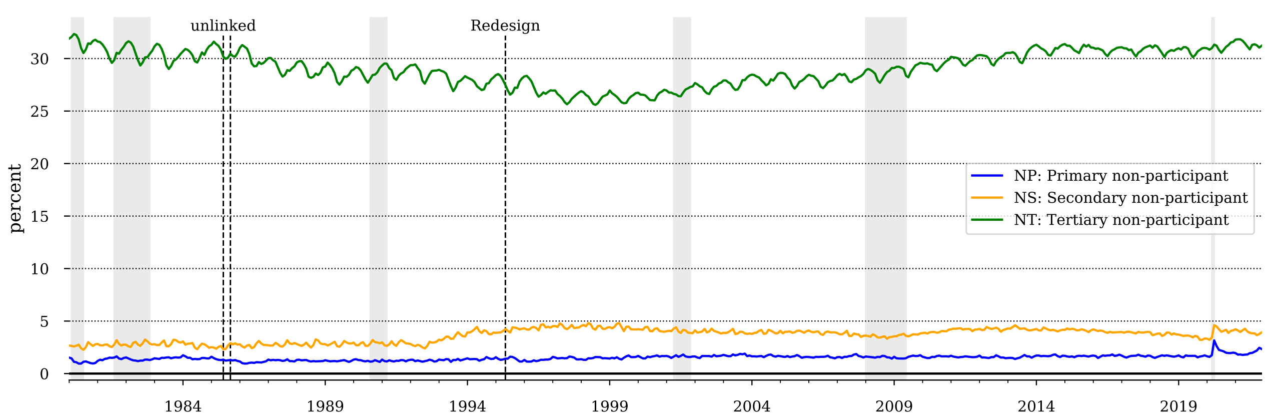
Source: CPS and authors' calculations.
Notes: Data and model do not match because of use of 12-month weights in model and imputation of missing values in model estimates.



(a) Employed



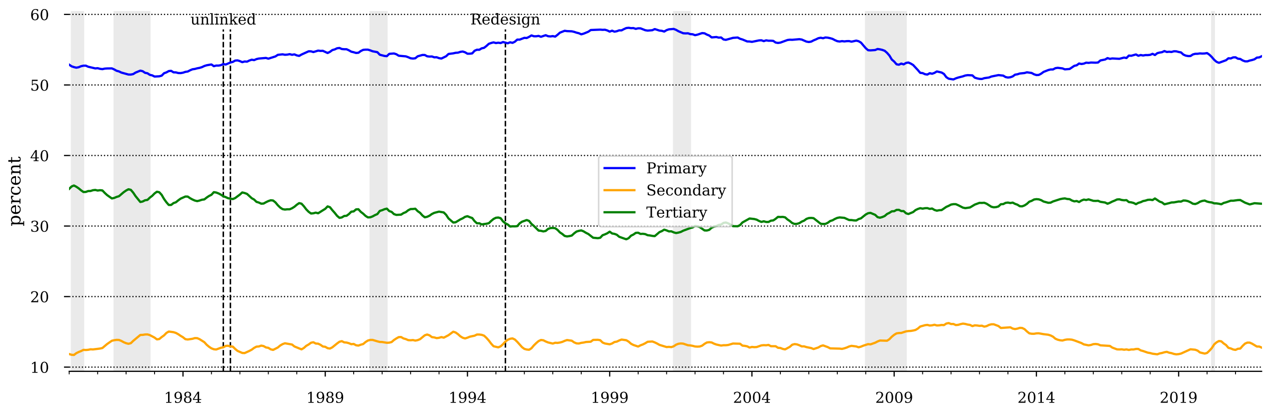
(b) Unemployed



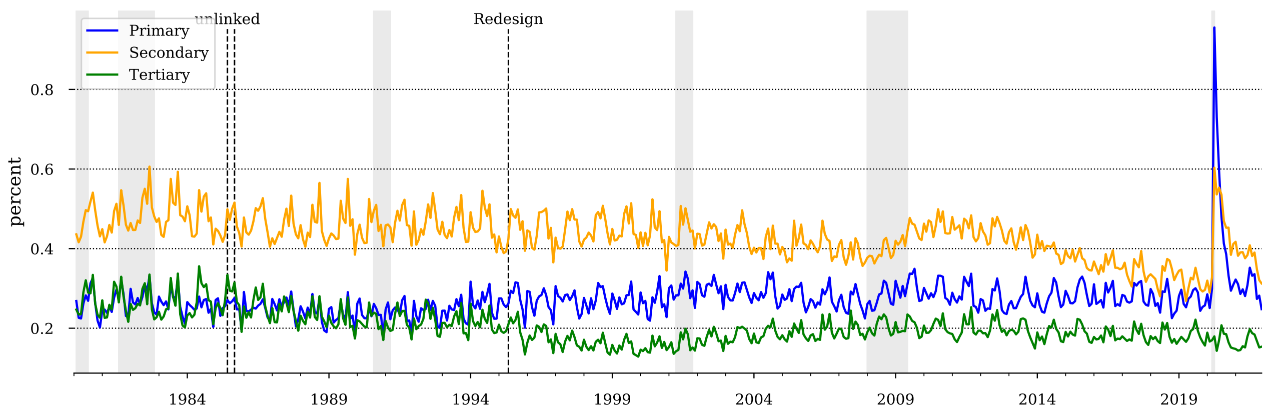
(c) Non-participant

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

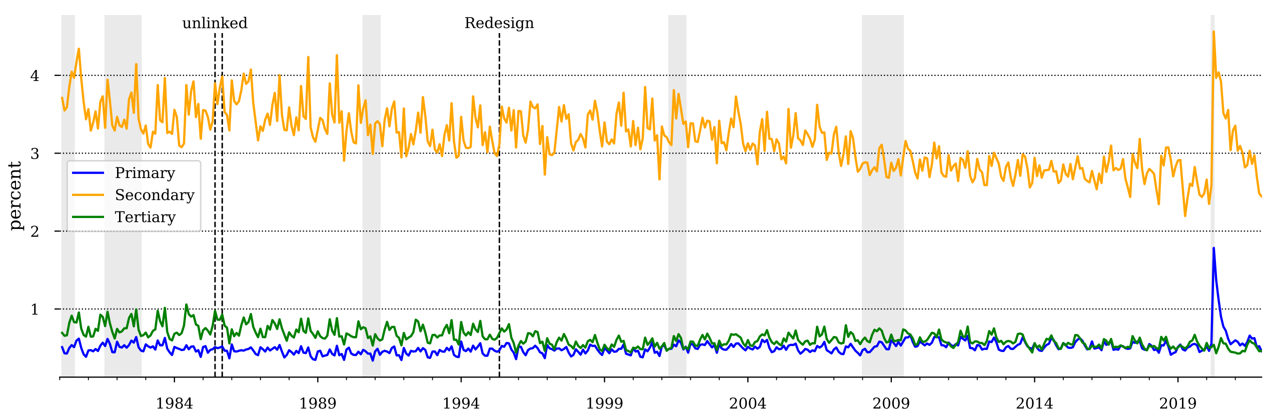
Source: CPS and authors' calculations.
Notes: "unlinked" is period where household identifiers in CPS are scrambled and 4-8-4 respondent histories are incomplete and "redesign"is the 1994 for which respondent histories are incomplete and the questions used to construct the emissions are changed.



(a) Share of population



(b) Annualized monthly flows per capita

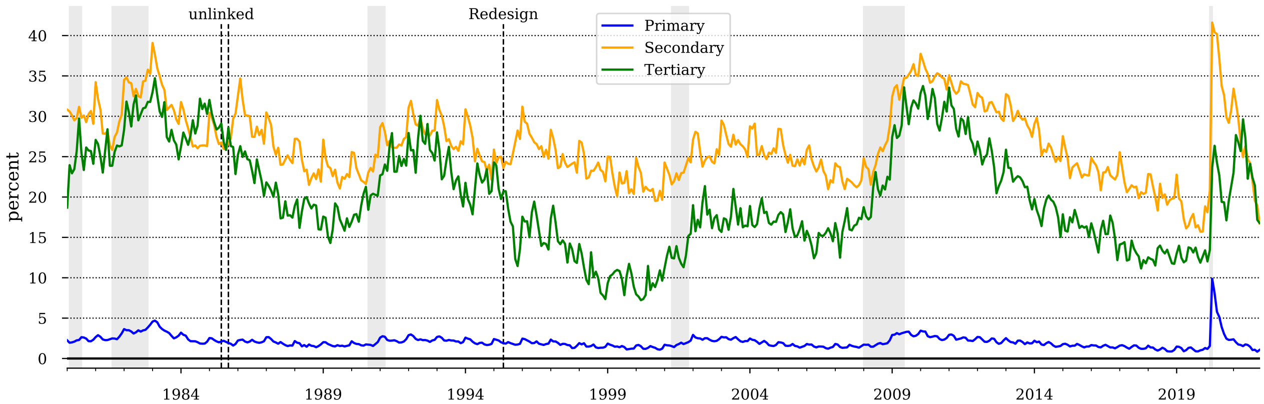


(c) Annualized monthly flows per person in segment

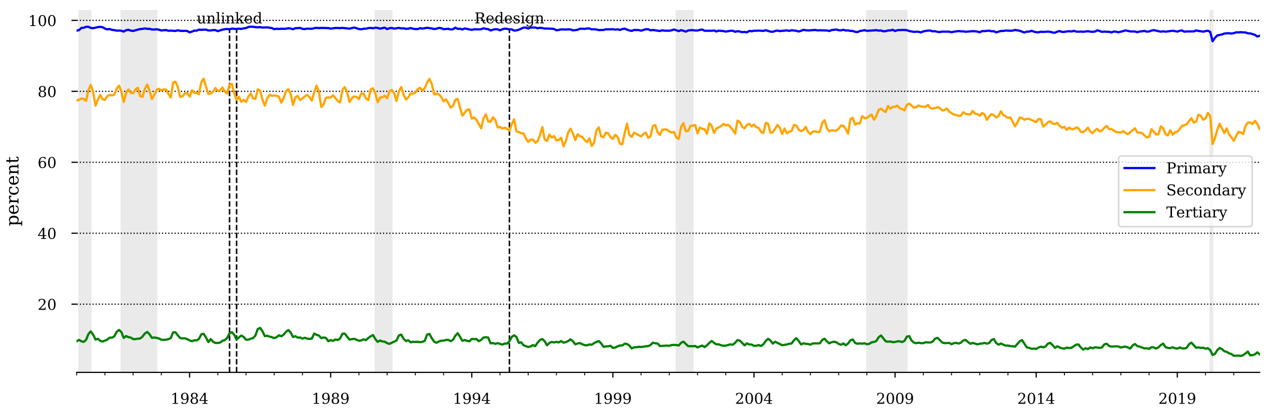
Figure B.3: Shares of population and flows per person by labor market segment.

Source: CPS and authors' calculations.

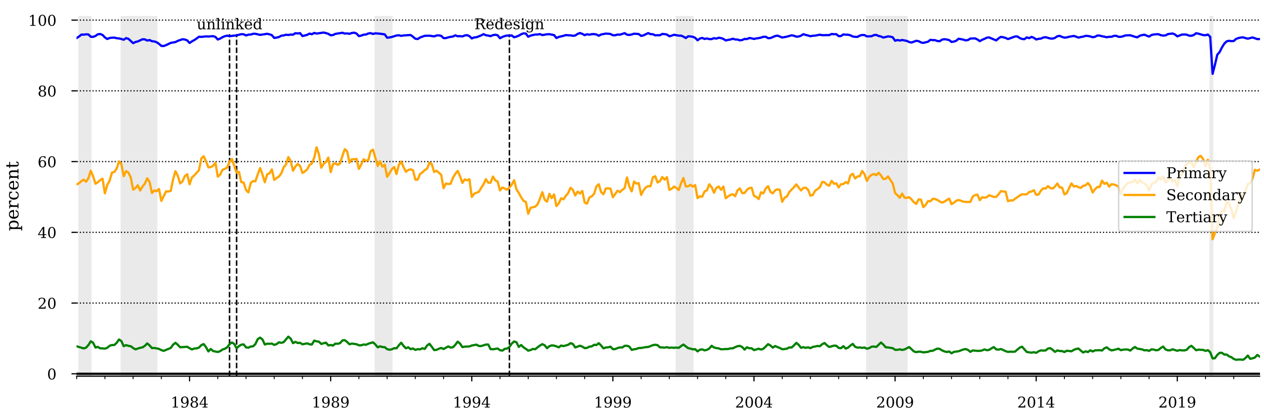
Notes: "unlinked" is period where household identifiers in CPS are scrambled and 4-8-4 respondent histories are incomplete and "redesign"is the 1994 for which respondent histories are incomplete and the questions used to construct the emissions are changed.



(a) Unemployment rates



(b) Labor force participation rates



(c) Employment-population ratios

Figure B.4: 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.

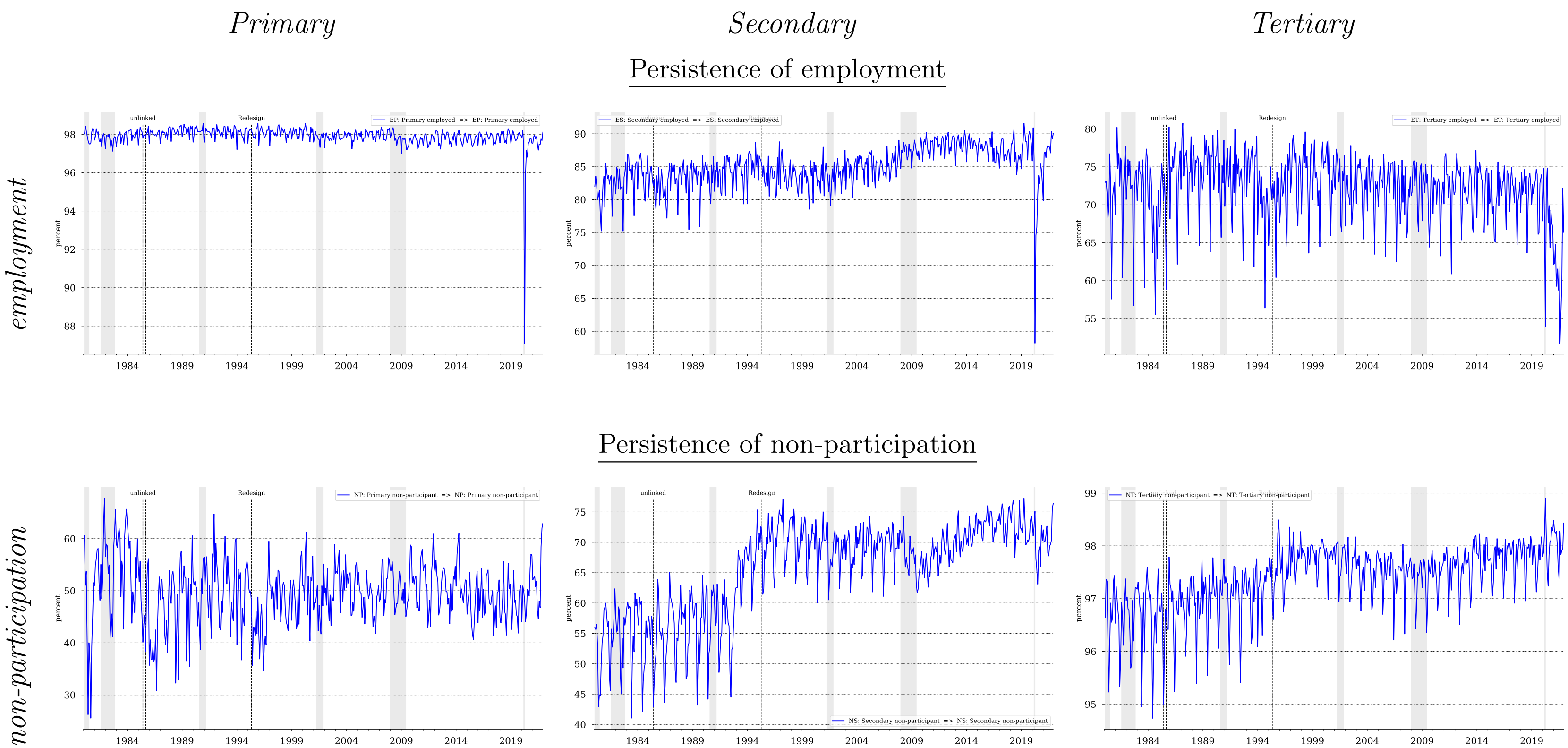


Figure B.5: Persistence of employment and non-participation

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.

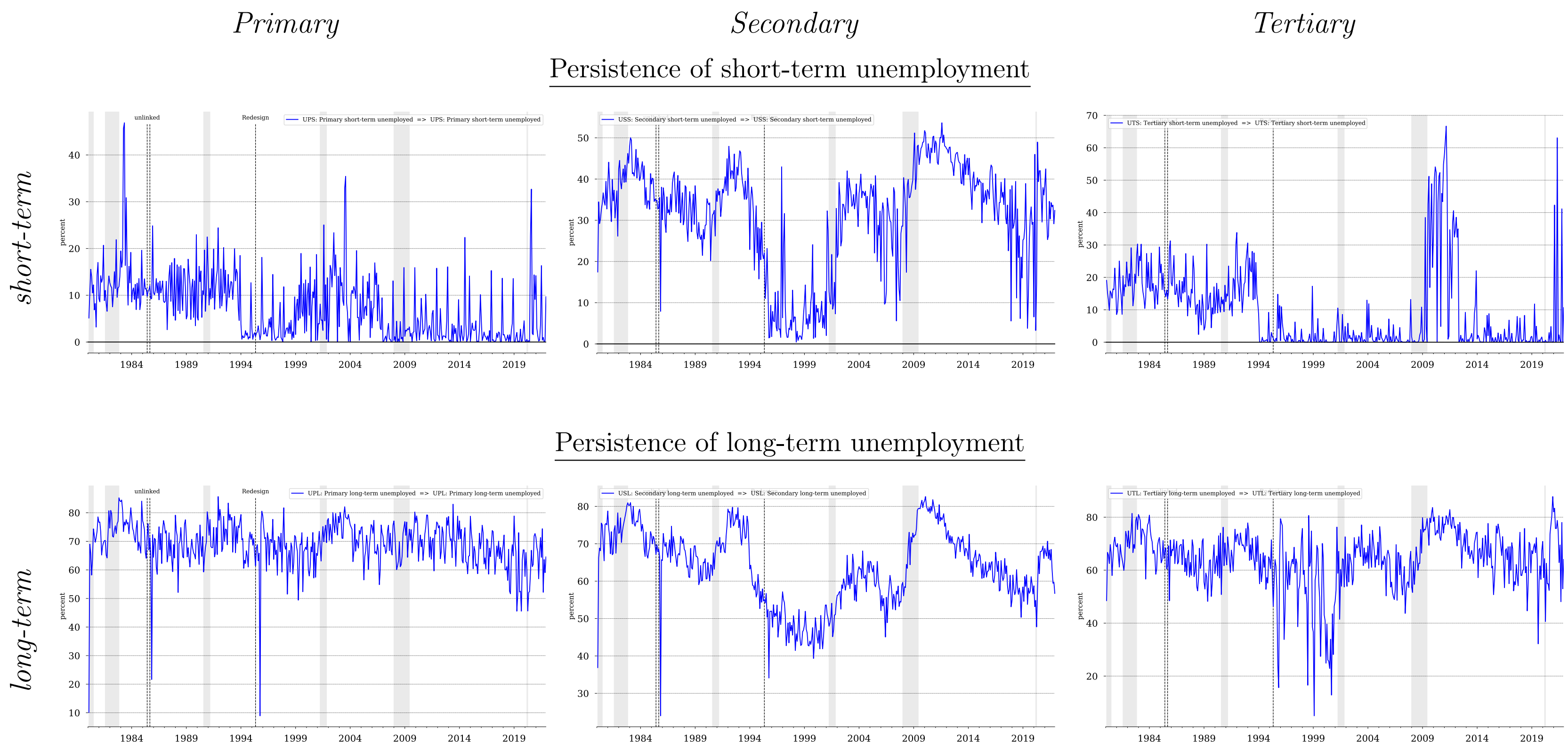


Figure B.6: Persistence of short- and long-term unemployment

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.

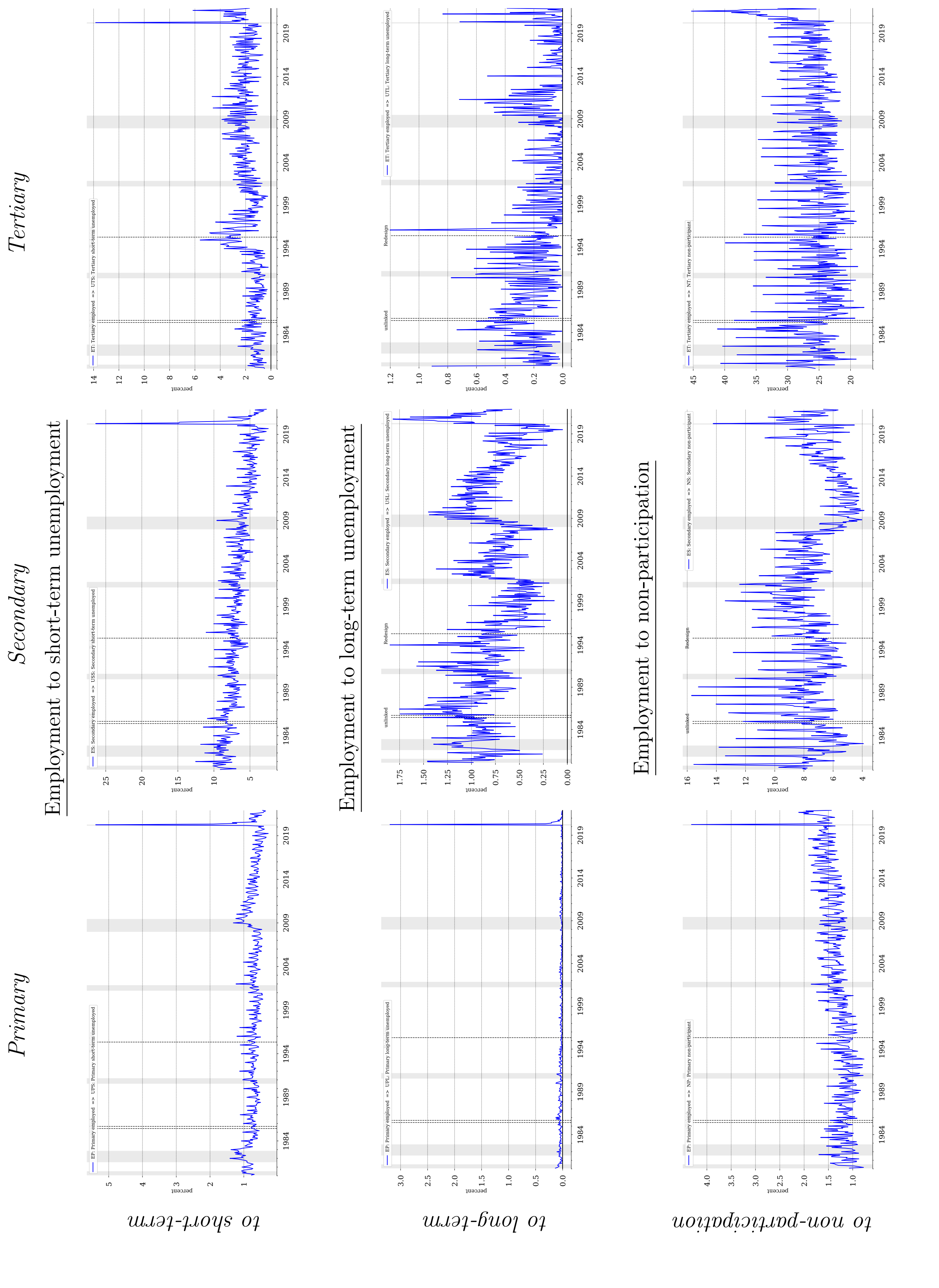


Figure B.7: Estimated transition probabilities from employment

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

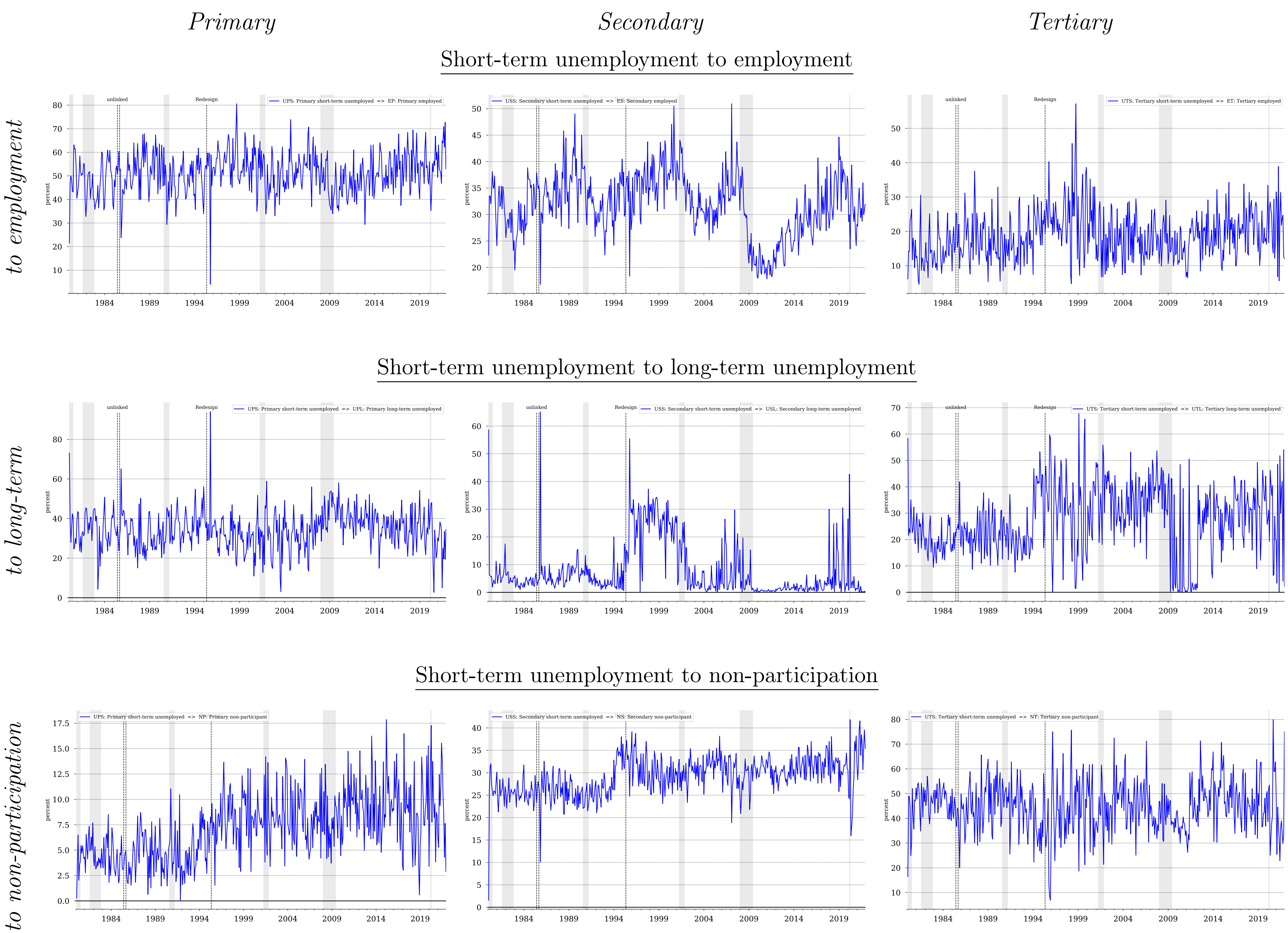


Figure B.8: Estimated transition probabilities from short-term unemployment

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

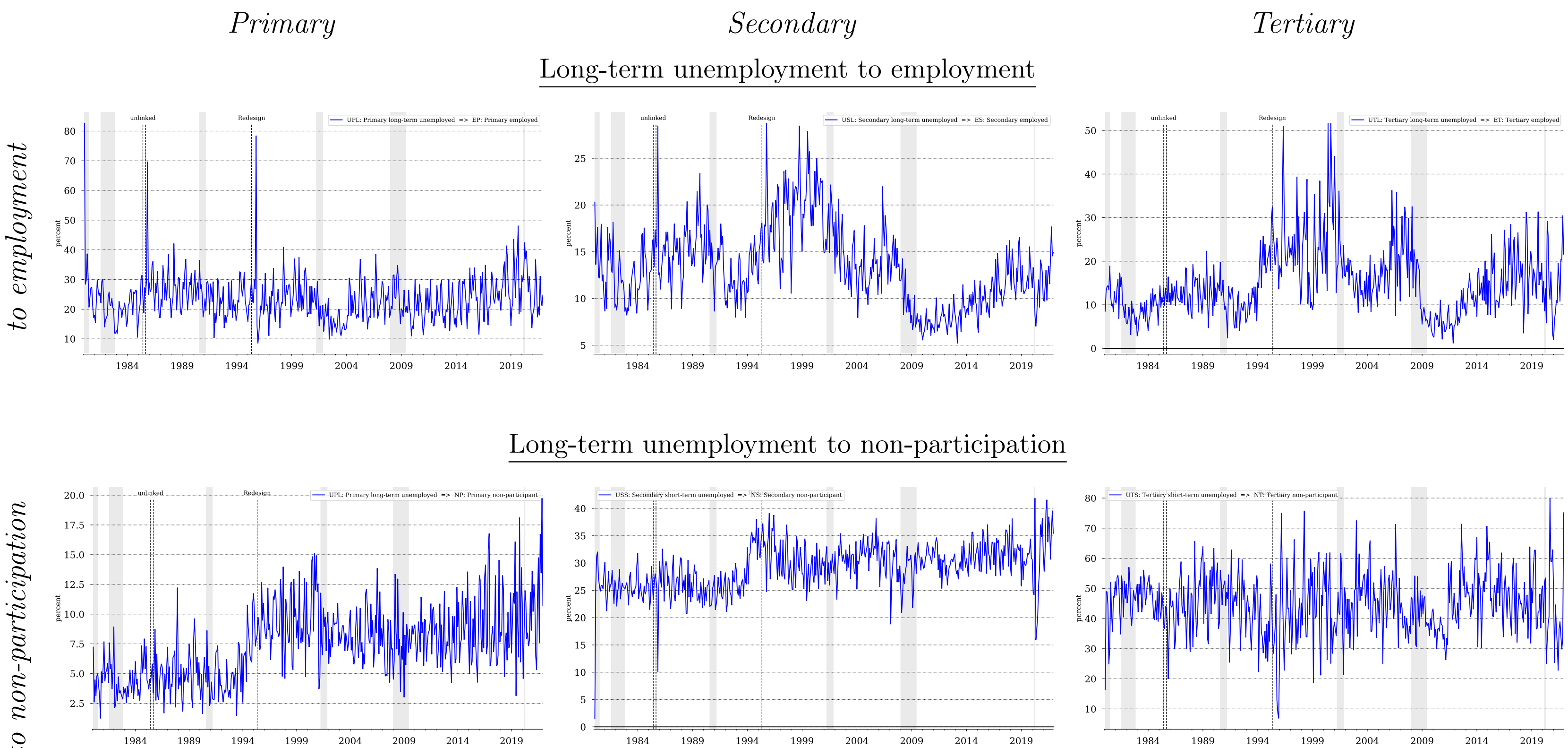


Figure B.9: Estimated transition probabilities from long-term unemployment

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.

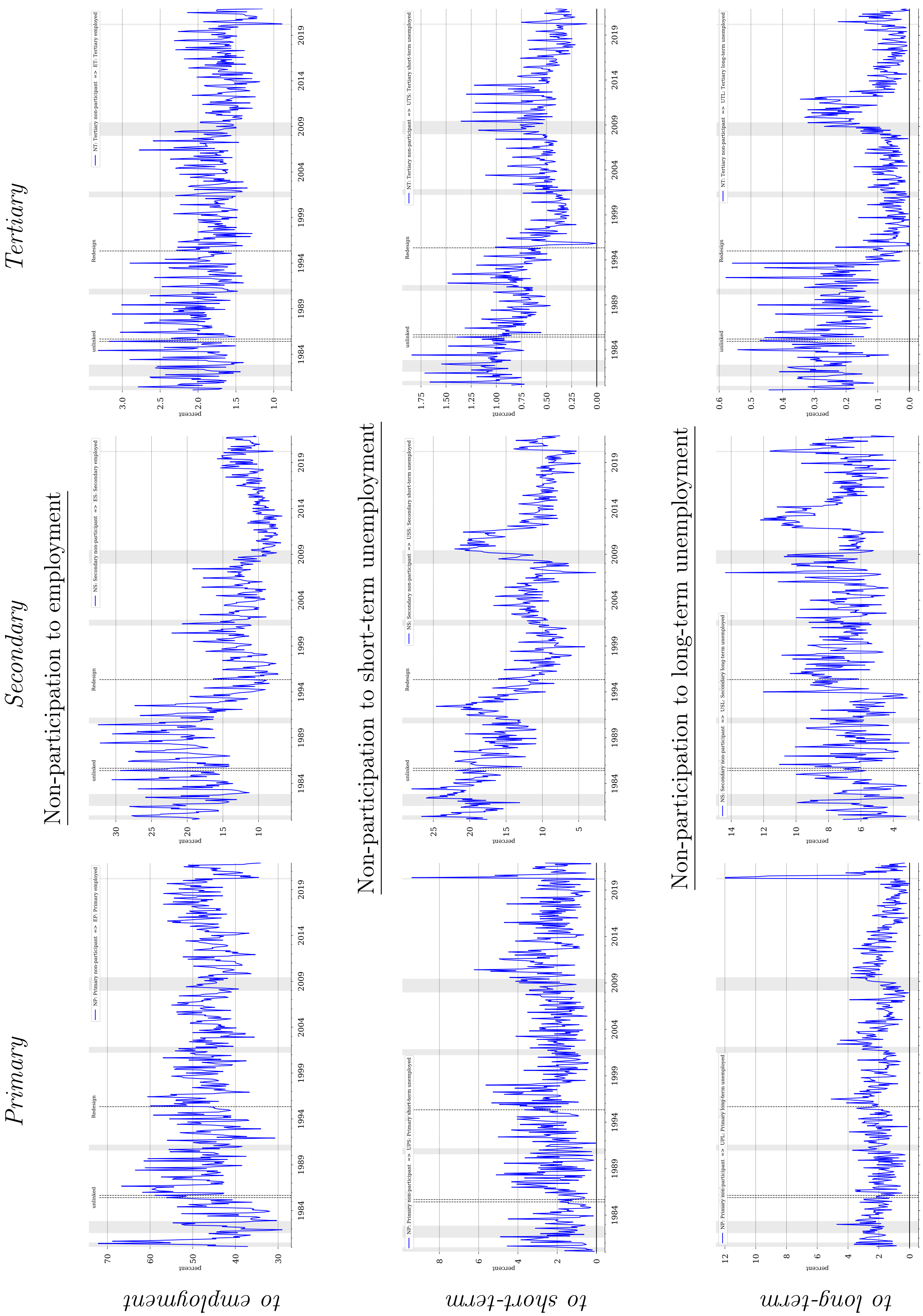
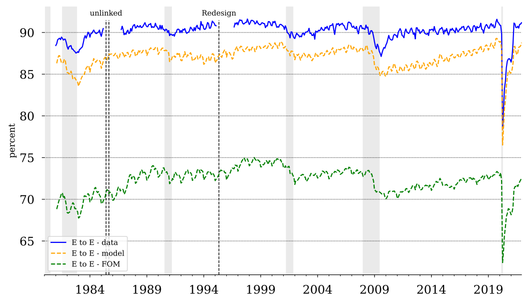
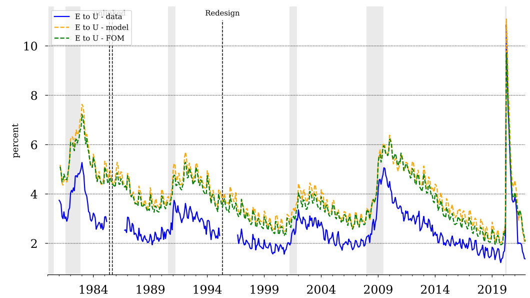


Figure B.10: Estimated transition probabilities from non-participation

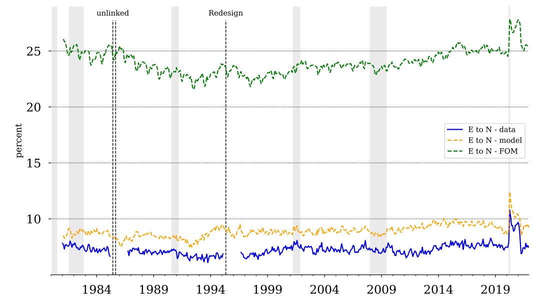
Source: CPS and authors' calculations. Monthly observations. Not seasonally adjusted.



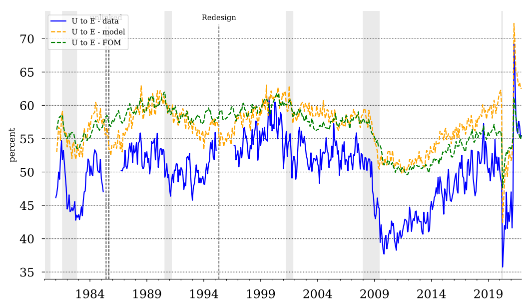
(a) $E \rightarrow E$



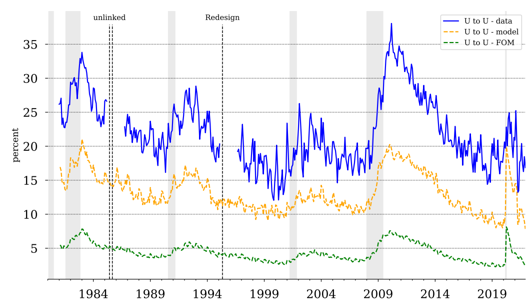
(b) $E \rightarrow U$



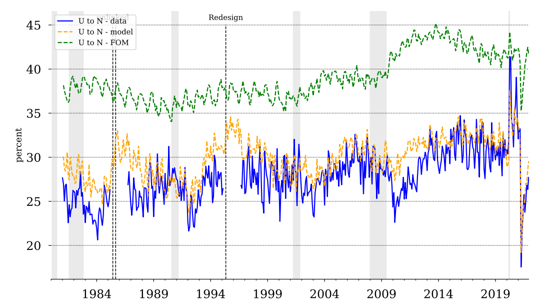
(c) $E \rightarrow N$



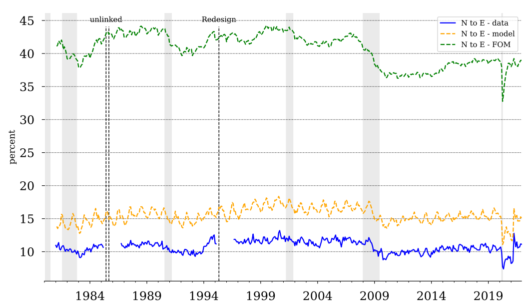
(d) $U \rightarrow E$



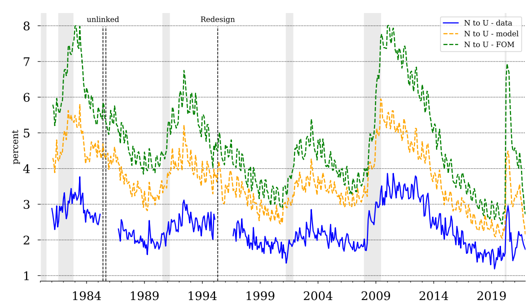
(e) $U \rightarrow U$



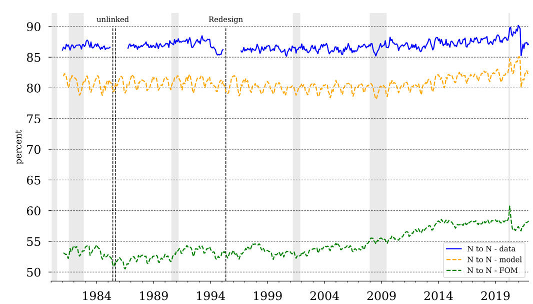
(f) $U \rightarrow N$



(g) $N \rightarrow E$



(h) $N \rightarrow U$



(i) $N \rightarrow N$

Figure B.11: Actual and estimated 12-month transition probabilities between E, U, and N

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

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. FOM time series are for First-Order Markov model with three states.