I Z A Institute of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15955

The Nature of Long-Term Unemployment: Predictability, Heterogeneity and Selection

Andreas I. Mueller
Johannes Spinnewijn

FEBRUARY 2023

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 15955

The Nature of Long-Term Unemployment: Predictability, Heterogeneity and Selection

Andreas I. Mueller
UT Austin, NBER, CEPR and IZA

Johannes Spinnewijn
LSE and CEPR

FEBRUARY 2023

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Nature of Long-Term Unemployment: Predictability, Heterogeneity and Selection*

This paper studies the predictability of long-term unemployment (LTU) and analyzes its main determinants using rich administrative data in Sweden. Compared to using standard socio-demographic variables, the predictive power more than doubles when leveraging the rich data environment. The largest gains come from adding job seekers' employment history prior to becoming unemployed. Applying our prediction algorithm over the unemployment spell, we show that dynamic selection into LTU explains at least half of the observed decline in job finding. While the within-individual declines are small on average, we find substantial heterogeneity in the individual-level declines and thus reject the commonly used proportional hazard assumption. Applying our prediction algorithm over the business cycle, we find that the cyclicality in average LTU risk is not driven by composition but rather by within-individual cyclicality and that individual rankings are relatively persistent across years. Finally, we evaluate the implications of our findings for the value of targeting unemployment policies and how these change over the unemployment spell and the business cycle.

JEL Classification: E24, J64

Keywords: long-term unemployment, heterogeneity, selection, duration dependence, business cycle, targeting

Corresponding author:
Andreas I. Mueller
Columbia Business School
Uris Hall 824
3022 Broadway
New York, NY 10027
USA
E-mail: amueller@columbia.edu

* We thank Peter Fredrikson, Francois Gerard, Gregor Jarosch, Maximilian Kasy, Philipp Kircher, Jonas Kolsrud, Camille Landais, Alan Manning, Rachel Meager, Peter Nilsson, Oskar Nordstrom Skans, Torsten Persson, Jean-Marc Robin, Jeff Smith, Josef Zweimüller and participants at seminars and workshops at Carnegie Mellon University, Cornell University, Dallas Fed, HEC Montreal, IIES, LSE, Stanford University, UCL, University of Essex, University of Zurich, Uppsala University, and UT Austin for their comments and suggestions and Sebastian Ernst, Miguel Fajardo-Steinhauser, Diego Ferreras Garrucho and Ada Zakrzewska for their excellent research assistance.

1 Introduction

Tackling the issue of long-term unemployment (LTU) is a long-standing challenge for labor market policy (Machin and Manning [1999]). Spells of prolonged joblessness are associated with worse economic outcomes, including lower job-finding rates, lower re-employment wages and less stable jobs. These concerns are heightened during bad economic times when the incidence of LTU rises dramatically, such as for example during the Great Recession in 2007-09 (Kroft et al. [2016]). The long-term unemployed are central in the policy debate on how to set unemployment benefits (Kolsrud et al. [2018] and are typically the target for active labor market programs (OECD [2019]).

Despite the policy-relevance of the issue, the sources of LTU are still subject to debate. A first source of contention centers around the observation that long-term unemployed workers typically exhibit lower job-finding rates. One interpretation is that likelihood of finding a job declines at the individual level of the unemployment spell and thus long-term unemployment is a trap, which is difficult to escape. An alternative interpretation, however, is that long-term unemployed workers are inherently different and generally less likely to find jobs throughout their spell. This dynamic selection thus could account for the lower observed job-finding among the long-term unemployed. A second unsettled debate concerns the rise of long-term unemployment in recessions. This rise may be due to a change in the composition of the pool of unemployed workers or be driven by a decrease in the within-individual job-finding chances during recessions for all or at least some unemployed workers.

Traditionally, research has found little role for heterogeneous job-finding rates in explaining the dynamics of job finding over the spell of unemployment or over the business cycle, using observable characteristics in survey data (e.g., Baker [1992]; Kroft et al. [2016]). This leaves an important potential role for within-individual dynamics, either over the unemployment spell or over the business cycle. Most recently, however, a series of papers contest the limited role of heterogeneity across unemployed job seekers, using a range of different methods (e.g., Hall and Kudlyak [2019], Gregory et al. [2021], Mueller, Spinnewijn and Topa [2021], Alvarez et al. [2022], Ahn et al. [2022]). This naturally raises the question whether studies, which found no or little role for heterogeneity, are limited by the range of observable characteristics typically available in survey data.

This paper uses administrative data on the universe of unemployment spells from Sweden for the years 1992-2016 and combines it with rich and detailed information on unemployed workers' characteristics, which includes – in addition to standard socio-demographics – information on income, employment and benefit histories, as well as information on prior employers, occupation, asset portfolios and IQ scores. We leverage the rich data environment to study the predictability of long-term unemployment risk and the heterogeneity in job-finding rates. We exploit the predictive power of our prediction model to study the dynamics of job finding over the unemployment spell and the business cycle. We first revisit the importance of heterogeneity in job finding across job seekers and the role of the resulting selection for the dynamics of the average job finding over the unemployment spell and the business cycle. In addition, the large-scale nature of our data both in terms of number of individuals and years allows us to study in detail the potential heterogeneity in *individual-level dynamics* over both dimensions.

We first develop a simple conceptual framework to infer ex-ante heterogeneity in job-finding rates from measures of predictive power from a prediction model. We show that the R-squared of the predicted job-finding rate with the actual job finding in a hold-out sample provides a lower bound for the variance in job-finding *probabilities*. In a second step, building on Mueller, Spinnewijn and Topa [2021], we distinguish between transitory and persistent heterogeneity in job-finding rates and show that the persistent heterogeneity can be inferred from the covariance between the predicted job-finding rate with actual job finding. Applying these results to the dynamics over the spell of unemployment, we show how our measure of persistent heterogeneity within an unemployment spell can be used to identify a lower bound for the contribution of dynamic selection to the observed decline in job finding. We also highlight the complementary value between our approach of using observable characteristics and uncovering *observable* heterogeneity vs. the approach of using multiple unemployment spells and uncovering *fixed* heterogeneity across spells (Honoré [1993], Alvarez et al. [2022]).

For the empirical analysis we employ standard machine learning (ML) techniques, training a prediction model on a training sample and then evaluating the predictive power in a hold-out sample. The strict separation between these two samples is necessary to deal with the problem of over-fitting in a data-rich environment. Our prediction model is an Ensemble ML model, which uses a weighted average of LASSO, Gradient Boosted Decision Trees and Random Forest [Einav et al., 2018]. We focus on predicting the probability of finding a job within 6 months from the start of the unemployment spell. We define this job-finding probability as one minus the probability to be still unemployed six months into the spell, which is a standard measure of long-term unemployment risk in labor-force statistics and a typical target in risk profiling models used by Public Employment Services. Our main analysis predicts the job-finding rates at the start of the spell in the year 2006, when we have all the different data sets available. We then extend our prediction exercise to other unemployment durations and to the years between 1992 and 2016, using a baseline set of characteristics, which are available consistently for each year in our sample and for the universe of unemployed workers in Sweden.

Our prediction analyses provide three sets of empirical results shedding new light on the nature of long-term unemployment risk and its determinants.

First, we find substantial heterogeneity in long-term unemployment risk. The predictive power, as measured by the hold-out sample R-squared between predicted job-finding rate and actual job finding, is more than twice as large than when adding income, employment and benefit histories relative to a prediction model that only uses basic socio-demographic variables such as education, age, gender, marital status, citizenship and the number and age of children. The prior employment history, even if only available for one or a few years, is particularly predictive. The further gains in predictive power when using information on occupation, assets, IQ scores, UI benefits, etc., which are only available for limited samples or years, are modest, suggestive of the saturation of our baseline model. However, using the sample of job seekers with repeated unemployment spells, we still find a substantial role for *unobserved* heterogeneity relative to the *predictable* heterogeneity, at least for the heterogeneity that is permanent across unemployment spells for this sample. We also find that the predictive power of a linear model is nearly as high compared to our baseline prediction model. This shows that the additional predictive power comes from the data rich environment rather than non-linearities exploited by the ML algorithm and that the risk of over-fitting using the universe of unemployment spells is limited.

Second, we apply our results on the predictability of job finding to their dynamics over the unemployment spell. Similar to other countries, the observed job-finding rate in Sweden declines strongly with the duration of an unemployment spell. In 2006, unemployed job seekers' 6-month job-finding rate was 70% at the start of the unemployment spell but then declined to 55% at 6 months. We repeat the prediction exercise at different durations for ongoing unemployment spells and infer the persistent heterogeneity in job finding by computing the covariance between predicted job finding from 6 to 12 months in the unemployment spell with actual job finding from 0 to 6 months. We find that nearly three quarters of the predictable heterogeneity in job finding is persistent. Applying the decomposition in our conceptual framework, the persistent heterogeneity accounts for a decline of job finding from 70% to 62.5% at 6 months, implying that dynamic selection accounts for 49% of the observed decline in job finding, or more, given the lower bound nature of our prediction exercise.

Our analysis thus suggests that the within-individual decline in job finding over the unemployment spell is relatively small, but our prediction exercises also allow us to investigate heterogeneity in these individual-level declines. We thus provide an empirical test for the key assumption in proportional hazard models, that job-finding rates decline at the same rate across job seekers. Specifically, we relate the predicted probability of job finding at 0, 6 and 12 months to the duration of unemployment and estimate a predicted decline over the unemployment spell for each individual in our data. This procedure uncovers substantial heterogeneity in the decline of job finding over the unemployment spell across individual, even when adjusting for the sampling error inherent in this exercise. Our results thus reject the assumption in the most commonly used model of job search. We also find that the individual-level declines are strongly negatively correlated with the job finding at the start of the spell. This convergence in within-individual job-finding rates further compresses the heterogeneity in job finding over the unemployment spell, above and beyond the dynamic selection.

Third, we turn to the time series of our sample and estimate our prediction model for each year from 1992 to 2016. We find strong persistence in the predictive power across years as the R-squared of actual job finding in 2006 with the predicted job-finding rate from the model of a different year remains high even if a distant year is used. The distribution of predicted job-finding risk, however, changes over the business cycle. Prior research has found that compositional changes in the pool of unemployed cannot account for the increased LTU risk in recessions [Baker, 1992; Kroft et al., 2016]. Using richer data, we still reject this so-called heterogeneity hypothesis, as we find that the compositional changes in the pool of unemployed do not translate into higher LTU risk in recessions. Unemployed workers are thus exposed to substantial changes in LTU risk over the business cycle. In parallel to our analysis on duration dependence, we also assess the heterogeneity in within-individual cyclicality of LTU risk, by relating the predicted job-finding rate for each individual to the unemployment rate in each year. In comparison with the duration-dependence analysis, we find only modest differences in the cyclicality in job finding. Recessions affect the employment prospects of both job seekers with higher and lower job-finding chances, but disproportionally hurt workers with lower education and income.

We conclude our analysis by gauging the implications of our empirical findings for the targeting of active labor market programs (ALMP). While it is beyond the scope of this paper to estimate treatment effects of these programs, our starting point is the observation that many countries either use long-term unemployment or a measure of the risk of long-term unemployment as a criterion for the assignment to these policies. The high predictability of LTU risk suggests that we can assign

ALMPs earlier and target them better. Indeed, the predicted LTU risk of the top 20 percent is nearly twice as high as the average LTU risk. These top 20% have the same probability of becoming long-term unemployed from the start of the spell as the average unemployed person 3 months into the unemployment spell. The gains of using prediction algorithms for targeting ALMPs are also larger in good times than in periods of high unemployment, because the distribution of LTU risk is more compressed in recessions. At the same time, given the high persistence of the prediction algorithm over time, there is little gain in adjusting the algorithm over the business cycle.

Related Literature. A long and distinguished literature has studied long-term unemployment and its causes and consequences. Our analyses aim to contribute to three important strands in this literature.

First, the duration-dependence in job finding and the importance of heterogeneous job finding therein has received wide attention, starting with the seminal work of Lancaster [1979] and Heckman and Singer [1984]. An important takeaway from this early literature is that identification is sensitive to functional and distributional form assumptions. A few recent papers have tried to overcome these challenges. Building on Honoré [1993] and his identification argument using multiple spell data, Alvarez et al. [2022] have developed and implemented a novel approach that estimates heterogeneity in job finding without relying on proportional hazards. Mueller, Spinnewijn and Topa [2021] instead focus on predictable heterogeneity from job seekers' reported beliefs about their own job-finding risk.[1] Both of these papers find substantial heterogeneity in job-finding rates, implying substantial dynamic selection over the unemployment spell, just like we do.[2] In comparison with Alvarez et al. [2022], our paper does not rely on identifying assumptions regarding the persistence of job-finding types across unemployment spells, but shows how to evaluate this empirically. In our data, we show that fixed heterogeneity across two spells is about equally as large as the heterogeneity uncovered by our prediction exercise. The nature of the heterogeneity, however, is different and thus we view the two approaches as highly complementary. In comparison with Mueller et al. [2021], we use rich observable characteristics from administrative data, including income, employment and benefit histories, rather than predictive beliefs. This gives us the statistical power to study heterogeneity in the dynamics of job-finding rates over the unemployment spell.

Second, we apply our approach to long time series of over 20 years of data, allowing us to speak to issues related to the business cycle. As already mentioned, prior work has found little role for observable heterogeneity for explaining the cyclicality of job finding or unemployment duration (e.g., Baker [1992]; Krueger et al. [2014]; Kroft et al. [2016]). Still, a number of papers have found compositional changes in the pool of unemployed over the business cycle. For example, Mueller [2017] documents changes in the composition of the unemployed in terms of prior wage but these compositional shifts have no bearance for the heterogeneity hypothesis because high- and low-wage workers have similar

[1] See also Arni et al. [2014] who look at the role of personality traits, beliefs and other behavior variables for predicting long-term unemployment.

[2] There is also direct evidence that call-back rates decline with unemployment duration (Eriksson and Rooth [2014]; Kroft et al. [2013]), though Farber et al. [2016] find no effects for older workers. Jarosch and Pilossoph [2018] also show that this may not necessarily translate into declining job-finding rates. Another related paper is Morchio [2020] who finds in data from the NLSY79 that workers with different unemployment histories in their 20s face different job-finding rates later in life.

job-finding rates. Elsby et al. [2015] show that the heterogeneity hypothesis matters for transition rates between unemployment and out of the labor force, but not for U-E transition rates. In contrast, Ahn and Hamilton [2020] estimated the heterogeneity in the context of a mixed proportional hazard model and find an important role of heterogeneity for cyclical movements in job finding. Hall and Schulhofer-Wohl [2018] find some evidence in favor of the heterogeneity hypothesis based on the reason for unemployment (layoff, quit, etc.) and show that the pool of job seekers sorts toward low-job-finding types in recessions. Overall, these papers typically rely on a small number of observable characteristics available in labor force survey data. Given the richness of our data, we believe it is important to re-evaluate the heterogeneity hypothesis of the cyclicality of long-term unemployment risk. While some earlier work has studied the heterogeneity in the cyclicality job finding by groups of workers characterized by their average wages and hours worked (Bils et al. [2012], Mueller [2017]), we go beyond this by studying heterogeneity in within-individual cyclicality.

Our paper more generally relates to recent papers that use machine learning and related techniques to classify workers into types based on labor force histories. Gregory et al. [2021] use a k-means algorithm to cluster workers based on the similarity of their employment histories in administrative data. Hall and Kudlyak [2019] and Ahn et al. [2022] infer worker types from their labor market transitions in labor force survey data. These papers generally find large amounts of heterogeneity in job finding, though they do not separately identify ex-ante heterogeneity and true duration dependence in job finding. Our paper is also complementary to the work studying the impact of job loss on earnings and wages in particular, focusing on the duration-dependence [Schmieder et al., 2016], the cyclicality [Schmieder et al., 2022] and predictable sources of heterogeneity more generally [Bertheau et al., 2022].

Finally, while the duration-dependence of unemployment benefits has received wide attention in the literature (e.g., Shimer and Werning [2008], Schmieder et al. [2012], Kolsrud et al. [2018]), the focus on the LT unemployed in the assignment of ALMPs has received less attention. A few theoretical papers have studied the optimal timing of ALMPs during the unemployment spell (e.g., Pavoni and Violante [2007], Spinnewijn [2013]). Even less empirical papers have studied the use of risk profiling models. Most notably, Black et al. [2003] use profiling scores and capacity constraints for identifying the effects of employment and training services in Kentucky.[3] A recent report [OECD, 2019] documents the growth in the use of profiling models in various OECD countries, with the most common outcome in profiling models being long-term unemployment at 6 or 12 months, but this has received little attention in spite of the large and growing literature evaluating active labor market programs (see Card et al. [2017] for a meta-analysis).

This paper proceeds as follows. Section 2 provides a conceptual framework that relates predictability of job finding to ex-ante heterogeneity in job finding and duration dependence. Section 3 presents the data and prediction model and analyses the predictability of job finding. Section 4 analyses the predictability of job finding over the spell of unemployment and quantifies its role for dynamic selection. Section 5 analyses the predictability of job finding over the business cycle. Section 6 draws implications of our findings for the targeting of unemployment policies. Section 7 concludes.

[3] Berger et al. [2001] also provide a general review of issues related to profiling and targeting of government services.

2 Conceptual Framework

We present a brief conceptual framework of job finding to show how to identify heterogeneity in job finding probabilities and how this heterogeneity is crucial for understanding the dynamics in job finding and the role of selection or sorting effects in particular.

We model the job-finding probability for individual i as

$$T_i = T(X_i) + \varepsilon_i,$$

where $T(X_i) = E(T_i|X_i)$ is the individual's type based on observable characteristics X_i. We will introduce a dynamic dimension (time or duration) further below. ε_i captures the unobservable heterogeneity, orthogonal to the observable characteristics, $E(\varepsilon_i|X_i) = 0$. While we cannot observe an individual's job-finding probability, we can observe whether or not an individual has found a job. The realization of the probability is

$$F_i = \begin{cases} 1 & \text{with prob } T_i \\ 0 & \text{otherwise.} \end{cases}$$

We posit a prediction model $F(.)$ such that

$$F_i = F(X_i) + e_i.$$

where e_i is a prediction error. Note that the source of prediction error is both sampling error as well as randomness in the outcome variable, which is the random realization of the underlying probability.

Ex-Ante Heterogeneity. We are interested in evaluating the ex-ante heterogeneity in job finding types $var(T_i)$ and the challenge is to separate this from the heterogeneity in ex-post outcomes, $var(F_i) = E(T_i)(1 - E(T_i))$. We can bound the ex-ante heterogeneity by using the variation in outcomes that is predictable ex-ante based on observables. We can state:

Proposition 1. *The hold-out sample R-squared between the prediction and the observed realization of the probability provides a lower-bound estimate of the variance in (observable) types relative to the variance in realizations:*

$$R^2(\hat{F}_i, F_i) \leq \frac{var(T(X_i))}{var(F_i)} \leq \frac{var(T_i)}{var(F_i)}. \tag{1}$$

Proof. See Appendix B.1.

Note that the R-squared depends on the covariance between the predictions and realizations scaled by their respective variances, $R^2(\hat{F}_i, F_i) = \frac{cov(\hat{F}_i, F_i)^2}{var(\hat{F}_i)var(F_i)}$. By evaluating the covariance in a hold-out sample rather than in the sample used for estimating the prediction model, we avoid any confounding correlation between the sample error underlying individual predictions and their specific outcomes. The main argument in the proof relies on the Cauchy-Schwarz inequality, which implies the lower bound. This lower bound becomes tight when the predictor is unbiased, i.e. $E(\hat{F}_i|X_i) = T(X_i)$. The hold-out sample covariance of the prediction and the observed realization then equals the variance in observable types, $cov(\hat{F}_i, F_i) = var(T(X_i))$ (see Proposition 3 in Appendix B.1).

Job Finding Dynamics. The heterogeneity in job finding determines how much the dynamics in average job finding can be driven by selection rather than dynamics at the individual level. Our empirical analysis will focus on two types of dynamics which have received wide attention in the literature: How do job-finding chances change over the unemployment spell? And how do job-finding chances change over the business cycles?

We conceptualize this by introducing the dynamic dimension δ and considering changes in job finding along this dimension. Hence, δ could denote calendar time, the duration of the ongoing unemployment spell or the stage of the business cycle. Changes in the average job finding along this dimension can be driven by within-individual changes in job finding or by changes in the selection of individuals who are unemployed along this dimension:

$$E_\delta\left(T_i^\delta\right) - E_{\delta'}\left(T_i^{\delta'}\right) = \underbrace{E_\delta\left[T_i^\delta - T_i^{\delta'}\right]}_{\text{within-individual dynamics}} + \underbrace{E_\delta\left(T_i^{\delta'}\right) - E_{\delta'}\left(T_i^{\delta'}\right)}_{\text{dynamic selection}}.$$

The two components are related to the heterogeneity in job-finding rates. On the one hand, it is clear that some heterogeneity in job finding is needed for selection effects to play any role. If job-finding probabilities do not vary, it does not matter which individuals select into long-term unemployment or which individuals become unemployed over the business cycle. The heterogeneity in job finding implies an upper bound on how important selection effects can be. On the other hand, within-individual changes in job finding may imply that the heterogeneity in job finding is not persistent. Indeed, when individuals are subject to differential changes in their job finding, those who are most at risk of remaining unemployed may change over the unemployment spell or over the business cycle. The persistence in heterogeneity, however, determines the potential importance of dynamic selection.

We illustrate the role of persistent heterogeneity for dynamic selection over the unemployment spell. Job seekers who select into long-term unemployment tend to have lower job-finding chances at the start of the spell, but whether or not they also have low job finding chances later in the spell depends on the persistence in the job-finding rates over the unemployment spell. In particular, Mueller, Spinnewijn and Topa [2021] showed that:

$$E_d\left(T_i^d\right) - E_{d+1}\left(T_i^{d+1}\right) = E_d\left(T_i^d - T_i^{d+1}\right) + \frac{cov_d\left(T_i^d, T_i^{d+1}\right)}{1 - E_d\left(T_i^d\right)}, \tag{2}$$

where d refers to the duration of the ongoing unemployment spell. The dynamic selection term $E_d\left(T_i^{d+1}\right) - E_{d+1}\left(T_i^{d+1}\right)$ is fully determined by the covariance between the job-finding chances in the current period of the spell vs. the next period of the spell. This captures the heterogeneity in job finding that remains persistent over the unemployment spell. Alternatively, we can decompose the duration-dependence in job-finding rates as,

$$E_d(T_i^d) - E_{d+1}(T_i^{d+1}) = E_{d+1}(T_i^d - T_i^{d+1}) + \frac{var_d(T_i^d)}{1 - E_d(T_i^d)}. \tag{3}$$

The dynamic selection term is now evaluated using the job finding at d, $E_d\left(T_i^d\right) - E_{d+1}\left(T_i^d\right)$, and fully determined by the variance. This alternative decomposition thus points to the importance of the overall heterogeneity for the dynamic selection, including any transitory heterogeneity. However, in this alternative decomposition, the within-individual dynamics are evaluated for the selected sample of individuals who are still unemployed at $d+1$. Any transitory heterogeneity will lead to mean reversion, which for this selected sample reduces the decline in job-finding chances over the spell or may even reverse its sign.[4]

Identifying Persistent Heterogeneity. Just like we can infer the overall heterogeneity from the covariance between the contemporaneous predictions and job finding realizations, we can infer the *persistent* heterogeneity from the covariance between predictions and realizations, but using lags or leads of the predictions instead:

Proposition 2. *If the predictor is unbiased, i.e.* $E_\delta(\hat{F}_i^\delta|X_i) = T^\delta(X_i)$*, then the hold-out sample covariance of the observed realization at time* δ *and the prediction model at time* δ' *evaluated in the sample of all unemployed at time* δ *and is an estimate of the covariance in observable types between time* δ *and time* δ'*:*

$$cov_\delta(F_i^\delta, \hat{F}_i^{\delta'}) = cov_\delta(T^\delta(X_i), T^{\delta'}(X_i)). \tag{4}$$

Proof. See Appendix B.1.

Note that this only identifies the persistence in individual types holding an individual's observables constant. There could be less persistence if the observables also change over time, the unemployment spell or the business cycle.[5] Given the decomposition of the observed duration dependence shown further above, we can use the covariance in equation (4) to identify an upper bound for the individual-level decline in job finding over the unemployment spell. We can proof the following corollary to Proposition 2:

Corollary 1. *If the predictor is unbiased, i.e.* $E_d(\hat{F}_i^d|X_i) = T^d(X_i)$*, and the unobserved heterogeneity is weakly persistent, i.e.* $cov_d(\varepsilon_i^d, \varepsilon_i^{d+1}) \geq 0$*, then the hold-out sample moments of the observed duration dependence in job finding and the covariance of the observed realization at duration d and the prediction model at duration* $d+1$ *provide an upper bound for the individual-level decline in job finding over the unemployment spell, as follows:*

$$E_d\left(T_i^d - T_i^{d+1}\right) \leq E_d\left(F_i^d\right) - E_{d+1}\left(F_i^{d+1}\right) - \frac{cov_d\left(F_i^d, \hat{F}_i^{d+1}\right)}{1 - E_d\left(F_i^d\right)}. \tag{5}$$

Proof. See Appendix B.1.

We will use equation (5) in the empirical analysis further below. Note that if we predict all of the individual-level heterogeneity or if unobserved heterogeneity is purely transitory, then equation (5) holds with equality.

[4] Indeed, $E_{d+1}(T_i^d - T_i^{d+1}) = E_d\left(T_i^d - T_i^{d+1}\right) + \frac{cov_d\left(T_i^d, T_i^{d+1} - T_i^d\right)}{1 - E_d(T_i^d)}$. Transitory heterogeneity makes the latter covariance term negative.

[5] In our empirical application of the dynamics of job finding over the unemployment spell, we only use predictors from the year prior to the start of the unemployment spell and thus the individual characteristics do not change over the unemployment spell.

Stylized Model. We briefly illustrate the dynamics of job finding in a stylized model where job finding types are determined by a permanent, dynamic and transitory component:

$$T_i^\delta = T_i + h(\delta) + \nu_i^\delta, \text{ where } \nu_i^\delta \text{ is iid.} \tag{6}$$

In this stylized model, the persistent heterogeneity is fully captured by the variance in the permanent component, $cov_\delta\left(T_i^\delta, T_i^{\delta'}\right) = var_\delta(T_i)$. The variance in the transitory component, however, contributes to the overall heterogeneity, $var_\delta\left(T_i^\delta\right) = var_\delta(T_i) + var_\delta(\nu_i^\delta)$, and determines the mean reversion underlying the dynamics of those remaining unemployed, $E_{\delta'}\left(T_i^\delta - T_i^{\delta'}\right) = h(\delta) - h(\delta') + \frac{var_\delta(\nu_i^\delta)}{1-E_\delta(T_i^\delta)}$. The transitory heterogeneity determines more generally how much variation there is in the within-individual dynamics, $var_\delta(T_i^\delta - T_i^{\delta'}) = 2var_\delta\left(\nu_i^\delta\right)$. Importantly, this type of heterogeneity is assumed away in the often-used proportional hazard model (e.g., Cox [1972], Lancaster [1979]), which relies on the persistence in observable heterogeneity to separately identify the within-individual dynamic component. Transitory heterogeneity can also be relevant over the business cycle and imply that those most at risk of long-term unemployment differ in different phases of the business cycle.

Ultimately, it is an empirical question how important transitory vs. persistent heterogeneity is, either over the unemployment spell or over the business cycle. In line with Propositions 1 and 2, the covariances of the observed realizations with the contemporaneous vs. lagged predictions allow to separate the variance in permanent vs. transitory observable types in our stylized model:

$$\begin{aligned} cov_\delta(\hat{F}_i^{\delta'}, F_i^\delta) &= var_\delta\left(T\left(X_i\right)\right) \text{ and} \\ cov_\delta(\hat{F}_i^{\delta}, F_i^\delta) - cov_\delta(\hat{F}_i^{\delta'}, F_i^\delta) &= var_\delta\left(\nu^\delta\left(X_i\right)\right), \end{aligned}$$

where $T(X_i)$ and $\nu^\delta(X_i)$ are the predictable components of T_i and ν_i^δ.

The stylized model thus indicates that to gauge the persistence in heterogeneity in the empirical analysis, we can leverage the difference in the predictability of job finding using contemporaneous vs. lags or leads of predictions. A first caveat here is that we cannot predict all of the individual-level heterogeneity, but only capture the observable heterogeneity. Hence, our covariances reflect a lower bound for the true variance in types, i.e. $var(T_i)$ and $var_\delta(\nu_i^\delta)$. We can, however, capture both observed and unobserved heterogeneity when observing multiple observations for the same individuals, i.e., F_i^δ and $F_i^{\delta'}$. This is the idea underlying the use of multiple unemployment spells as proposed by Honoré [1993] and more recently Alvarez et al. [2022]. The disadvantages of the multiple-spell approach, however, are that it only identifies heterogeneity that is permanent across dynamic states δ and δ' and relies on a potentially selected sample of individuals that are observed to be unemployed in those two states. We study this further in Section 3.5, where we use a sample of workers with multiple spells to identify heterogeneity in types and compare results with our approach. A second caveat is that in practice heterogeneity in job finding is not the simple sum of a dynamic component and permanent and transitory heterogeneity, as shown in equation (6). We can, however, study the heterogeneity in within-individual dynamics more directly by analyzing the relation between individuals' predictions across dynamic states, $\hat{F}_i^\delta$ and $\hat{F}_i^{\delta'}$. We turn to this in Sections 4 and 5.

3 The Predictability of Long-Term Unemployment

This section presents the prediction model, evaluates its predictive value and shows how much different sets of variables contribute. We start by describing the data and institutional context.

3.1 Data and Context

We merge several data sources from Sweden for the universe of prime-age job seekers starting an unemployment spell between 1992 and 2016.

First of all, we use data on unemployment spells from the Public Employment Service (PES), merged with data on UI benefit payments from the UI funds in Sweden. Unemployment benefits replace 80% of pre-unemployment earnings for workers who have worked for at least 6 months prior to being displaced and contributed to the UI system for at least 12 months. The unemployment benefit level is subject to a maximum and a minimum. Before 2001, the benefits were constant during the unemployment spell. Downward steps have been introduced in subsequent reforms for both the replacement rate and the maximum level. UI benefits are typically received for 450 business days, after which the unemployed must accept to participate in counselling activities and, potentially, active labor market programs (ALMP).[6] The PES organizes various ALMPs for unemployed workers with training programs being the cornerstone of Swedish labor market policy for many years. The ALMPs are targeted to the long-term unemployed or those who are 'typically at risk' of long-term unemployment [Richardson and van den Berg, 2013]. Our data contains information on the date the unemployed registered with the PES (which is a pre-requisite to start receiving UI benefits), unemployment benefits received and participation in the ALMPs. We define an unemployment spell as a spell of non-employment during which an individual receives unemployment benefits. To define the start date of an unemployment spell, we use the registration date at the PES. The end of a spell is defined as finding any employment or leaving the PES for other reasons.

Second, the unemployment data are linked with the longitudinal dataset LISA which merges several administrative and tax registers for the universe of Swedish individuals aged 16 and above. In addition to socio-demographic information (such as age, family situation, education, county of residence, immigration, etc.), LISA contains exhaustive information on earnings, taxes and transfer and capital income on an annual basis. It also includes data on the occupation and industry of workers. We use the matched employer-employee register (RAMS) to obtain further information on workers' employers and their tenure prior to becoming unemployed.

Third, for selected years and sub-samples, we have access to additional data sources. This includes granular data on asset portfolios, including liquid bank accounts, outstanding debt and other financial and real asset holdings for the universe of households, but only up to 2007 when the wealth tax in Sweden was abolished. We use data on union membership and contributions to the UI system [Landais and Spinnewijn, 2021]. Sweden is with Iceland, Denmark and Finland, one of the only four countries in the world to have a voluntary UI scheme, in which an important role is played by the unions who administer the unemployment benefits. Workers who have not contributed to obtain the comprehensive UI coverage receive the minimum benefit level instead. The premium for comprehensive

[6] See Kolsrud et al. [2018] for more details on the UI schedule in Sweden.

Table 1: VARIABLES INCLUDED IN BASELINE PREDICTION MODEL

Socio - demographics	Migration history	Income (last 5 years)	Employment (last 5 years)	Industry and municipality
Educational attainment	Years since migration	Labour Income	Number of unemployment spells	3-digit industry code
Age	Citizenship	Income from other sources	Days on UI	Municipality
Gender		Income of other HH members	Days on DI	
Marital status			Employment status	
Number of kids			Number of employers	
Age of kids			Pre-unemployment job tenure	
Foreign citizenship			Pre-unemployment firm (size, growth rate, layoff rate)	

Notes: For income histories, we include variables capturing annual income during the year before the unemployment spell and total income during the preceding 4 years. For employment histories, we compute the variables over the last 2 years and last 5 years.

UI coverage was heavily subsidized, but this subsidy has been reduced in a reform in 2007. Around 80-90 percent of workers have been covered by comprehensive UI. Finally, we have IQ data for men from military enlistments and we have also merged survey data from SILC (Survey of Income and Living Conditions) to our administrative data with questions related to health and mental well-being.

Table 1 lists the range of variables used in the baseline model, which are generally available for all years in the sample period and thus allows estimating the baseline model for every year in a consistent manner.[7] We only use pre-determined variables, i.e. variables that predate the unemployment spell in question. The set of variables thus includes basic socio-demographics, which are usually available in labor force survey data; migration history, including citizenship and years since immigration; yearly income, both wage and non-wage, as well as income from other household members, for each year over the 5 years pre-dating the year of the unemployment spell; employment history over the five years, pre-dating the year of the unemployment spell, including the employment status in November of each year, the number of unemployment spells, days on UI, days on DI, and the number of employers. In our baseline model, we include these variables for both over the last 2 and the last 5 years to capture the timing of the employment histories. We also include job tenure at the pre-unemployment firm and its characteristics (size and employment growth and layoff rate); finally, our baseline model includes 3-digit industry and municipality dummies. Additional variables - which are not available in all years or only for a subset of individuals - are used in extensions of the baseline model for the year 2006.

Table 2 shows descriptive statistics for the sample of unemployed job seekers in our sample. Panel A compares the sample of unemployed to the overall population and shows that the sample of unemployed is selected towards the young, foreign and low-education and low-income population. Overall, our data over the years 1992 to 2016 features over 7 million unemployment spells. An important observation

[7] A limitation to this is that, since the LISA panel only starts in 1990 and even later for days spent on UI and DI, income and employment histories are partly censored for the earliest years in our sample. We impute these censored observations using the individual's history when partially observed, falling back on the population mean in 1995 when the full history is missing.

Table 2: DESCRIPTIVE STATISTICS

A. Unemployed Sample vs. Population	**Mean**	
	Unemployed	Population
Age	36.0	40.0
Female	48%	49%
Foreign	16%	8%
Primary Edu.	21%	16%
Secondary Edu.	57%	54%
Tertiary Edu.	21%	29%
Labour Income (2006 SEK)	100,500	199,700
Other Income (2006 SEK)	12,100	14,900
Household Income (2006 SEK)	106,200	156,200
Number of unemployment spells	7,259,797	
Spells interrupted by training	1,628,080	

B. Employment History	**Mean**	**Percentile**		
		25th	50th	75th
Days on UI (2y)	113	0	48	190
Days on UI (5y)	251	0	156	409
Days on DI (2y)	34	0	0	0
Days on DI (5y)	66	0	0	19
Unemp. Spells (5y)	1.4	0	1	2
Employers (5y)	1.9	1	2	3
Tenure (years)	1.8	1	1	3
Firm Size	4,460	15	147	3,063
Firm Layoff Rate	35%	17%	25%	48%

Notes: Descriptive statistics for the baseline sample for the years 1992-2016 and ages 25-54.

is that many of these spells include the participation in ALMPs including training and job search assistance, though a majority of them start at 12 months of unemployment or later. We include these spells in our baseline prediction exercise, but in a series of robustness checks we perform the prediction exercise without spells that enter ALMPs in the first 6 months of the unemployment spell. In Panel B, we show additional statistics for employment histories of the unemployed in our sample, including days on UI and DI, number of employers and employment spells, the pre-unemployment tenure, prior employer size and its layoff rate in the prior year. Overall, the table shows that there are substantial differences in these variables across the unemployed in our sample.

3.2 Prediction Model

Our baseline model predicts the long-term unemployment risk for spells starting in 2006. To be precise, we predict the probability that someone does not leave unemployment in the first 6 months of the unemployment spell. We will mostly report the complement of this probability, which we will refer to as the 6-month *job-finding probability* at the start of the unemployment spell.[8] We then redo the prediction exercise for every year in the data and also estimate the 6-month job-finding probabilities for workers at 6 and 12 months of the unemployment spell in 2006. As in any prediction exercise, there is a trade off between improving predictive power and over-fitting when including more variables. To address this issue, we use machine learning techniques to optimally choose variables to include in model and generate predictions. More specifically, we use the Ensemble Model (see, e.g., Einav et al. [2018]), which combines three different Machine Learning (ML) algorithms: LASSO, Gradient Boosted Decision Trees, and Random Forest. The Ensemble Model trains each ML algorithm separately and, in a final step, assigns each of these three algorithms a linear weight. All results presented here use the predictions of this Ensemble Model, but evaluated in the hold-out sample. We provide more detail in Appendix C.

To evaluate the accuracy of our prediction model, we compare predictions and outcomes in the hold-out sample for the year 2006 in Panel A of Figure 1. The figure shows a binned scatter plot of the averages of the 6-month job-finding rate for 20 equally sized bins of the predicted 6-month job-finding probability. Typically, attenuation of outcomes from the 45 degree lines suggest issues with sampling error due to limited sample size, which is not a significant issue here as the prediction exercise yields outcomes that are close to the 45 degree line. The individual-level linear regression, shown as a red line, of a dummy for finding a job within 6 months on the predicted 6-month job-finding probability has a slope of .93 (.01).

Our prediction algorithm thus seems to work very well, producing predictions that are subject to minimal bias. In Appendix C, we gauge this further by performing the prediction exercise separately for different groups. We find that the slope remains close to one if we split the sample by income, citizenship, gender, education, days on UI and days on DI. We also split the sample into 40 groups based on income decile, gender and days on UI and compute the average job-finding rate for each group in the hold-out sample. We find again that the slope is very close to one. This remains true even if we split the sample further into 144 groups based on income decile, gender, citizenship, days on DI and days on UI. This shows that our prediction exercise also does well for different sub-groups.

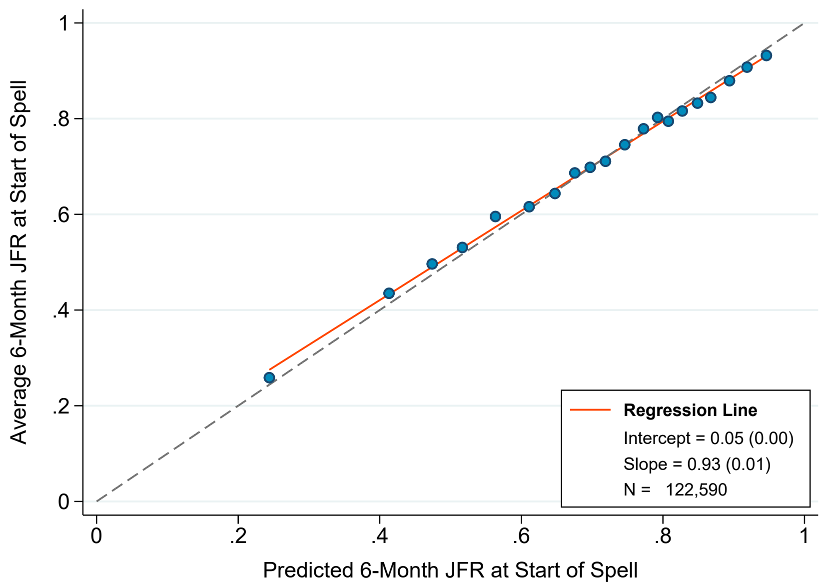
Finally, we estimate a linear model with the same variables as in the baseline ML model and again compare its accuracy on the hold-out sample. As the least square estimates of the linear model are unbiased, this is assessing the importance of sampling error in attenuating the relation between predictions and outcomes. As Appendix Figure A2 shows, we again find that the slope is very close to 1 and similar to the baseline ML model. This suggests that our approach does not suffer from any biases that may arise in the non-linear ML algorithm.

Overall, we conclude that our prediction exercise performs very well and we do not detect any biases in the prediction, neither for job seekers with specific predicted job-finding probabilities, nor for job seekers belonging to specific observable groups.

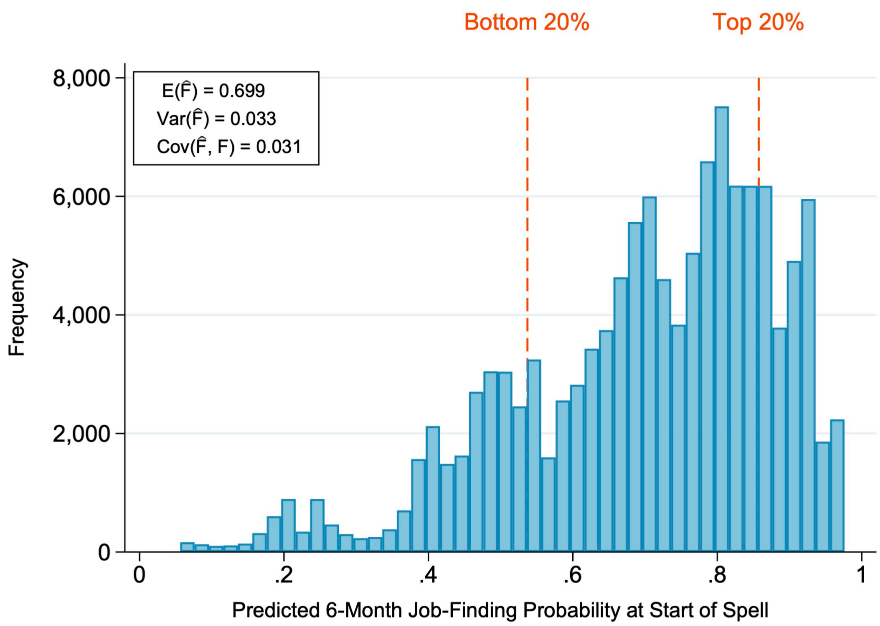
[8] In Section 3.3 below, we probe the robustness of our results to more narrow definitions of job finding and find that the predictive power of our model remains very similar.

Figure 1: Baseline Prediction Model in 2006

A. Comparing Predictions to Outcomes



B. Distribution of Predicted Job-Finding Probability



Notes: Panel A presents a binned scatter plot of observed and predicted job finding. That is to say, we split the hold-out sample into 20 vigintiles of predicted 6-month job-finding probability and report, for each bin, mean observed and predicted 6-month job-finding rates at the start of the spell. The red line shows the results of a linear regression at the individual level of a dummy for finding a job within 6 months on the predicted 6-month job-finding probability. Panel B shows the distribution of the predicted job-finding probability in the hold-out sample for the year 2006.

3.3 Predictive Value

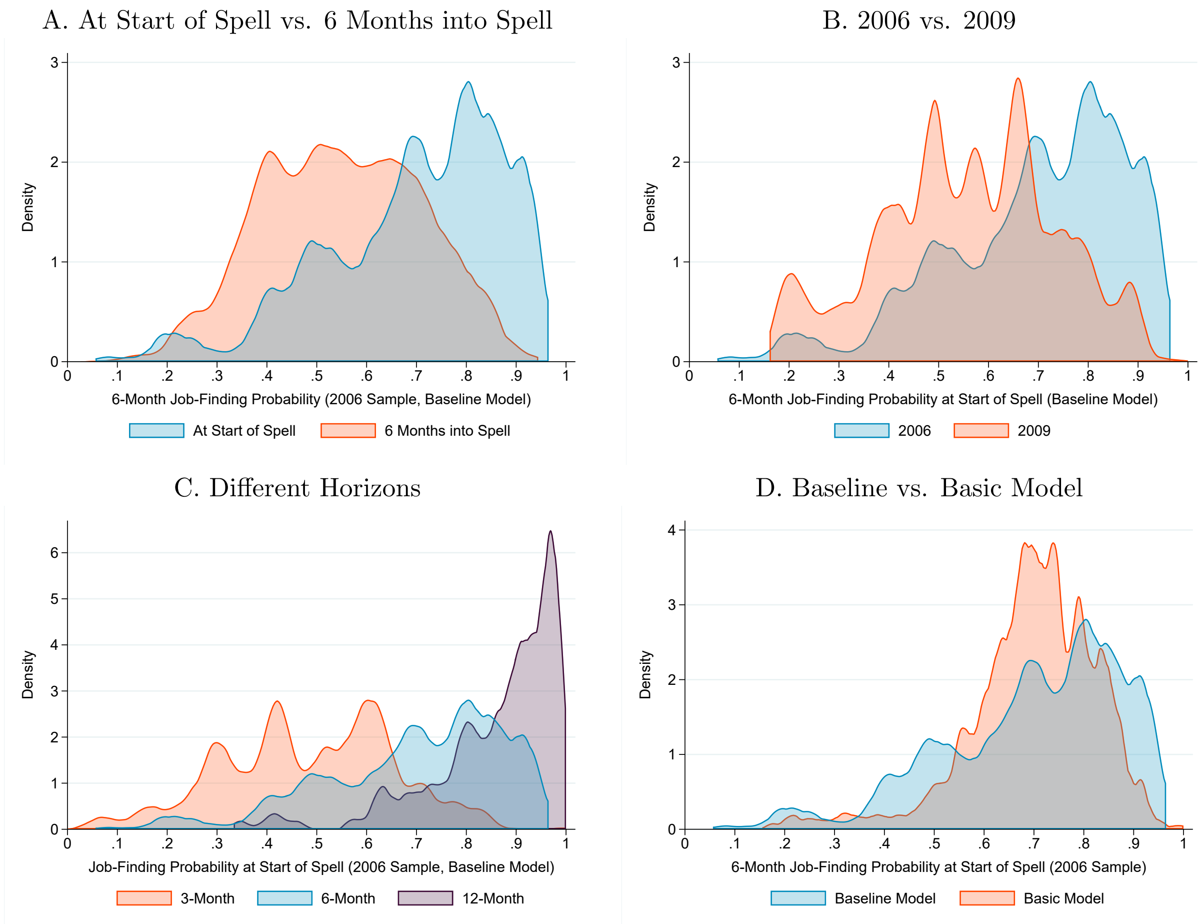
We use our prediction model to assess the heterogeneity in long-term unemployment risk. Panel B of Figure 1 shows the distribution of the predicted 6-month job-finding probability in the hold-out sample for the year 2006. The figure shows substantial dispersion in the job-finding probability, with our predictions covering almost the entire range from 0 to 100 percent. The average job-finding probability equals 70 percent. For the 80th percentile of the sample, the predicted job-finding probability is as high as 86 percent. The distribution has a long left tail. The 20th percentile corresponds to a predicted job-finding probability of 53 percent, the 5th percentile to a probability of 38 percent. The predicted job-finding probability is subject to error in the prediction model and thus Figure 1 can in principle exaggerate the dispersion in the 6-month job-finding risk. As shown above, the covariance of the predicted job-finding rate with actual job finding in the hold out sample gives an accurate estimate of the variance in job-finding risk. Our estimate of the covariance is 0.031, which is substantial and only slightly below the variance of 0.033. This suggests that the prediction error is rather small, confirming the results from Panel A of Figure 1. We also find that the R-squared of the predicted job-finding rate with actual job finding is 0.136. To address the issue of sampling error, we also bootstrapped the 95% confidence interval for these moments and find its fairly tight with [0.030, 0.031] for the covariance and [0.132, 0.140] for the R-squared.[9] The R-squared may appear fairly low, but the dependent variable is a binary, random realization of a probability. The low R-squared thus reflects the noise associated with this binary realization.[10] A complementary way to evaluate the predictive power of the model is to consider the area under the ROC. This curve contrasts the false-positive rate to the true-positive rate in the hold-out sample, depending on the threshold used to map the predicted job-finding probabilities into a binary outcome. The area under the ROC equals 0.72, compared to a lower bound of 0.5 for a random prediction model and an upper bound of 1 for a model with perfect foresight.

Our baseline model studies job-finding probabilities at the start of unemployment spells that initiated in 2006, which is prior to the Great Recession. Panels A and B of Figure 2 illustrate how the distribution of predicted job-finding rates changes when considering instead job seekers who are long-term unemployed or became unemployed during the Great Recession. In particular, Panel A compares the distributions for individuals six months into the spell vs. at the start of the spell, while Panel B compares the distributions for individuals at the start of the spell in 2009 vs. 2006. Not surprisingly, both for the long-term unemployed and the recession year, the job-finding rates are substantially lower on average. In both cases, few individuals remain who are almost certain to find a job in the next six months, while the predicted job-finding probabilities become more compressed in the bottom range of the distribution. As discussed in the conceptual framework, the observed differences in job finding can result form compositional changes in the pool of unemployed job seekers, but also from within-individual dynamics in job-finding chances, which can be heterogeneous too. We aim to separate the different forces in the next two sections.

[9] We draw 500 bootstrap samples from the hold-out sample in 2006, following Mullainathan and Spiess [2017].

[10] For example, if T_i is uniformly distributed on interval [0, 1], then $R^2(F_i, T_i) = 1/3$.

Figure 2: Distribution of Predicted Job-Finding Probability



Notes: This figure reports the distribution of various predicted job-finding probabilities. In all four panels, the baseline (in blue) is the predicted 6-month job-finding probability at the start of the spell for the 2006 holdout sample. Panel A compares the baseline with the predicted distribution 6 months into the spell. Panel B compares with the predicted distribution in 2009. Panel C shows the predicted job-finding probabilities over 3-month and 12-months horizons. Finally, Panel D contrasts the baseline model, which uses all the variables shown in Table 1, with the basic model, which only uses the socio-demographic variables shown in the table.

Robustness. We perform a series of robustness checks where we evaluate the predictive value over different horizons, for alternative models and for different samples.

First, while our focus is on job finding over a 6 month-horizon, as it corresponds to the standard measurement of long-term unemployment risk, we can evaluate the job-finding rates over different horizons too. Panel C of Figure 2 shows substantial dispersion in predicted job finding when considering instead job finding over the first 3 or 12 months of the unemployment spell. The prediction exercise yields a similar predictive power when the job-finding rate is measured at 3, 6 and 12 months of the unemployment spell (see Appendix Table A1). Moreover, when we compute the R-squared of the prediction for the 3- and 12-month job-finding rate with actual job finding over the first 6 months, we find that the R-squared are very similar compared to using the 6-month model. This is re-assuring and suggests that our results do not rely on a particular horizon chosen for the job-finding rate.

Second, we evaluate the predictive power of the linear model using all the variables of the baseline model. Interestingly, we find that the R-squared is slightly higher than for our machine learning model, as reported in Appendix Table A2. This suggests that over-fitting is not a first-order issue when the sample of unemployment spells is large. The potential non-linearities leveraged by ML methods are not particularly important either for the high predictive power of our baseline model. Instead, what matters is the rich data going into the model.

Third, we gauge how our prediction is affected by our definition of job finding and the transition through active labor market policies. We first consider active labor market policies (ALMP) started in the first 6 months of the unemployment spell. Figure A12 in the Appendix shows that these spells are not strongly correlated with predicted long-term unemployment risk. We return to this issue in Section 6. Table A2 in the Appendix also shows that the R-squared and covariance change little when we exclude these spells from the hold-out sample. We also re-estimate the model after excluding all spells with ALMP starting in the first 6 months and we find that the R-squared remains similar both in the hold-out sample of spells with the ALMP spells included ($R^2 = 0.134$) and excluded ($R^2 = 0.133$), see also the Appendix Figure A12.

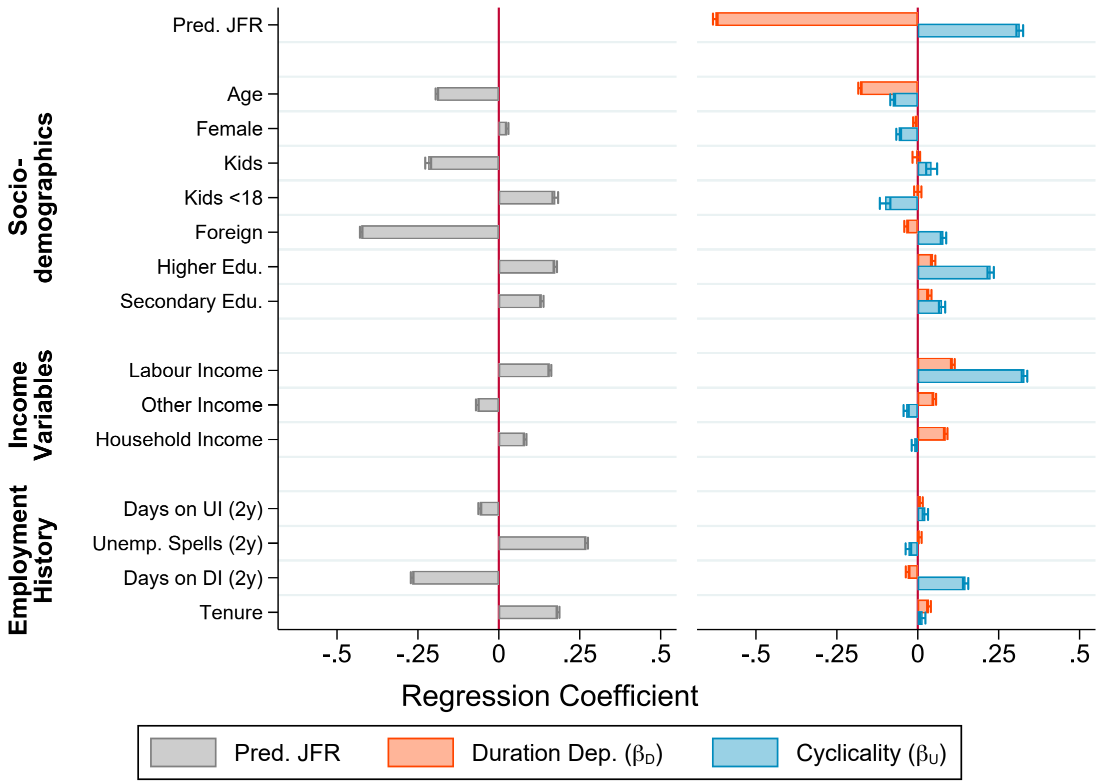
Finally, we evaluate the predictive power of our baseline model when we use additional information on the end of the unemployment spell. As shown in Appendix Table A2, when we re-define job finding in the hold-out sample as those spells which ended but did not take up education other than training, the predictive power of the baseline model remains nearly unaffected. When we re-define job finding as those spells which ended but did not take up education and did not end for unknown reasons, the predictive power is somewhat lower, but still high with an R-squared of 0.117.

3.4 Predictive Variables

While our prediction model does not allow us to evaluate the causal impact of job seekers' characteristics on their unemployment risk, we are interested in how much different sets of variables contribute to the predictive value of the model. Prior work using labor force surveys to predict long-term unemployment has been mostly limited to socio-demographic variables, including age, gender, family composition and education. A simple linear regression already reveals that various other characteristics relate significantly to long-term unemployment risk, conditional on these socio-demographic variables. To make the estimates comparable and help with the interpretation, we first standardize both the explanatory variables and the predicted job-finding probabilities. The left panel of Figure 3 shows that on average job seekers face a higher risk of long-term unemployment when they are older and less educated. This risk is also further increased for job seekers who had lower income and tenure in their prior employment and spent more days on UI and DI prior to the unemployment spell.

We can evaluate more formally how the predictive value of our baseline model compares to a basic model, which only uses socio-demographic variables (see Column 1 of Table 1). Panel D of Figure 2 shows that the predicted dispersion using more limited information is substantially smaller. To quantify how much smaller, Table 3 reports the R-squared in the hold out sample for various sub-models. The basic model with only socio-demographic variables has a predictive power that is less than half of our full model (R-squared of 0.057). We then sequentially add variables that are increasingly unlikely to be available in surveys and have been rarely used in earlier work. As shown in

Figure 3: Heterogeneity in Job Finding, Duration Dependence and Cyclicality



Notes: This figure presents results from linear regressions of the predictions on a subset of observables. Both right-hand-side and left-hand-side variables have been standardized by subtracting the sample mean and dividing by the sample standard deviation, so the coefficients can be interpreted as the standard-deviation change in the outcome associated with a one-standard-deviation change in the covariate. The left panel shows the OLS coefficients of a regression of the predicted 6-month job-finding probability at the start of the spell, from the baseline model in 2006, on the variables listed on the y-axis. The right panel shows the coefficients from regressions of the duration dependence parameter β_D (see Section 4) and the cyclicality parameter β_U (see Section 5) on the same covariates and the predicted job-finding rate. In both panels, standard 95% confidence intervals are shown around the point estimates.

the top panel of the table, the subsequent inclusion of income variables and the employment history substantially improves the predictive power of the model and these variables thus seem key for the predictability of long-term unemployment risk. Moreover, once we have added these variables, further adding income history, migration history, and location and industry effects do not add much predictive power.[11] Clearly, the ordering of variables used for various sub-models in Table 3 matters as different predictors are correlated. To compare the marginal contributions of different sets of variables, we add each of them separately to the socio-demographics in the basic model. The bottom panel of the table reveals that adding either the income variables, the migration history, or industry fixed effects increase the R-squared by about 50 percent and thus realizes half of the gain in predictive power when

[11] Note that the R-squared declines slightly for some specifications as we add more variables, which is likely due to sampling error in the training and hold-out samples.

expanding from the basic to the baseline model.[12]

The most predictive set of variables regards the individuals' employment history prior to the unemployment spell. This includes number of days spent unemployed, the number of unemployment spells, but also DI receipt and number of job switches (see column 4 of Table 1). Decomposing this history, we find that information in the year prior to unemployment is sufficient to realize most of the gain in predictive power, with an increase in R-squared of 87% relative to an increase of 5% or less for any additional year (see Appendix Table A4). The marginal contribution is highest when using the most recent year, but still sizeable at 23% when only adding information from five years earlier.[13]

We explore to what extent the predictive power of the baseline model can be improved further by adding additional variables available in 2006. As these variables are not available for all years between 1992 and 2016 and some are only available for a subsample, we did not include them in the baseline model. Panel C of Table 3 shows that using information on the prior occupation (at the 3-digit level) significantly increases the predictive power of the baseline model, in line with employment variables and history prior to the unemployment spell being most predictive. The R-squared increases by 5.2%, even though we observe prior occupation for only 59.5% percent of the spells. Adding information on UI benefits increases the R-squared by 4.7%, which we observe for 54.6% percent of the spells. This is again a significant increase, but arguably small in light of the attention given to the design of UI benefits and the corresponding moral hazard. Of course, the UI benefit level received depends on the pre-unemployment earnings and employment history, which are included in the baseline model. However, even relative to the basic model including only basic socio-demographics, the UI benefits and thus job seekers' potential search responses explain relatively little of the variation in employment outcomes (see Appendix Table A6). Adding workers' choices to get comprehensive UI or not increases the R-squared by only 2.0%. However, we note that most workers do get the comprehensive UI (70.3%) and prior work has shown relatively limited risk-based selection into comprehensive UI with most of it explained by observables included in our prediction model [Landais and Spinnewijn, 2021]. Another striking observation is that adding detailed information on individuals' financial and real assets adds very little predictive power relative to the baseline model (increase in R-squared of 1.0%). Adding information on IQ or on union membership does not even increase the predictive power of our model. The final column in Panel C of Table 3 shows that the additional information from the various administrative registers jointly increases the R-squared for our prediction to 0.144 (relative to 0.135 for our baseline model). Other than information on DI receipt, our prediction model uses no information on job seekers' health. In Appendix Table A5 we briefly explore the potential predictive power of health status using survey data. The results indicate that information on mental health in particular adds additional explanatory power above and beyond our model predictions, although we cannot perfectly address concerns of over-fitting in the small survey samples.[14]

[12]As non-linearities do not seem very relevant, we can evaluate the contribution of different sets of variables by sequentially adding them to a linear regression of the predictions and observing how the R^2 changes. For this, we calculate a Shapley-Owen decomposition of the R^2, as described by Grömping [2007] and Huettner and Sunder [2012]. Intuitively, when the regressors are correlated, the change in the R^2 that we attribute to each variable depends on the order in which it is introduced into the regression. The Shapley-Owen decomposition overcomes this limitation by computing the average change across all possible orderings of the variables. The conclusions are very similar, as shown in Appendix Table A10.

[13]Panel C in Appendix Table A4 also shows that none of the specific features of the pre-unemployment history (e.g., UI receipt, number of unemployment spells, employment switches) by themselves are driving the predictive power.

[14]After constructing a mental health index using principal component analysis, we find that adding this index to our

Table 3: R^2 FOR VARIOUS MODELS IN THE YEAR 2006

A. Sub-models of Baseline - Sequential

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$R^2(\hat{F}_{i,0}, F_{i,0})$	0.057	0.086	0.094	0.128	0.132	0.136	0.139	0.136
Change (j) vs $(j-1)$	-	+51.0%	+8.9%	+37.2%	+2.7%	+3.2%	+2.3%	-2.4%
Socio-demographics	X	X	X	X	X	X	X	X
Labour Income		X	X	X	X	X	X	X
Other Income			X	X	X	X	X	X
Employment History				X	X	X	X	X
Income History					X	X	X	X
Migration History						X	X	X
Industry							X	X
Municipality								X

B. Sub-models of Baseline - Marginal

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$R^2(\hat{F}_{i,0}, F_{i,0})$	0.057	0.086	0.072	0.129	0.092	0.078	0.085	0.064
Change (j) vs (1)	-	+51.0%	+25.9%	+126.4%	+61.9%	+36.9%	+49.5%	+12.1%
Socio-demographics	X	X	X	X	X	X	X	X
Labour Income		X						
Other Income			X					
Employment History				X				
Income History					X			
Migration History						X		
Industry							X	
Municipality								X

C. Extensions of Baseline

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$R^2(\hat{F}_{i,0}, F_{i,0})$	0.135	0.143	0.135	0.137	0.134	0.138	0.142	0.144
Change (j) vs (1)	-	+5.2%	-0.5%	+1.0%	-1.4%	+2.0%	+4.7%	+6.5%
Baseline Variables	X	X	X	X	X	X	X	X
Occupation		X						X
Union member			X					X
Wealth				X				X
IQ					X			X
UI choice						X		X
UI benefits							X	X

Notes: The table shows the R^2 of the predicted 6-month job-finding probability and a dummy for actual job finding in the hold-out sample for the year 2006 for various models. Panel A starts from the basic model in (1) and adds variable groups sequentially until all of the groups included in the baseline model are incorporated in (8). Panel B adds the same variable groups one at a time. Finally, Panel C starts from the baseline model in (1) and adds additional information from other administrative data sets, first one at a time and then all at once in column (8).

Overall, this confirms that our prediction model seems saturated along some key dimensions of heterogeneity. However, it also reminds us that the predicted heterogeneity remains a lower bound on the overall heterogeneity across job seekers.

3.5 Identifying Unobserved Heterogeneity in Multiple Spell Data

As mentioned before, our prediction model identifies a lower bound on the extent of ex-ante heterogeneity in job finding as it does not capture heterogeneity that is orthogonal to observable characteristics and histories. An alternative approach, which identifies both observed and unobserved heterogeneity, is to use data on multiple unemployment spells per person, see Honoré [1993] and more recently Alvarez et al. [2022].[15] Identification through multiple spell data, however, has two potential disadvantages relative to our approach using observables: First, the multiple-spell approach identifies only ex-ante heterogeneity that is *fixed* between spells whereas our approach allows for changes in job-finding risk between spells. Second, the approach is limited to a sample of multiple unemployment spells, which is likely to be more homogeneous, especially given our finding that unemployment histories themselves have predictive power for job finding.

In Sweden we observe the universe of unemployment spells for the period 1992-2016 and thus observe multiple unemployment spells for a sizeable share (69%) of the individuals in our data. This allows us to estimate heterogeneity following the multiple-spell approach, but also to quantify the advantages and disadvantages compared to our approach. To do so, we restrict our sample to all individuals in the hold-out sample with at least two spells, choosing two spells at random if the individual has more than two. We then compute the covariance in observed job finding between the two spells. As shown in the Appendix B.1, this identifies the ex-ante heterogeneity in types that is fixed between the two spells.[16] Column 1 of Table 4 shows that this is 0.031, which is sizeable. We contrast this with the covariance of the predicted risk with the actual outcomes for a random sample of spells, including individuals with one spell only. The results are shown in Column 4, which is 0.029.[17]

Our analysis in Sweden thus provides similar lower bounds on the heterogeneity when either using the identification approach relying on multiple spells and or the identification approach relying on observables.[18] However, the nature of heterogeneity identified by the two approaches can be very different. To assess this further, we report in Column 2 the covariance of the predicted risk of the first spell with the actual outcome of job finding of the second spell and find that the covariance is substantially lower at 0.018. This shows that observable characteristics identify only about 58% of the full extent of the fixed heterogeneity in job finding. In Column 3, we compute instead the covariance

model predictions in a linear regression model increases the R-squared by 23 percent.

[15] Both papers prove identification in this context. The latter paper also shows empirically that ex-ante heterogeneity is important in the context of data of multiple unemployment spells in Austria.

[16] Proposition 4 in the Appendix shows that covariance in actual job finding is equal to the covariance in the underlying probabilities, so it identifies the persistence in types across spells. In a corollary, in the context of our stylized model in Section 2, we show that the covariance of actual job finding across two spells is equal to the variance of fixed types, T_i.

[17] To be precise, for every pair of spells in the two-spell sample, we draw a random spell from the same year, but including individuals with one spell only. This keeps the distribution of number of spells by year the same in the two-spell and control sample. This explains why this estimates deviates slights from estimate of 0.031 reported earlier.

[18] Note that the R^2 is substantially lower in Column 1 compared to the other Columns. This is natural since only using information from one unemployment spell in the past should have less predictive power than our baseline prediction model.

Table 4: INDIVIDUALS WITH REPEATED SPELLS

	Two-spell sample			Control Sample
	$F^1_{i,0}, F^2_{i,0}$	$\hat{F}^1_{i,0}, F^2_{i,0}$	$\hat{F}^2_{i,0}, F^2_{i,0}$	$\hat{F}^2_{i,0}, F^2_{i,0}$
	(1)	(2)	(3)	(4)
Cov$(\cdot)$	0.031	0.018	0.025	0.029
$R^2(\cdot)$	0.019	0.053	0.102	0.122
N	791,524	791,524	791,524	791,524

Notes: This table reports key statistics for the sample of individuals with multiple unemployment spells between 1992 and 2016. For the first three columns, we construct the two-spell sample by restricting the analysis to individuals with at least two spells in our hold-out sample. We take two of the spells, choosing at random if there are three or more, and randomly label them as "first" and "second". For the last column, we construct a control sample by matching each second spell with a spell chosen at random from the hold-out sample of the same calendar year, without excluding unique spells. We report the covariance and the R-squared for several combinations of observed and predicted job finding in the first 6 months of unemployment. Column (1) corresponds to *observed* job finding in the first ($F^1_{i,0}$) and second ($F^2_{i,0}$) spell. Column (2) matches predicted job-finding in the first spell ($\hat{F}^1_{i,0}$) and observed job-finding in the second spell. Columns (3) and (4) relate predicted ($\hat{F}^2_{i,0}$) and observed job finding during the second spell for the respective samples described above.

between the predicted risk of the second spell with the actual outcome of the second spell and find a value of 0.025. This suggests that there is a non-negligeable part of risk that is not fixed between spells. The overall predictable heterogeneity is about 39% higher than the predictable heterogeneity in Column 2 that is fixed between spells.[19] We come back to the issue of persistence in job-finding over time in the next Section 5 when we turn to job-finding dynamics over the business cycle. Finally, the difference in the covariance reported in Column 3 and 4 relates to sample composition and shows that the the restriction to a sample of two spells misses some additional predictable heterogeneity.

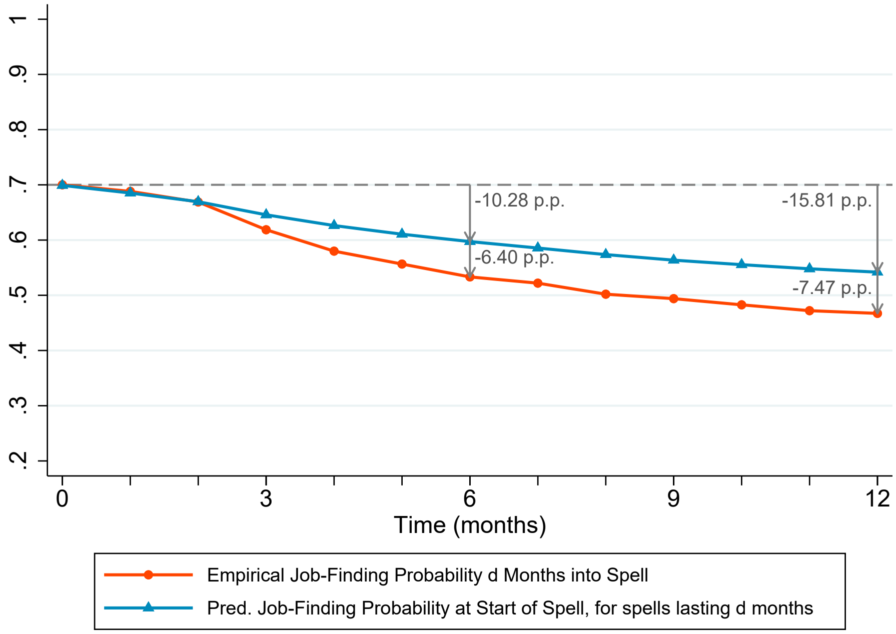
To sum up, while both approaches highlight a large role for heterogeneity in job-finding risk and of similar magnitude, they identify different elements of heterogeneity. The two approaches are thus highly complementary and jointly potentially imply a larger role for heterogeneity in job-finding risk than identified by our lower bound estimate.

4 Job Finding over the Unemployment Spell

The substantial heterogeneity in job-finding risk at the start of the spell implies that the long-term unemployed will differ from the short-term unemployed. This section studies how the job-finding chances change over the unemployment spell and revisits the question how much changes in the composition of unemployed job seekers contribute to this. We also study heterogeneity in the dynamics

[19]This result is confirmed in Appendix Table A7, where we restrict the two-spell sample further to spells two or five years apart. We find that the further apart the spells, the lower is the covariance in Columns 1 and 2, whereas the values in Columns 3 and 4 are not affected.

Figure 4: Dynamic Selection over the Unemployment Spell



Notes: The figure compares the evolution of the empirical 6-month job-finding rate d months into the spell with the predicted 6-month job-finding rate at the beginning of the spell for individuals who reach the d-th month of unemployment, in the hold-out sample for the year 2006.

individuals experience over the unemployment spell.

4.1 Dynamic Selection into Long-Term Unemployment

We first study how much compositional changes contribute to the average decrease in job-finding rates over the unemployment spell. Figure 4 provides a graphical illustration of the dynamic selection, using the prediction model estimated on job seekers unemployed at the start of the unemployment spell. The figure shows how the average of these predictions changes for the pool of job seekers still unemployed at different unemployment durations. The decline is substantial. For the job seekers still unemployed six months into the spell the average predicted job finding is 10.3 percentage points lower compared to all the unemployed starting a spell. This is 62 percent of the observed drop in job-finding rates of 16.7 percentage points. For those still unemployed after one year, the average predicted job finding is 15.8 percentage points lower, corresponding now to 68 percent of the observed drop. This analysis is constrained by the observables used in our prediction model. Indeed, using our basic model with socio-demographic variables only, we explain a drop of 8.1 percentage points instead and would thus assess the dynamic selection as potentially only half as important.

In order to interpret the residual drop in observed job finding as the dynamic effect of unemployment, i.e., as *true* duration-dependence, the job-finding rates would need to be fully persistent over the spell. If not, the implied mean reversion in job-finding rates among the surviving unemployed would imply that the residual drop underestimates the true-duration dependence (if there is no sig-

Table 5: PREDICTABLE HETEROGENEITY DURING THE UNEMPLOYMENT SPELL

Sample	**Model**	N	$E(F)$	$E(\hat{F})$	$Var(\hat{F})$	$Cov(\hat{F},F)$	$R^2(\hat{F},F)$
At Start of Spell	0M Model	122,590	0.700	0.699	0.033	0.031	0.136
	6M Model	122,590	0.700	0.604	0.027	0.022	0.090
	12M Model	122,590	0.700	0.536	0.025	0.011	0.024
6M into Spell	6M Model	41,647	0.547	0.550	0.026	0.018	0.054
	12M Model	41,647	0.547	0.508	0.025	0.010	0.016
12M into Spell	12M Model	20,305	0.485	0.494	0.025	0.007	0.008

Notes: The table reports summary statistics about models trained on different unemployment durations in 2006. The first three rows correspond to the hold-out sample at the start of the unemployment spell. We generate three different predictions for this sample, with models trained at the start of the spell (the contemporaneous predictions), 6 months into the spell and 12 months into the spell. Rows 4 and 5 deal with the hold-out sample 6 months into the spell; for this sample, we generate predictions using the models trained contemporaneously and 12 months into the spell. Row 6 presents results for the hold-out sample 12 months into the spell, using the contemporaneous model.

nificant selection on unobservable heterogeneity). This difference corresponds to the two alternative decompositions (2) and (3) presented in Section 2. The former decomposition allows to bound the dynamic effects, using the *persistent* heterogeneity in job finding over the unemployment spell instead. To gauge the persistence, we study the relative predictive value of the prediction model estimated on the long-term unemployed vs. on the unemployed at the start of the spell for the outcomes for the latter. In Table 5, we report $R_0^2(\hat{F}_{i,6}, F_{i,0}) = 0.090$ in the hold-out sample of 2006, compared to $R_0^2(\hat{F}_{i,0}, F_{i,0}) = 0.136$ when using the contemporaneous predictions. The lower sub-script 0 refers to the fact that the covariance is evaluated in the sample of all newly unemployed and not just those who remain unemployed for at least 6 months. Our estimates imply that two thirds of the (observable) heterogeneity is persistent over the first six months of the unemployment spell. Following Corollary 1, $cov_0(F_{i,0}, \hat{F}_{i,6})/(1 - E_0(F_{i,0}))$ provides a conservative lower-bound of 7.5 percentage points on the average decline in job finding due to selection, considering only persistent heterogeneity. This implies that the true duration-dependence can explain at most 51% of the observed decline in the 6-month job-finding rate between the start of the spell and at 6 months of the spell.

We perform the same calculation for the decline in the job-finding rate between 6 and 12 months of the spell and find a lower-bound decline due to selection of 0.022. This is about 36% of the total decline in the 6-month job-finding rate from 0.55 to 0.49. The declining role of selection over the unemployment spell can be attributed to the role of selection earlier in the unemployment spell. The reason is that, if heterogeneity is persistent, then the variance in predicted job-finding rates among survivors should decline over the unemployment spell. Indeed, we find that the hold-out sample variance in predicted 6-month job-finding rates is 0.031 at 0 months of unemployment, but declines to 0.018 at 6 months and 0.007 at 12 months of unemployment.

Robustness. We perform two robustness checks on these results. First, we address issues of sampling error due to the smaller sample sizes of unemployed at 6 and 12 months of duration.[20] To do this, we pool the years 2006 and 2007 and redo our prediction exercise. Appendix Table A8 reports the corresponding results, which are very similar to the results in Table 5. In fact, we find that $R_0^2(\hat{F}_{i,6}, F_{i,0}) = 0.100$ in the hold-out sample of 2006, compared to $R_0^2(\hat{F}_{i,0}, F_{i,0}) = 0.155$, and following Corollary 1, we find a very similar number for $cov_0(F_{i,0}, \hat{F}_{i,6})/(1-E_0(F_{i,0}))$, accounting for 45% of the observed decline in the 6-month job-finding rate between the start of the spell and at 6 months of the spell. This suggests that sampling error is not biasing our conclusions. Second, we redo the same analysis for 2009-2010 when both unemployment and LTU risk were higher and report the results in Table A9. We find that $R_0^2(\hat{F}_{i,6}, F_{i,0}) = 0.079$ in the hold-out sample of 2009-2010, compared to $R_0^2(\hat{F}_{i,0}, F_{i,0}) = 0.123$. While the persistent share is estimated to be equally important, the predictable heterogeneity is thus somewhat smaller in these recession years. The observed duration-dependence in job-finding rates is, however, is smaller as well during the recession years: the relative decline in the 6-month job-finding rate 6 months into the spell equals 19% in 2009-2010 compared to 24% in 2006-2007. This is consistent with prior evidence for the US in Krueger et al. [2014]. Separating out the role of dynamic selection following Corollary 1, we find that it accounts for at least 44% of this observed decline in the 6-month job-finding rate between the start of the spell and at 6 months of the spell. This is only slightly lower than the lower bound during the pre-recession years 2006-2007.

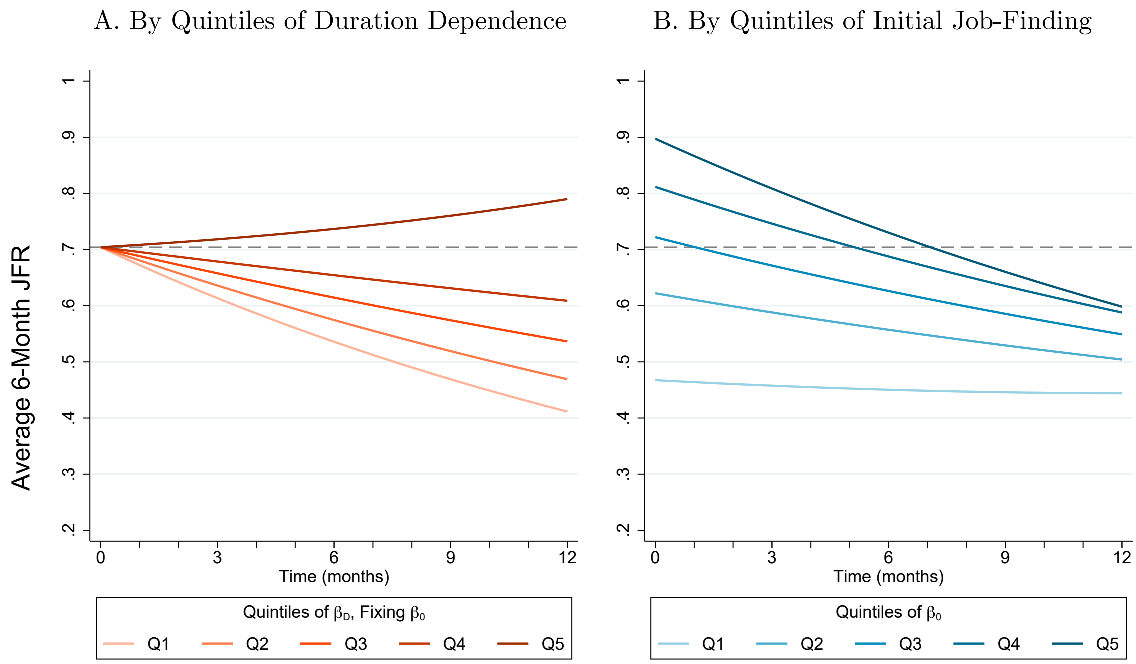
Discussion. Overall, our estimates show that predictable heterogeneity is important in explaining the observed dynamics of job finding over the spell of unemployment. This corroborates the findings in recent work. Mueller et al. [2021] document substantial predictability based on elicited beliefs about the job-finding probability in U.S. survey data. Using a model of beliefs, they then show that selection accounts for 85% of the observed decline of job finding over the first 12 months of the unemployment spell. Alvarez et al. [2022] also find a large role for heterogeneity using repeated unemployment spells.[21] As discussed in section 3.5, their method identifies both observed and unobserved heterogeneity in job-finding risk, but only to the extent that it is persistent between spells. We showed in Table 4 above that the extent of unobserved heterogeneity in the multiple-spell sample of our data is about 58% of the extent of observed heterogeneity. If one assumes that observed heterogeneity accounts for 58% of all of the heterogeneity, i.e., not only for fixed *but also* for heterogeneity in job finding that changes between spells, then one could upscale the covariance in Column 4 of Table 5 above by a factor of $\frac{1}{0.58} = 1.72$. Using the same calculation as above, this would imply that we can account for 84% of the decline in job finding over the spell of unemployment, which is close to the estimate in Mueller et al. [2021].

These results indicate a lesser role for true duration-dependence than often suggested in earlier

[20] Note that these predictions models are indeed estimated on much smaller samples and thus subject to more sampling error as can be seen from the attenuation when comparing outcomes to predictions in Panels B and C of Appendix Figure A14. However, splitting the sample again into 40 groups based on income decile, gender and days on UI and computing the average observed and predicted job-finding rate for each group in the hold-out sample, we find again that the slope is very close to one, suggesting that the predictions remain unbiased.

[21] Note that it is difficult to express their findings in a comparable number as the observed duration dependence in job finding in the Austrian data exhibits a hump-shaped pattern, mostly due to temporary layoffs (Nekoei and Weber [2015]). Still, they do find that dynamic selection is particularly important over the first 6 months of the unemployment spell, which we find as well in our data.

Figure 5: Heterogeneity in Individual Duration-Dependence



Notes: Panel A shows the predicted individual change in the job-finding rate for the five quintiles of the distribution of β_D, assuming a job-finding rate of .70 at the start of the unemployment spell. Panel B shows the predicted change at the individual level in the job-finding rate for the five quintiles of the distribution of β_0.

work. For example, using a resume-audit study, Kroft et al. [2013] find that call-back rates to interviews decline by about 40% over the first 12 months of the unemployment spell. While their study is on the U.S. labor market and for job seekers ages 40 and younger, we believe it is still useful to relate this to our findings. We find that, in the Swedish context, dynamic selection can account for a 10 percentage point decline of the observed decline in job finding from 70% to 49% over the first 12 months of the unemployment spell. This leaves at most a 11 percentage point decline of job finding, or 16 percent decline relative to the initial job-finding rate, for true duration dependence in job finding. Assuming contexts are comparable, this implies that true duration dependence in job finding is substantially less important than duration dependence in call back rates. This is in line with Jarosch and Pilossoph [2018] who show that a decline in call-back rate may not result in a meaningful decline in the job-finding rate.[22]

4.2 Heterogeneity in Individual Duration-Dependence

We can add more structure to the dynamics in job finding and use the data to estimate the observable heterogeneity in these dynamics. In particular, we assume the following model for predicted job-finding

[22]Note that Kroft et al. [2013] find that the decline in call-back rates is less pronounced when the unemployment rate is high. Comparing our estimates prior and during the Great Recession, we find that the observed duration-dependence indeed decreased, but our lower-bound on the role of dynamic selection, if anything, decreases.

rates by duration of unemployment d:

$$\log(\hat{F}_d(X_i)) = \beta_{0,i} + \beta_{D,i} \times d + \varepsilon_{i,d}, \tag{7}$$

where $\hat{F}_d(X_i)$ is the predicted job-finding rate for individual i at unemployment duration d. An important feature of this exercise is that we can estimate this model for any individual, regardless of when he or she found a job, relying on the model predictions given the individual's observables X_i prior to becoming unemployed. That is, we compute $\hat{F}_d(X_i)$ at durations $d = 0, 6$ and 12 months for all individuals unemployed at the start of the spell based on X_i. We then use these predicted job findings to estimate an individual-specific intercept $\beta_{0,i}$ and slope $\beta_{D,i}$. Of course, we should again interpret the estimated slope as an upper-bound, as some of the decrease in job finding for individuals with observables X_i who remain unemployed could be driven by selection on unobservables driving low job-finding. The estimated coefficients $\hat{\beta}_{0,i}$ and $\hat{\beta}_{D,i}$ are also measured with error and thus the dispersion in the estimated coefficients reflects both their true dispersion as well as sampling error. To address this issue, we shrink each individual prediction using the standard errors of the same regressions, as follows:

$$\tilde{\hat{\beta}}_{j,i} = E(\hat{\beta}_{j,i}) + \sqrt{\frac{var(\hat{\beta}_{j,i}) - \hat{\sigma}^2_{j,i}}{var(\hat{\beta}_{j,i})}} (\hat{\beta}_{j,i} - E(\hat{\beta}_{j,i})) \text{ for } j \in \{0, D\}, \tag{8}$$

where $var(\hat{\beta}_{j,i})$ is the sample variance of the predictions and $\hat{\sigma}_{j,i}$ is the standard error corresponding to prediction $\hat{\beta}_{j,i}$.

Our estimates reveal substantial heterogeneity in individual job finding dynamics. The distribution of the (shrunken) estimates of β_D, shown in Appendix Figure A6, displays significant dispersion. We find an estimated standard deviation of 0.0223 with a bootstrapped confidence interval of [0.0222, 0.0225].[23] Because we estimate our prediction in terms of the logs in equation (7) above, this by itself implies that job-finding rates do not differ proportionally throughout the unemployment spell. Importantly, this rejects a common assumption in models of job search. Indeed, a large number of papers have estimated mixed proportional hazard models, assuming full persistence in the job-finding rates over the unemployment spell along observable dimensions. Our rejection corraborates the findings in Alvarez et al. [2022], who find evidence of non-proportionality in data on repeated unemployment spells in Austria.

Panel A of Figure 5 illustrates the magnitude of the estimated heterogeneity in individual dynamics by showing how the job-finding rates evolve over the unemployment spell for the five quintiles of β_D, all starting from the average job-finding rate of 70% for comparability. This clearly evinces the significant dispersion. In fact, for the top quintile, our exercise predicts that the job-finding rate increases over the spell of unemployment, whereas for the bottom quintile it declines by close to 50% over a 12 month spell.[24]

[23] Again following Mullainathan and Spiess [2017], we draw 500 bootstrap samples from the 2006 hold-out sample.

[24] The average decline is 9 percentage points after 6 months, which is close to our estimate of 7.5 percentage points of the residual decline in job-finding after accounting for the role of dynamic selection.

Panel B of Figure 5 shows how the predicted individual level dynamics relate to the initial job-finding rate. More specifically, it shows the predicted changes by quintile of the distribution of the intercept β_0. Clearly, the decline in the job-finding rate is strongly correlated with the initial job-finding rate. In fact, for the bottom quintile of the initial job-finding rate, we do not predict any decline at all. For the top quintile, however, we predict a decline of 30 percentage points. Hence, the predicted job-finding rates would converge as the spell continues for *all* job seekers in the spell. As a result, both the heterogeneous dynamics and the dynamic selection contribute to the compression of the job finding distribution among job seekers who remain unemployed for longer. We note again that we cannot rule out that the heterogeneity in the predicted individual dynamics is driven by differential dynamic selection based on unobservables. However, it seems unlikely that the dynamic selection would be so important for job seekers with the highest predicted job-finding rates and fully absent for job seekers with the lowest predicted job-finding rates at the start of the unemployment spell.

Finally, we briefly study which type of job seekers are more at risk of declining job-finding rates. Following the same approach as for the heterogeneity in job finding at the start of the spell, we correlate observable characteristics with the dynamic component β_D in the right panel of Figure 3. Like in the left panel, we again standardize both the outcome and explanatory variables. We confirm the finding that the most pronounced gradient is in the individuals' job finding at the start of the spell. However, the dynamics in job-finding rates do not only differ across initial job-finding rates, but also conditional on initial job-finding rates. This necessarily leads to rank reversals in predicted job finding over the unemployment spell. Indeed, conditional on initial job finding, we for example find that job seekers experience stronger declines in job finding during unemployment when they are older, have lower income and are less educated.[25] While several observables correlate significantly with the dynamic component β_D, the estimates are all relatively small.[26]

To sum up, we find significant heterogeneity in duration-dependence across individuals with different observed characteristics. Our findings reject the common assumption of proportionality used in standard models of job finding.

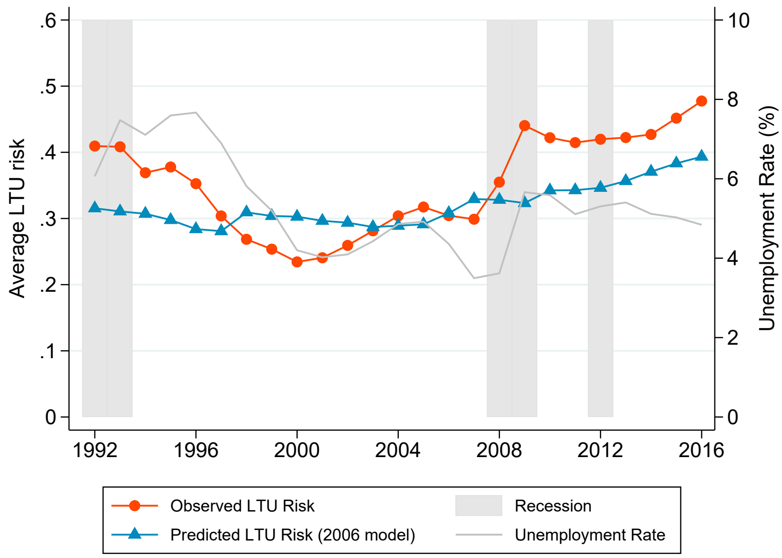
5 Job Finding over the Business Cycle

This section turns from the dynamics in job-finding rates over the unemployment spell to the dynamics in job-finding rates over the business cycle. Like for the dynamics over the spell, an important, outstanding question is to what extent the dynamics over the business cycle are driven by compositional changes in the pool of unemployed or by changes in the job-finding rates themselves. In the latter case, the natural follow-up question is whether the cyclicality in the job-finding rates varies across job seekers.

[25] Our findings are somewhat different from Eriksson and Rooth [2014] who find using randomized CVs that the call back rate is significantly lower only for job seekers after 9 months into unemployment, though they do find heterogeneity in the duration profile as the call backs only decline for job seekers with low-medium skills jobs. Again, they study only call-backs, which may not translate into job finding.

[26] Appendix Figure A9 shows the corresponding bivariate correlations for comparison. We also report the Shapley-Owen decomposition of the R^2 to assess the explanatory power of the different groups of variables in Appendix Figure A10, mirroring the analysis done for the baseline model. The results again confirm that the job-finding rates at the start of the spell jump out in explaining the variation in duration-dependence.

Figure 6: Compositional Effects over the Business Cycle



Notes: The figure shows the averages of 1 minus the 6-month job-finding rate in the hold-out sample for the years 1992-2016, the averages of the predicted long-term unemployment risk using the 2006 model, and the yearly averages of the unemployment rate (see Appendix Figure A5 for a comparison with OECD data). The grey shaded areas correspond to periods with two consecutive quarters of negative growth in Gross Domestic Product.

5.1 Cyclical Selection into Unemployment

We first study the role that selection into the pool of unemployed plays for the cyclicality of the average long-term unemployment risk. Figure 6 shows the average long-term unemployment risk for each year in our sample period as well as the aggregate unemployment rate. There is a strong positive correlation between the two, with a correlation coefficient of 0.33. The increase in long-term unemployment risk is particularly notable during the Great Recession period, as well as the decrease after the large recession in Sweden in the beginning of the 1990s. The substantial variation in long-term unemployment risk over the business cycle is also a feature of other major developed economies, including the U.S. (see, e.g., Elsby et al. [2009], Shimer [2012], among many others).

We revisit the heterogeneity hypothesis, i.e., whether compositional changes in the pool of unemployed do translate into higher LTU risk in recessions, by using the 2006 model to predict the long-term unemployment risk of newly unemployed job seekers in each year in our sample. Figure 6 shows how the average long-term unemployment risk predicted by the 2006 model but in the sample of newly unemployed in each year changes over time.[27] By keeping the prediction model fixed, changes in this counter-factual long-term unemployment risk only reflect compositional changes in the pool of unemployed. As is clear from the figure, there appears to be no distinct relationship of the

[27] Since the LISA panel only starts in 1990 and even later for days spent on UI and DI, we impute the pre-unemployment variables for spells in the earliest years of our sampling variable. In particular, we use the individual's history when partially observed, but use the population mean in 1995 when the individual's history is entirely missing. Given our finding that employment histories from the prior year are the most predictive, the imputation of earlier years does not seem restrictive.

predicted long-term unemployment risk with the aggregate unemployment rate, much in contrast to the observable long-term unemployment risk. As shown in Appendix Table A10 the raw correlation of the predicted and the aggregate unemployment rate is actually slightly negative. The bi-variate regression coefficient with the log unemployment rate as dependent variable turns slightly positive (0.119) when we control for a linear time trend but is only at 15.2% of the one of the regression with observable LTU risk (0.782).[28]

Overall, this shows that – even when using a rich set of observables that is highly predictive of job finding – there is little support for the heterogeneity hypothesis. That is, the increase in long-term unemployment risk in recessions cannot be attributed to changes in the composition of unemployed workers.[29]

5.2 Heterogeneity in Individual Cyclicality

Since compositional changes cannot account for the large increase in long-term unemployment risk in recessions, this means that the same individuals face very different risks depending on when they become unemployed. An important question taken up in the literature is whether individuals differ in their cyclicality of long-term unemployment risk. For example, Bils et al. [2012] find that those who tend to work fewer hours have a lower cyclicality of job finding out of unemployment, but there are little differences in the cyclicality based on education and prior wages [Mueller, 2017]. We revisit this question, simply using richer data and leveraging the extendability of our approach to different years. We perform our prediction exercise for each year in our sample from 1995 to 2016.[30]

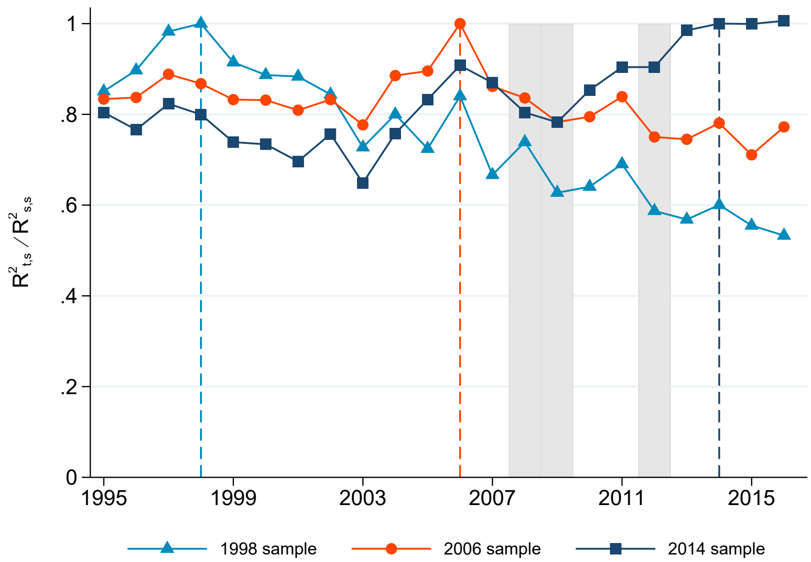
We first consider again how persistent the heterogeneity in job finding, but now over the cycle rather than over the spell. We compare the relative predictive value of models estimated in a different year t than the year of the hold-out sample. In the stylized model presented in Section 2, this relative predictive value separates how important the heterogeneity in permanent types is relative to transitory differences. Figure 7 shows the relative predictive value of models estimated in a different year t for three different hold-out samples (resp. 1998, 2006 and 2014). For example, the orange dotted line shows the R-squared in the hold-out sample of 2006 for models from different years relative to the R-squared for the model of 2006. Not surprisingly, the model from 2006 does best but the fall off in predictive power for other years is relatively modest. But even when using models estimated in more distant years, the decrease in the predictive power is limited. E.g., for the 2014 hold-out sample, the predictive power of the model estimated 20 years earlier is still 80 percent of the predictive power of the model estimated in 2014. This suggests that the features that predict long-term unemployment are relatively stable and that while job-finding rates are cyclical, there are only limited rank reversals in predicted job-finding over time. The estimated persistence in job-finding over time is consistent with our analysis in section 3.5 of multiple spell data, which show a high covariance of job finding across spells.

[28] We find that these conclusions are the same if we restrict the sample to the years where we have complete employment histories available.

[29] Note that selection into long-term unemployment is driven both by selection into unemployment, i.e. at the start of the spell, and selection that occurs over the spell. We find, however, similar patterns when we look at the predicted risk of remaining unemployed from 6 to 12 months or from 12 to 18 months, see Appendix Figure A4.

[30] We start only in 1995 to reduce the censoring of pre-unemployment histories.

Figure 7: Persistence in the Predictive Value over Time



Notes: The figure shows the R^2 in the hold out sample of 1998, 2006 and 2014 using the prediction model from each year *relative to* R^2 in the hold out sample from the same year as the prediction model.

We use the data to estimate how the heterogeneity evolves over the business cycle in parallel to our analysis of heterogeneity in individual duration dependence in the previous section. An advantage of our prediction exercise is that we can predict a hypothetical job-finding rate for each individual and each year in our data, conditional on their observable characteristics. We can then estimate the cyclicality in the predicted 6-month job-finding rates by relating it – for each individual's combination of observable characteristics – to the log of the aggregate unemployment rate, u_t, as follows:

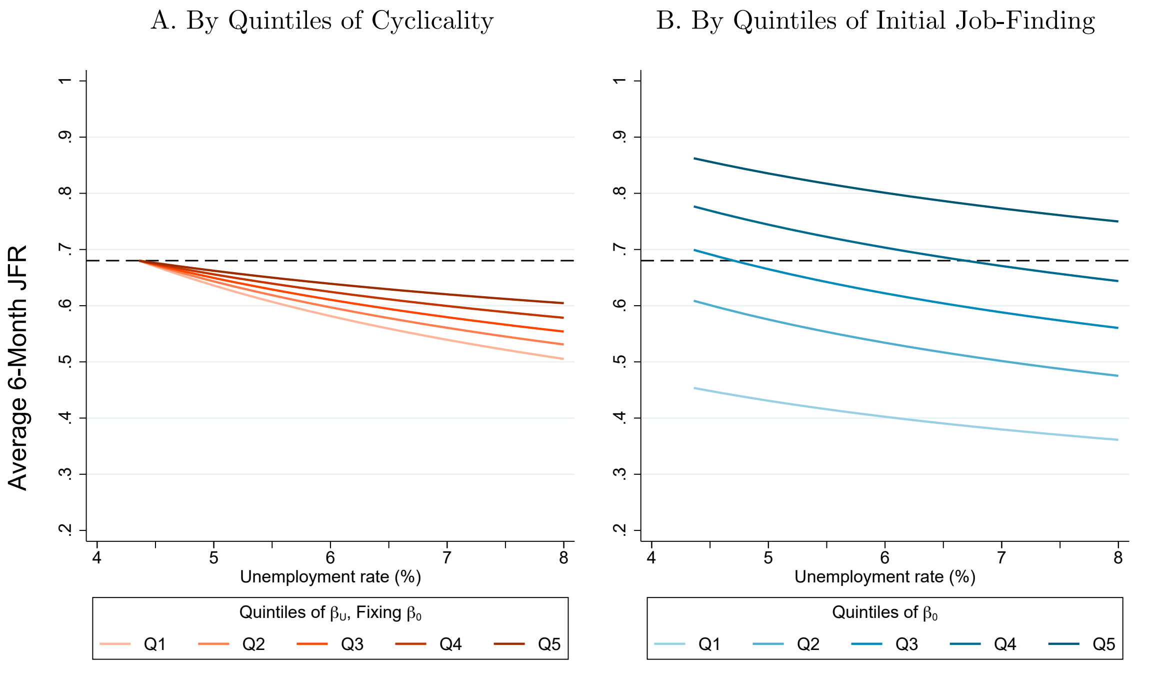
$$\log(\hat{F}_t(X_{i,t_0})) = \beta_{0,i} + \beta_{U,i}(u_t - u_{t_0}) + \beta_{T,i}(t - t_0) + \varepsilon_{i,t}, \tag{9}$$

where we use 2006 as the reference year t_0. Note that we include only individuals who were actually unemployed in this reference year. Moreover, this exercise holds characteristics of individual job seekers constant over time. As we have shown that there is little movement in composition of types, this suggests that the conclusions from our exercise are not affected by these restrictions. We again shrink the distributions of $\hat{\beta}_j$ using the standard errors of the same regressions, as for the duration-dependence estimates (see equation 8 above).

Panel A in Figure 8 illustrates the heterogeneity in the individual cyclicality of job finding by splitting the sample into the five quintiles based on the cyclicality estimate β_U, normalizing the job-finding rate to the one for the 2006 unemployment rate. It shows that there is only moderate dispersion in the changes in the predicted job-finding rate for the range of the unemployment rates observed in Sweden over our sample period, especially in comparison to the dispersion in duration-dependence discussed earlier (see Figure 5).[31] Panel B of Figure 8 illustrates how the predicted individual level

[31] Appendix Figure A7 shows the estimated distributions of the permanent and cyclical component of the job-finding rate. The distribution of the permanent component in Panel A closely resembles the distribution of predicted job-finding

Figure 8: HETEROGENEITY IN INDIVIDUAL CYCLICALITY



Notes: Panel A shows the predicted individual change in the job-finding rate for the five quintiles of the distribution of β_U, normalizing the job-finding rate to the one in 2006. Panel B shows the predicted change at the individual level in the job-finding rate for the five quintiles of the distribution of the intercept β_0.

dynamics relate to the job-finding rate in the reference year. More specifically, it shows the predicted changes by quintile of the distribution of the intercept β_0. The declines in the job-finding rate are quite similar across the five quintiles, very much in contrast with the dynamics over the unemployment spell, where we found large differences in the slopes.

The right panel of Figure 3 shows again how observable characteristics correlate with the cyclical component β_U. We confirm that in contrast to the duration dependence analysis job seekers with higher job-finding rates do not experience larger, but slightly smaller (relative) declines as the unemployment rate increases. We again find that, conditional on initial job finding, job seekers are more shielded against declines in the job-finding rate when they are more educated and had higher labor income prior to unemployment. The estimated differences are more pronounced compared to the duration-dependence analysis.[32]

To sum up, we find moderate dispersion in the individual cyclicality of job finding despite the important cyclicality overall and the large predictable differences in individual job-finding rates at a given moment in time. The predictable heterogeneity in job finding is thus persistent over time, which is also reflected in the persistence of the predictive power of the prediction models.

rates for the year 2006 in Figure 2, as we choose the year 2006 as our reference year. Panel B shows the distribution of the cyclical component. The average cyclicality equals -0.34, implying that if the unemployment rate doubles, the job-finding rate decreases by 34 percent. For the cyclical component, $\hat{\beta}_U$, the shrunken standard deviation is 0.120, which is moderate, and the bootstrapped 95% confidence interval is tight at [0.119, 0.120]. As in previous exercises, we draw 500 bootstrap samples from the 2006 hold-out sample for this last step.

[32]The estimated Shapley values in Appendix Table A10 confirm the relative importance of characteristics other than the 2006 predicted job-finding in explaining the individual cyclicality.

6 Targeting Unemployment Policies

An important consequence of predictable heterogeneity across the unemployed is that unemployment policies can be targeted. We study the potential for targeting unemployment policies to job seekers at risk of long-term unemployment, drawing our motivation from the fact that long-term unemployment is the predominant criteria for the assignment of active labor market policies (ALMP). A number of Public Employment Services (PES) in fact use the expected risk of long-term unemployment to assign job seekers to job search counselling or training programs, either based on specific dimensions (e.g., age, education), as assessed by the case worker or as predicted by statistical profiling models [OECD, 2019]. Alternatively, some programs are limited to the long-term unemployed (e.g., wage subsidies), or are required for long-term unemployed workers to remain eligible for unemployment benefits (e.g., workfare).

Unemployment Policies in Sweden. The Swedish PES has used a wide range of different policies and interventions, as presented in Appendix Table A11. The first category consists of training programs designed to learn vocational skills, non-vocational training and job search assistance and includes 15.7% percent of all unemployment spells in Sweden. The second category consists of work experience and workfare interventions, corresponding to 13.7% percent of unemployment spells. The final category includes wage subsidies provided to employers who hire workers who have been long-term unemployed. Appendix Table A11 reports the timing and duration of the different type of unemployment policies. About 14% percent of the unemployed job seekers in 2006 were subject to an intervention already in the first six months of the spell. However, around 44% of the unemployment policies are taken up by long-term unemployed, even though only 30% of job seekers become long-term unemployed. For the workfare programs and the wage subsidies, the average job seeker has been unemployed for over a year and longer when they get initiated. Only the training and assistance programs start earlier in the spell with about 55% of them initiated before a job seeker reaches six months of unemployment.

The PES case workers in Sweden play a key role in the allocation of ALMPs, but are in some cases limited explicitly by law. The PES centres have also become more centralized, further limiting the discretion by case workers. While the targeting of wage subsidies for example was selective before 2007, *all* unemployed job seekers are now eligible, but only if unemployed for at least one year [Lombardi et al., 2018]. The training programs are targeted already earlier in the spell to workers who 'typically' experience long unemployment spells [Richardson and van den Berg, 2013]. For a few years, the PES even used a risk profiling model to help case workers with the targeting [Bennmarker et al., 2007].

We can use our prediction model to evaluate how successful the targeting of the Swedish unemployment policies has been. Panel A of Figure A12 plots the distribution of predicted job-finding risks for individuals assigned to the ALMPs defined in the narrow sense (i.e., training and job search) in the first six months. If well targeted, we would expect most participants to be in the lower range of the distribution. While the distributions are distinct, there is substantial overlap with the distribution for the non-participants. On average, the predicted LTU risk is 39% for those who have taken up the ALMPs relative to an average of 30% in the total sample of unemployed. Hence, their long-term unemployment risk is larger than average, but the difference is modest, suggesting an important

Table 6: Predicted Job-Finding Probabilities for Different Targets

	At Start of Spell				6 Months into Spell	
	All	**ALMP**	**Target 20%**	**Target 10%**	**LTU**	**Target 67% LTU**
	(1)	(2)	(3)	(4)	(5)	(6)
Pred. Job Finding ($\hat{F}_{i,d}$)	0.70	0.61	0.41	0.33	0.55	0.46
Duration (d)	0	0	0	0	6	6
N	122,590	9,256	24,712	12,259	41,647	27,765

Notes: The table reports mean predicted 6-month job-finding probabilities in different targeted samples in the year 2006. Column (1) corresponds to the full hold-out sample at the beginning of the spell. Column (2) is the subset of spells that included any ALMPs in the narrow sense during the first 6 months of unemployment. Columns (3) and (4) are the bottom quintile and decile of predicted job-finding probability at the start of the spell, respectively. Column (5) looks at those 6 months into the unemployment spell and reports the predicted job-finding probability from 6 to 12 months of the spell. Finally, Column (6) considers the bottom two thirds of predicted job-finding probability 6 months into the spell.

potential for targeting these ALMPs better to those with the highest LTU risk.[33]

Risk-Profiling for Assignment. The use of prediction models can change the assignment of unemployment policies in two ways. First, unemployment policies initiated early in the spell can be targeted to job seekers with a higher risk of long-term unemployment. Second, unemployment policies designed for the long-term unemployed can be assigned earlier in the spell. As noted before, the profiling models used in practice, including by the PES in Sweden, aim to target individuals with the highest long-term unemployment risk. To be efficient, this requires that individuals prone to be long-term unemployed gain more from the interventions and that these gains do not increase as a given individual remains unemployed for longer. Of course, this remains to be shown but is beyond the scope of this paper.[34]

We can easily evaluate these potential advantages of targeting using our prediction model. By targeting the 20% of individuals with the worst employment prospects at the start of the spell, one would reach individuals with a LTU risk of 59 percent on average, see Table 6. If we targeted only the bottom 10%, corresponding to the share of job seekers starting either training programs or job search assistance early, the average LTU risk would be further increased to 67%. This is significantly higher than the average LTU risk of 30%, but also than the predicted LTU risk of 39% for the about 10% of job seekers assigned to training and assistance programs. The prediction model can thus help to significantly increase the predicted long-term unemployment risk of the pool of individuals assigned to ALMPs. Moreover, rather than using the prediction model at the start of the spell, one could have also waited for 3.5 months before assigning job seekers to programs. At this duration, the average probability to be still unemployed six months into the spell has increased to 67%, which corresponds

[33] We find that the differences also remain similar when we retrain the model on the sample of those who did not enter ALMPs in the first 6 months of the spell; see panels C and D of Appendix Figure A12, which show the corresponding distributions of predicted job-finding rates.

[34] Note that Cockx et al. [2023] do find differences in the effect of training programs by unemployment duration, but cannot separate selection vs. true duration-dependence in the treatment effects. There is some theoretical work on this issue. For example, Spinnewijn [2013] links the optimal timing of training programs to the loss of human capital upon job loss vs. its depreciation during unemployment.

Figure 9: The Value of Targeting over the Business Cycle

Notes: The figure shows the ratio of the observed LTU risk for the bottom 20% relative to the average LTU risk in each year.

to the average long-term unemployment risk for the bottom 10%.[35] Hence, presuming that we want to assign programs to job seekers most likely to experience unemployment spells of at least 6 months, we can use the prediction model for assignment at the start of the spell and avoid 3.5 months of unemployment to screen out the low-risk job seekers.

Targeting by Duration of Unemployment. An alternative objective is to target those with the lowest job-finding rate among all the unemployed, i.e. including job seekers at all durations of unemployment. For this, the length of the ongoing unemployment spell can be a useful screening device as those who remain unemployed for longer have lower job-finding rates. However, our analysis suggests that using the duration of unemployment is less effective than using the prediction model. Six months into unemployment, the average job-finding rate over a six-month horizon has decreased from 70% to 55%, for the 34% of job seekers remaining. Even one year into the spell, when only 17% of job seekers remain, the average job-finding rate is still 49%. This remains higher than the 41% probability for the 20% job seekers who can be targeted at the start of the spell (see Table 6). Of course, one can also wait before targeting and apply the prediction model to those who have been unemployed for some time. Still, in the Swedish context, the returns to delaying the targeting of

[35] Appendix Figure A13 shows how the probability to be still unemployed six months into the spell evolves over the first six months of unemployment. Note that this probability increases mechanically as job seekers have a shorter horizon left to find a job before becoming long-term unemployed, but also because of the dynamic selection of less employable types into longer unemployment spells.

workers are limited because of the persistence in heterogeneity.[36]

Targeting over the Cycle. We finally gauge how the value of targeting changes over the business cycle as well as how much one benefits from updating the prediction model over time. We first compare the LTU risk for the bottom 20% relative to the average LTU risk in the hold-out sample for each year between 1995 and 2016. Figure 9 shows that the value of targeting is strongly pro-cyclical. It reaches its peak just before the Great Recession, when the LTU of the targeted group is almost twice as high as the average LTU risk. The value clearly drops during the recessions observed in our sample period.[37] Note that we found earlier that the predicted job finding varies close to proportionally with the unemployment rate, starting from different job-finding rates, but this is not true when converting into predicted LTU risks as illustrated in Appendix Figure A8. Figure 9 also shows the time series of the same statistic, but only using the prediction model in 2006 to rank individuals and determine the bottom 20%. This thus shows the LTU for this group – which would be targeted based on the 2006 model – relative to the average LTU risk, again for each year in our sample period. The comparison of the two time series shows that the loss of not updating the prediction model to select the group of individuals to assign ALMPs is relatively minor. This corroborates the earlier finding in Figure 7, showing that the heterogeneity in job-finding rates is very persistent over the business cycle.

7 Conclusion

This paper uses rich administrative data from Sweden to study the predictability and determinants of long-term unemployment. We find substantial predictable heterogeneity of LTU risk that is driven by the use of comprehensive data on income, employment and benefit histories and show how the predictability in LTU risk relates to issues of selection over unemployment spell and the business cycle.

First, we show that, over the spell of unemployment, our results of substantial predictability have important implications for dynamic selection and the observed duration dependence in job finding. In particular, we show that the persistence in the predictability of job finding is a key statistic that pins down the extent of dynamic selection. We show empirically that the the predictability of job finding is indeed very persistent over the spell of unemployment and that, as a consequence, at least 50% of the observed decline in job finding is driven by dynamic selection. This finding complements recent research that has found an important role for selection but with different and complimentary approaches [Alvarez et al., 2022; Mueller et al., 2021].

Second, we examine to what extent the cyclicality in job finding and LTU risk is driven by cyclical changes in the pool of unemployed. Prior work has found little role for composition, but was limited to socio-demographic variables in survey data. Despite the richer data and the high predictive power for job finding in the cross-section of job seekers, we find no evidence that the rise in LTU risk is

[36] For example, as shown in Column 6 of Table 6, the bottom two thirds of the long-term unemployed (wich corresponds to 23% of the sample at the beginning of the spell) have a predicted chance of 46% to find a job in the next 6 months. Note that this is higher than the 41% probability for the bottom 20% at the start of the spell. Moreover, we would have to wait for six months before assigning them to a policy.

[37] These patterns are robust to excluding training spells or targeting a fixed number of unemployed instead of a fixed percent of unemployed, see Figure A11 in the Appendix.

driven by composition and thus confirm the previous evidence over the business cycle (e.g., Baker [1992] and Kroft et al. [2016]).

Our approach of estimating prediction models for job finding at different stages of the unemployment spell or the business cycle also allows us to predict the duration dependence and the cyclicality of job finding for each individual in our data. We find substantial heterogeneity in the dynamics over the unemployment spell that are inconsistent with the common assumption of proportionality in models of job search. Over the business cycle, instead, we only find limited heterogeneity in the cyclicality of job finding.

Finally, we apply our prediction exercise to the question of targeting and timing of unemployment policies. While we do not know to what extent treatment effects of these policies are heterogeneous by predictable LTU, our starting point is that a common criteria for search policies is the duration of unemployment or the risk of long-term unemployment from profiling models similar to ours. We find that targeting the bottom of the predicted distribution of job finding can save valuable time and is associated with a much higher LTU risk than for the average unemployed. This is particularly so in times of low unemployment, suggesting a procyclical value for the targeting of unemployment policies. A valuable avenue for future research, which is beyond the scope of this paper, would be to identify the effects of unemployment policies by predicted LTU risk, by duration of unemployment as well as the state of the business cycle.

References

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

_ , Bart Hobijn, and Ayşegül Şahin, "The Dual US Labor Market Uncovered," Working Paper 2022.

Alvarez, Fernando E, Katarína Borovičková, and Robert Shimer, "Decomposing Duration Dependence in a Stopping Time Model," Working Paper 2022.

Arni, Patrick, Marco Caliendo, Steffen Künn, and Robert Mahlstedt, "Predicting the Risk of Long-Term Unemployment: What can we learn from Personality Traits, Beliefs and other Behavioral Variables," Technical Report, Working Paper 2014.

Baker, Michael, "Unemployment Duration: Compositional Effects and Cyclical Variability," *The American Economic Review*, 1992, *82* (1), 313–321.

Bennmarker, Helge, Kenneth Carling, and Anders Forslund, "Vem blir långtidsarbetslös?," Technical Report, Institutet för arbetsmarknadspolitisk utvärdering (IFAU) 2007.

Berger, Mark C., Dan Black, and Jeffrey A. Smith, "Evaluating profiling as a means of allocating government services," in Michael Lechner and Friedhelm Pfeiffer, eds., *Econometric Evaluation of Labour Market Policies*, Physica-Verlag HD Heidelberg 2001, pp. 59–84.

Bertheau, Antoine, Edoardo Maria Acabbi, Cristina Barcelo, Andreas Gulyas, Stefano Lombardi, and Raffaele Saggio, "The Unequal Cost of Job Loss across Countries," Working Paper 29727, National Bureau of Economic Research February 2022.

Bils, Mark, Yongsung Chang, and Sun-Bin Kim, "Comparative advantage and unemployment," *Journal of Monetary Economics*, 2012, *59* (2), 150–165.

Black, Dan A., Jeffrey A. Smith, Mark C. Berger, and Brett J. Noel, "Is the Threat of Reemployment Services More Effective Than the Services Themselves? Evidence from Random Assignment in the UI System," *American Economic Review*, September 2003, *93* (4), 1313–1327.

Card, David, Jochen Kluve, and Andrea Weber, "What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations," *Journal of the European Economic Association*, 10 2017, *16* (3), 894–931.

Cockx, Bart, Michael Lechner, and Joost Bollens, "Priority to unemployed immigrants? A causal machine learning evaluation of training in Belgium," *Labour Economics*, 2023, *80*, 102306.

Cox, D. R., "Regression Models and Life-Tables," *Journal of the Royal Statistical Society. Series B (Methodological)*, 1972, *34* (2), 187–220.

Einav, Liran, Amy Finkelstein, Sendhil Mullainathan, and Ziad Obermeyer, "Predictive modeling of U.S. health care spending in late life," *Science*, 2018, *360* (6396), 1462–1465.

Elsby, Michael W. L., Ryan Michaels, and Gary Solon, "The Ins and Outs of Cyclical Unemployment," *American Economic Journal: Macroeconomics*, January 2009, *1* (1), 84–110.

Elsby, Michael W.L., Bart Hobijn, and Ayşegül Şahin, "On the importance of the participation margin for labor market fluctuations," *Journal of Monetary Economics*, 2015, *72*, 64–82.

Eriksson, Stefan and Dan-Olof Rooth, "Do Employers Use Unemployment as a Sorting Criterion When Hiring? Evidence from a Field Experiment," *American Economic Review*, March 2014, *104* (3), 1014–39.

Farber, Henry S., Dan Silverman, and Till von Wachter, "Determinants of Callbacks to Job Applications: An Audit Study," *American Economic Review*, May 2016, *106* (5), 314–18.

Gregory, Victoria, Guido Menzio, and David G Wiczer, "The Alpha Beta Gamma of the Labor Market," Working Paper 28663, National Bureau of Economic Research April 2021.

Grömping, Ulrike, "Estimators of Relative Importance in Linear Regression Based on Variance Decomposition," *The American Statistician*, 2007, *61* (2), 139–147.

Hall, Robert E and Marianna Kudlyak, "Job-Finding and Job-Losing: A Comprehensive Model of Heterogeneous Individual Labor-Market Dynamics," Working Paper 25625, National Bureau of Economic Research February 2019.

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

Heckman, J. and B. Singer, "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 1984, *52* (2), 271–320.

Honoré, Bo E., "Identification Results for Duration Models with Multiple Spells," *The Review of Economic Studies*, 01 1993, *60* (1), 241–246.

Huettner, Frank and Marco Sunder, "Axiomatic arguments for decomposing goodness of fit according to Shapley and Owen values," *Electronic Journal of Statistics*, 2012, *6* (none), 1239 – 1250.

Jarosch, Gregor and Laura Pilossoph, "Statistical Discrimination and Duration Dependence in the Job Finding Rate," *The Review of Economic Studies*, 09 2018, *86* (4), 1631–1665.

Kolsrud, Jonas, Camille Landais, Peter Nilsson, and Johannes Spinnewijn, "The Optimal Timing of Unemployment Benefits: Theory and Evidence from Sweden," *American Economic Review*, April 2018, *108* (4-5), 985–1033.

Kroft, Kory, Fabian Lange, and Matthew J. Notowidigdo, "Duration Dependence and Labor Market Conditions: Evidence from a Field Experiment*," *The Quarterly Journal of Economics*, 06 2013, *128* (3), 1123–1167.

_ , _ , _ , and Lawrence F. Katz, "Long-Term Unemployment and the Great Recession: The Role of Composition, Duration Dependence, and Nonparticipation," *Journal of Labor Economics*, 2016, *34* (S1), S7–S54.

Krueger, Alan B, Judd Cramer, and David Cho, "Are the Long-Term Unemployed on the Margins of the Labor Market?," *Brookings Papers on Economic Activity*, 2014, *2014* (1), 229–299.

Lancaster, Tony, "Econometric Methods for the Duration of Unemployment," *Econometrica*, 1979, *47* (4), 939–956.

Landais, Camille and Johannes Spinnewijn, "The Value of Unemployment Insurance," *The*

Review of Economic Studies, 09 2021, *88* (6), 3041–3085.

Lombardi, Stefano, Oskar Nordström Skans, and Johan Vikström, "Targeted wage subsidies and firm performance," *Labour Economics*, 2018, *53*, 33–45. European Association of Labour Economists 29th annual conference, St.Gallen, Switzerland, 21-23 September 2017.

Machin, Stephen and Alan Manning, "Chapter 47 The causes and consequences of longterm unemployment in Europe," in "Handbook of Labor Economics," Vol. 3, Elsevier, 1999, pp. 3085–3139.

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

Mueller, Andreas I., "Separations, Sorting, and Cyclical Unemployment," *American Economic Review*, July 2017, *107* (7), 2081–2107.

_ , Johannes Spinnewijn, and Giorgio Topa, "Job Seekers' Perceptions and Employment Prospects: Heterogeneity, Duration Dependence, and Bias," *American Economic Review*, January 2021, *111* (1), 324–63.

Mullainathan, Sendhil and Jann Spiess, "Machine Learning: An Applied Econometric Approach," *Journal of Economic Perspectives*, May 2017, *31* (2), 87–106.

Nekoei, Arash and Andrea Weber, "Recall Expectations and Duration Dependence," *American Economic Review*, May 2015, *105* (5), 142–46.

OECD, "Statistical profiling in public employment services," Technical Report 224, OECD 2019.

Pavoni, Nicola and G. L. Violante, "Optimal Welfare-to-Work Programs," *The Review of Economic Studies*, 01 2007, *74* (1), 283–318.

Richardson, Katarina and Gerard J. van den Berg, "Duration Dependence versus Unobserved Heterogeneity in Treatment Effects: Swedish Labor Market Training and the Transition Rate To Employment," *Journal of Applied Econometrics*, 2013, *28* (2), 325–351.

Schmieder, Johannes F, Till M von Wachter, and Jörg Heining, "The Costs of Job Displacement over the Business Cycle and Its Sources: Evidence from Germany," Working Paper 30162, National Bureau of Economic Research June 2022.

Schmieder, Johannes F., Till von Wachter, and Stefan Bender, "The Effects of Extended Unemployment Insurance Over the Business Cycle: Evidence from Regression Discontinuity Estimates Over 20 Years," *The Quarterly Journal of Economics*, 04 2012, *127* (2), 701–752.

_ , _ , and _ , "The Effect of Unemployment Benefits and Nonemployment Durations on Wages," *American Economic Review*, March 2016, *106* (3), 739–77.

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

_ and Ivan Werning, "Liquidity and Insurance for the Unemployed," *American Economic Review*, December 2008, *98* (5), 1922–42.

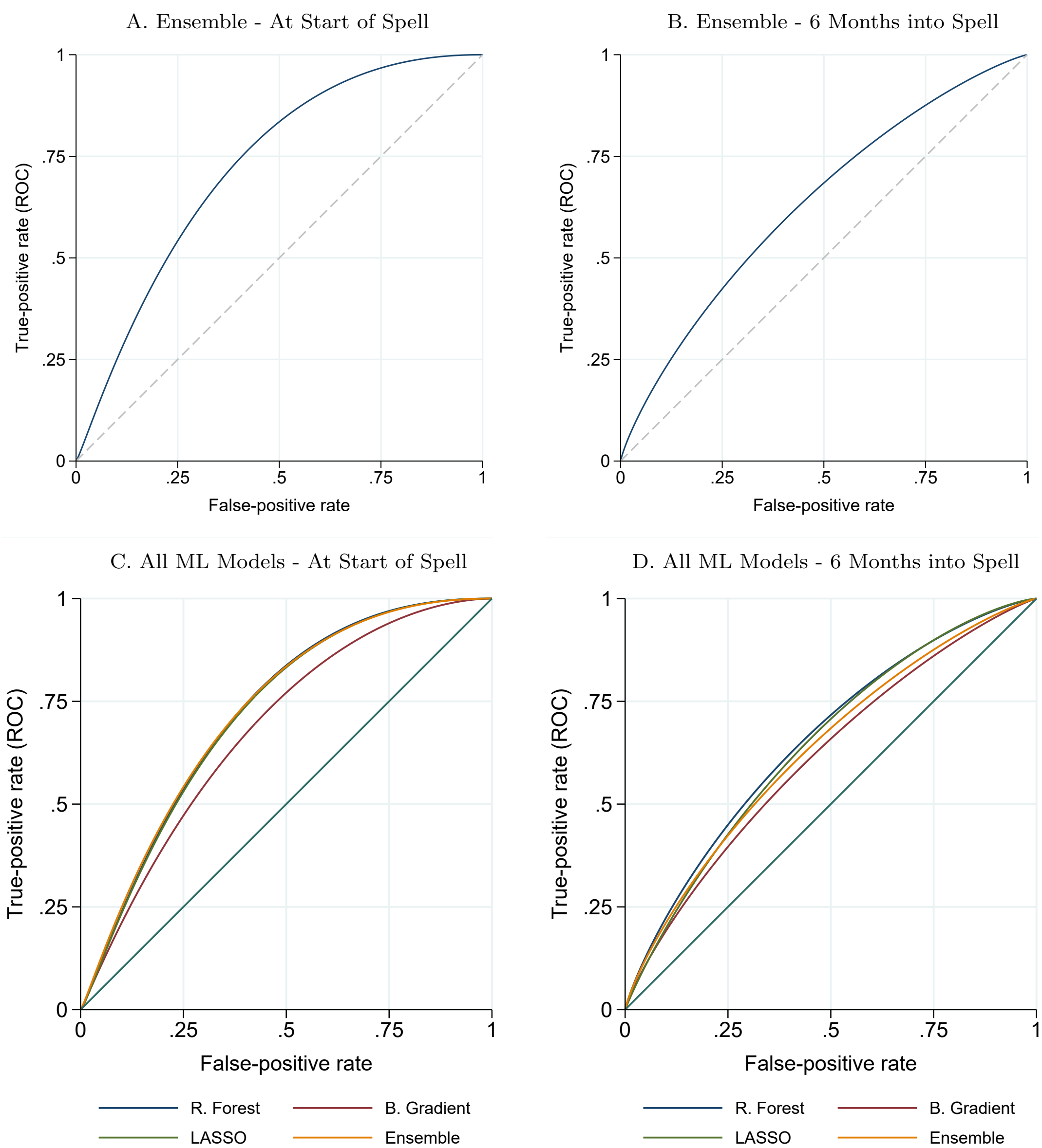
Spinnewijn, Johannes, "Training and search during unemployment," *Journal of Public Economics*, 2013, *99*, 49–65.

Appendix

A Additional Figures and Tables

A.1 Predictive Power: ROC Curves

Figure A1: Predictive Power of Ensemble Model



Notes: The Receiver Operating Characteristic (ROC) curves plot the combinations of true-positive and false-positive rates attained by binary classifiers based on various thresholds of our predicted job-finding probabilities. Panels A and B focus on the ensemble model, while panels C and D also show the three underlying ML models (random forest, gradient-boosted decision trees and lasso). All curves shown correspond to the hold-out sample for the year 2006.

A.2 Predictive Power: Different Horizons and Robustness

Table A1: Robustness: Job Finding over Different Horizons

		$E(\cdot)$		$Var(\cdot)$		$Cov(\cdot)$		$R^2(\cdot)$	
Job Finding Horizon	N	F	$\hat{F}$	F	$\hat{F}$	$\hat{F}, F$	$\hat{F}, F_{6m}$	$\hat{F}, F$	$\hat{F}, F_{6m}$
3 Months	122,590	0.483	0.484	0.250	0.030	0.027	0.028	0.097	0.122
6 Months	122,590	0.700	0.699	0.210	0.033	0.031	0.031	0.136	0.136
12 Months	122,590	0.860	0.861	0.120	0.017	0.016	0.021	0.132	0.131

Notes: The table reports summary statistics about observed and predicted job-finding probabilities at the start of the spell over different horizons. We consider job finding over three horizons: three months, six months (the baseline) and twelve months since the beginning of the spell.

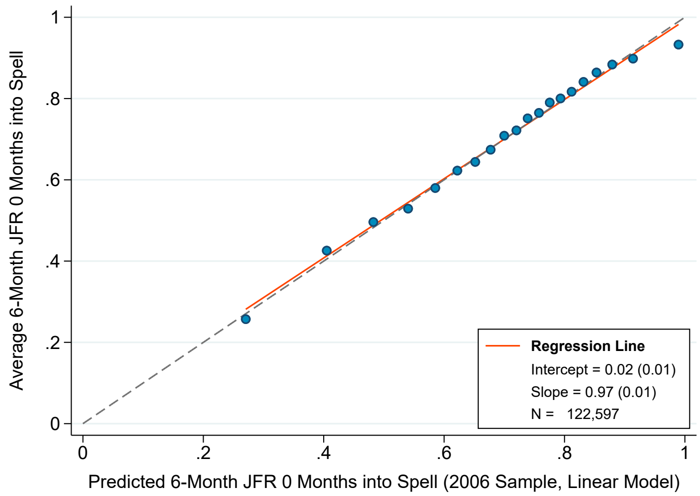
Table A2: Robustness: ALMPs and Job Finding Definition

Model	Sample	N	$E(F_{i,0})$	$E(\hat{F}_{i,0})$	$Var(F_{i,0})$	$Var(\hat{F}_{i,0})$	$Cov(\hat{F}_{i,0}, F_{i,0})$	$R^2(\hat{F}_{i,0}, F_{i,0})$
A. Baseline								
Baseline	All	122,590	0.700	0.699	0.210	0.033	0.031	0.136
B. Robustness to ALMPs								
Baseline	No ALMPs	113,334	0.730	0.706	0.197	0.032	0.029	0.130
No ALMPs	All	122,590	0.700	0.726	0.210	0.030	0.029	0.134
No ALMPs	No ALMPs	113,334	0.730	0.731	0.197	0.030	0.028	0.133
C. Robustness to job finding definition								
Baseline	No AvOrs 7-8	105,300	0.683	0.698	0.216	0.033	0.031	0.133
Baseline	No AvOrs 5-8	65,590	0.742	0.743	0.192	0.026	0.024	0.117
D. Robustness to functional form								
Linear	All	122,590	0.700	0.700	0.210	0.030	0.030	0.138

Notes: The table reports summary statistics about observed and predicted job-finding probabilities at the start of the spell for different combinations of models and samples. Models considered include the baseline 2006 model, a model trained on the subset of spells that did not include ALMPs (in the narrow sense) during the first six months of unemployment ("No ALMPs"), and the linear regression model ("Linear"). The samples on which we evaluate the predictions are the full 2006 hold-out sample ("All") and several subsets thereof: excluding spells that include ALMPs during the first six months ("No ALMPs"), excluding spells that ended because the job seeker entered education other than training or died ("No AvOrs 7-8") and excluding spells that ended because the job seeker terminated contact with PES for unspecified or unknown reasons, entered education other than training or died ("No AvOrs 5-8").

A.3 Predictive Power: Linear Model

Figure A2: Comparing Predictions to Outcomes: Linear Model



Notes: The figure shows a binned scatter plot of observed job finding and the predictions of the linear model. That is to say, we split the hold-out sample into 20 vigintiles of predicted 6-month job-finding probability and report, for each bin, mean observed and predicted 6-month job-finding rates at the start of the spell. The red line shows the results of a linear regression at the individual level of a dummy for finding a job within 6 months on the predicted 6-month job-finding probability.

Table A3: R^2 for various submodels in the year 2006: ML Model vs Linear Model

	A. ML Model							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$R^2(\hat{F}_{i,0}, F_{i,0})$	0.057	0.086	0.094	0.128	0.132	0.136	0.139	0.136
Change (j) vs $(j-1)$	-	+51.0%	+8.9%	+37.2%	+2.7%	+3.2%	+2.3%	-2.4%
	B. Linear model							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$R^2(\hat{F}_{i,0}, F_{i,0})$	0.062	0.086	0.092	0.123	0.125	0.131	0.133	0.138
Change (j) vs $(j-1)$	-	+37.5%	+7.8%	+33.4%	+1.7%	+4.2%	+2.1%	+3.4%
Socio-demographics	X	X	X	X	X	X	X	X
Labour Income		X	X	X	X	X	X	X
Other Income			X	X	X	X	X	X
Employment History				X	X	X	X	X
Income History					X	X	X	X
Migration History						X	X	X
Industry							X	X
Municipality								X

Notes: The table shows the R^2 of the predicted 6-month job-finding probability and a dummy for actual job finding in the hold-out sample for the year 2006 for various models. Panel A reproduces Panel A in Table 3 for convenience. Panel B shows results from linear regression models that use the same variable groups, starting from the basic model in (1) and adding variable groups sequentially until all of the groups included in the baseline model are incorporated in (8).

A.4 Predictive Power: The Role of Pre-Unemployment History Variables

Table A4: R^2 DEPENDING ON PRE-UNEMPLOYMENT HISTORY VARIABLES

A. Groups of variables: sequential sub-models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$R^2(\hat{F}_{i,0}, F_{i,0})$	0.058	0.108	0.114	0.117	0.118	0.120	0.125
Change (j) vs $(j-1)$	-	+87.0%	+5.7%	+2.5%	+0.5%	+2.3%	+4.1%
Basic Socio-demographics	X	X	X	X	X	X	X
Individual History in $t-1$		X	X	X	X	X	X
Individual History in $t-2$			X	X	X	X	X
Individual History in $t-3$				X	X	X	X
Individual History in $t-4$					X	X	X
Individual History in $t-5$						X	X
Firm Characteristics							X

B. Groups of variables: marginal sub-models

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$R^2(\hat{F}_{i,0}, F_{i,0})$	0.058	0.108	0.087	0.070	0.076	0.071	0.087
Change (j) vs (1)	-	+87.0%	+51.4%	+21.2%	+32.0%	+23.1%	+50.0%
Variables:							
Basic Socio-demographics	X	X	X	X	X	X	X
Individual History in $t-1$		X					
Individual History in $t-2$			X				
Individual History in $t-3$				X			
Individual History in $t-4$					X		
Individual History in $t-5$						X	
Firm Characteristics							X

C. Individual variables: sequential sub-models

	(1)	(2)	(3)	(4)	(5)
$R^2(\hat{F}_{i,0}, F_{i,0})$	0.058	0.070	0.076	0.097	0.110
Change (j) vs $(j-1)$	-	+20.8%	+9.7%	+26.4%	+13.9%
Basic Socio-demographics	X	X	X	X	X
Days on UI (2y)		X	X	X	X
Unemp. Spells (2y)			X	X	X
Number of Firms (2y)				X	X
Days on DI (2y)					X

D. Individual variables: marginal sub-models

	(1)	(2)	(3)	(4)	(5)
$R^2(\hat{F}_{i,0}, F_{i,0})$	0.058	0.070	0.074	0.079	0.075
Change (j) vs (1)	-	+20.8%	+28.3%	+37.1%	+30.8%
Basic Socio-demographics	X	X	X	X	X
Days on UI (2y)		X			
Unemp. Spells (2y)			X		
Number of Firms (2y)				X	
Days on DI (2y)					X

Notes: The table shows the R^2 of the predicted 6-month job-finding probability and a dummy for actual job finding in the hold-out sample for the year 2006 for various models. Panel A starts from the basic model in (1) and adds years of pre-unemployment history variables (including days on UI, days on DI, number of unemployment spells and number of firms) and firm characteristics (tenure, size, change in size and layoff rate)) sequentially, while Panel B adds the same groups one at a time. Panels C and D do the same for a selection of individual variables.

A.5 Predictive Power: SILC Survey Data

Table A5: Regressions with SILC survey data

	General Health (GH)			**Mental Health (MH)**			**GH for MH sample**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pred. JFR	1.082	1.054		0.992	0.890		0.957	0.973	
	(0.098)	(0.101)		(0.347)	(0.353)		(0.348)	(0.360)	
Health PC1		0.014	0.040					-0.008	0.019
		(0.011)	(0.012)					(0.043)	(0.043)
Mental Health PC1					0.046	0.064			
					(0.033)	(0.034)			
R^2	0.153	0.155	0.016	0.095	0.117	0.044	0.089	0.090	0.003
Adj. R^2	0.142	0.142	0.014	0.083	0.094	0.031	0.078	0.066	-0.010
N	735	735	735	80	80	80	79	79	79

Notes: This table presents output from linear regressions of observed job finding on the predicted job-finding rate and measures of general and mental health obtained from the EU-SILC survey. Our measure of general health ("Health PC1") is constructed from three survey questions: general health (PH010), suffering from any chronic illness (PH020) and limitation in activities because of health problems (PH030). The mental health index ("Mental Health PC1") is constructed from five questions: overall life satisfaction (PW010), meaning of life (PW020), being very nervous (PW050), feeling "down in the dumps" (PW060) and feeling downhearted or depressed (PW080). In both cases, the index used in the regressions is the first principal component of the matrix of relevant survey answers. For the regressions, we match individual spells in our hold-out samples from 1992 to 2016 with responses to the survey, with Columns (1)-(3) including spells matched with general health answers, (4)-(6) with mental health answers and (7)-(9) with both. Note that the mental health module was only included in the 2013 version of the survey, hence the lower number of matches in Columns (4)-(9). The results show that general health does not add much explanatory power, potentially because our prediction model already incorporates this information via the number of days spent on DI in the years before the unemployment spell. In contrast, adding mental health increases the R^2 by 23% and adjusted R^2 by 13%, although the small sample size raises concerns about overfitting. A simple placebo exercise, where we perform the general health regressions on the mental health sample and find virtually no change in the R^2 and a decrease in the adjusted R^2, suggests that the added explanatory power of the mental health variables is not simply due to the small sample size.

A.6 Predictive Power: Extended Models

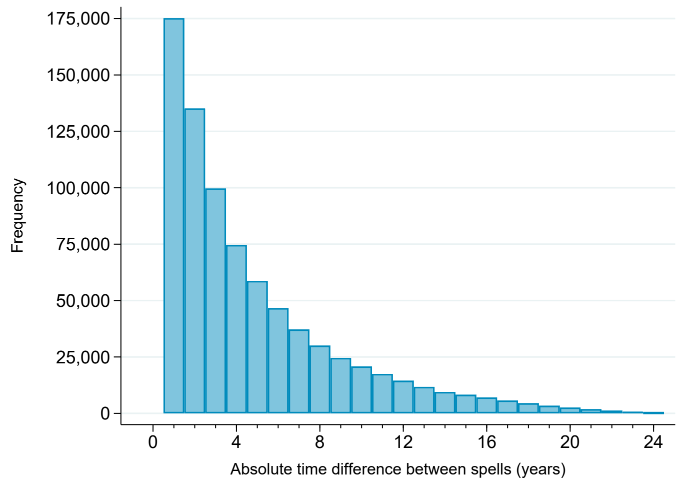
Table A6: R^2 FOR EXTENDED MODELS: STARTING FROM BASIC

	Extensions of Basic							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$R^2(\hat{F}_{i,0}, F_{i,0})$	0.057	0.082	0.067	0.068	0.063	0.068	0.068	0.091
Change (j) vs (1)	-	+43.5%	+18.2%	+19.7%	+9.8%	+18.9%	+19.3%	+60.6%
Socio-demographics	X	X	X	X	X	X	X	X
Occupation		X						X
Union member			X					X
Wealth				X				X
IQ					X			X
UI choice						X		X
UI benefits							X	X

Notes: The table shows the R^2 of the predicted 6-month job-finding probability and a dummy for actual job finding in the hold-out sample for the year 2006 for various models. We start from the basic model using only socio-demographic information in column (1) and add additional information from other administrative data sets, first one at a time and then all at once in column (8).

A.7 Multiple Spell Data and Results

Figure A3: Two-Spell Sample: Time Difference Between Spells



Notes: The figure shows the distribution of the calendar year difference, in absolute terms, between the start of the two unemployment spells for individuals in our two-spell sample, as described in Table 4. The resulting sample consists of 791,524 individuals.

Table A7: IDENTIFYING HETEROGENEITY USING REPEATED SPELLS VS. OBSERVABLES

	Two-spell sample			Control Sample
	$F^1_{i,0}, F^2_{i,0}$	$\hat{F}^1_{i,0}, F^2_{i,0}$	$\hat{F}^2_{i,0}, F^2_{i,0}$	$\hat{F}^2_{i,0}, F^2_{i,0}$
	(1)	(2)	(3)	(4)
	A. Spells in different calendar years			
Cov(·)	0.031	0.018	0.025	0.029
$R^2(\cdot)$	0.019	0.053	0.102	0.122
N	791,524	791,524	791,524	791,524
	B. Spells more than 2 years apart			
Cov(·)	0.026	0.014	0.026	0.029
$R^2(\cdot)$	0.012	0.032	0.105	0.123
N	481,205	481,205	481,205	481,205
	C. Spells more than 5 year apart			
Cov(·)	0.019	0.009	0.025	0.030
$R^2(\cdot)$	0.007	0.013	0.097	0.124
N	248,173	248,173	248,173	248,173

Notes: This table reports key statistics for the sample of individuals with multiple unemployment spells between 1992 and 2016. For a description of the samples and statistics, see the note to Table 4. Panel A reproduces Table 4, while panels B and C simply restrict the samples in A to spells more than 2 and 5 years apart, respectively.

A.8 Dynamics over Spell: Robustness

Table A8: Models Trained of Different Samples: Pooled 2006-2007 Data

Sample	**Model**	N	$E(F)$	$E(\hat{F})$	$Var(\hat{F})$	$Cov(\hat{F}, F)$	$R^2(\hat{F}, F)$
At Start of Spell	0M Model	220,439	0.704	0.702	0.035	0.034	0.156
	6M Model	220,439	0.704	0.594	0.023	0.022	0.100
	12M Model	220,439	0.704	0.525	0.021	0.013	0.041
6M into Spell	6M Model	72,347	0.538	0.537	0.024	0.020	0.063
	12M Model	72,347	0.538	0.489	0.020	0.012	0.031
12M into Spell	12M Model	37,242	0.481	0.474	0.020	0.009	0.018

Notes: The table reports summary statistics about models trained on different unemployment durations in 2006 and 2007. The first three rows correspond to the hold-out sample at the start of the unemployment spell. We generate three different predictions for this sample, with models trained at the start of the spell (the contemporaneous predictions), 6 months into the spell and 12 months into the spell. Rows 4 and 5 deal with the hold-out sample 6 months into the spell; for this sample, we generate predictions using the models trained contemporaneously and 12 months into the spell. Row 6 presents results for the hold-out sample 12 months into the spell, using the contemporaneous model.

Table A9: Models Trained of Different Samples: Pooled 2009-2010 Data

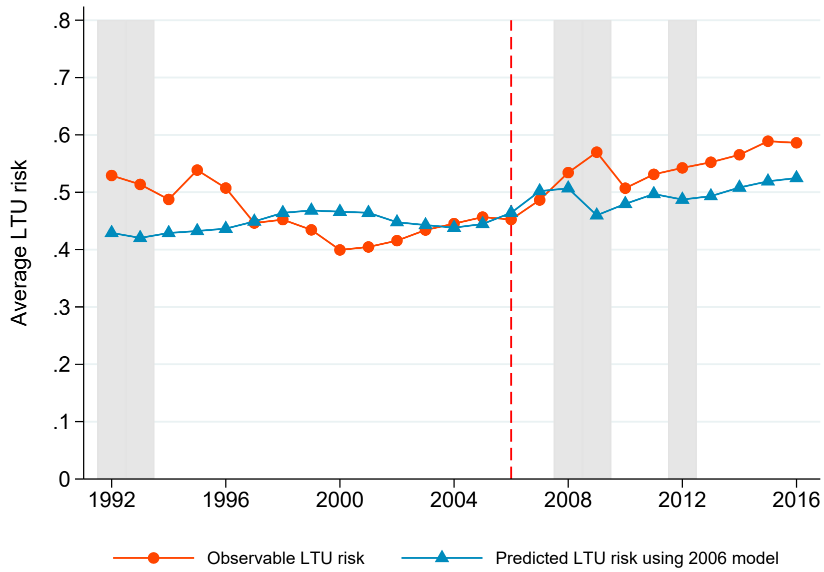
Sample	**Model**	N	$E(F)$	$E(\hat{F})$	$Var(\hat{F})$	$Cov(\hat{F}, F)$	$R^2(\hat{F}, F)$
At Start of Spell	0M Model	229,444	0.577	0.572	0.031	0.031	0.123
	6M Model	229,444	0.577	0.522	0.020	0.020	0.079
	12M Model	229,444	0.577	0.482	0.017	0.011	0.029
6M into Spell	6M Model	96,336	0.469	0.473	0.018	0.014	0.044
	12M Model	96,336	0.469	0.450	0.017	0.009	0.019
12M into Spell	12M Model	47,504	0.426	0.426	0.016	0.009	0.022

Notes: The table reports summary statistics about models trained on different unemployment durations in 2009 and 2010. The first three rows correspond to the hold-out sample at the start of the unemployment spell. We generate three different predictions for this sample, with models trained at the start of the spell (the contemporaneous predictions), 6 months into the spell and 12 months into the spell. Rows 4 and 5 deal with the hold-out sample 6 months into the spell; for this sample, we generate predictions using the models trained contemporaneously and 12 months into the spell. Row 6 presents results for the hold-out sample 12 months into the spell, using the contemporaneous model.

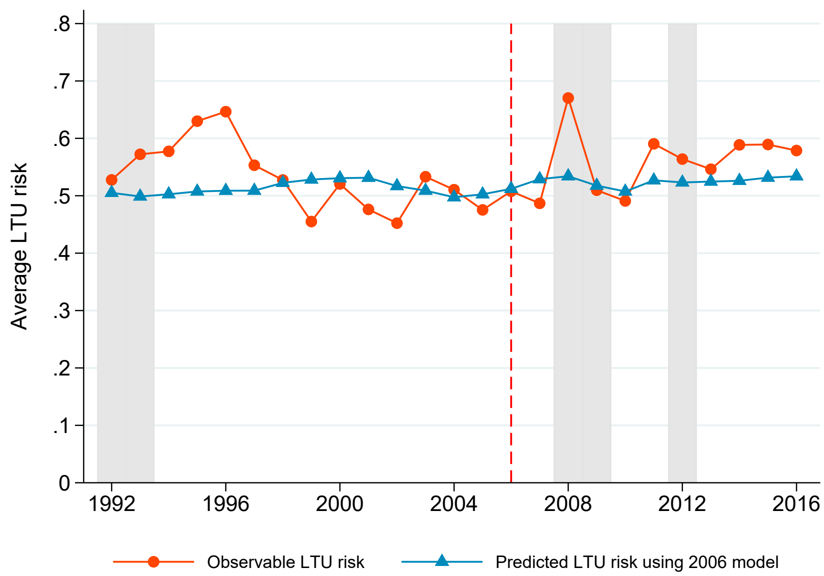
A.9 Selection over Business Cycle: Robustness

Figure A4: Average Risk and Selection into Long-Term Unemployment

A. 6 Months into Spell

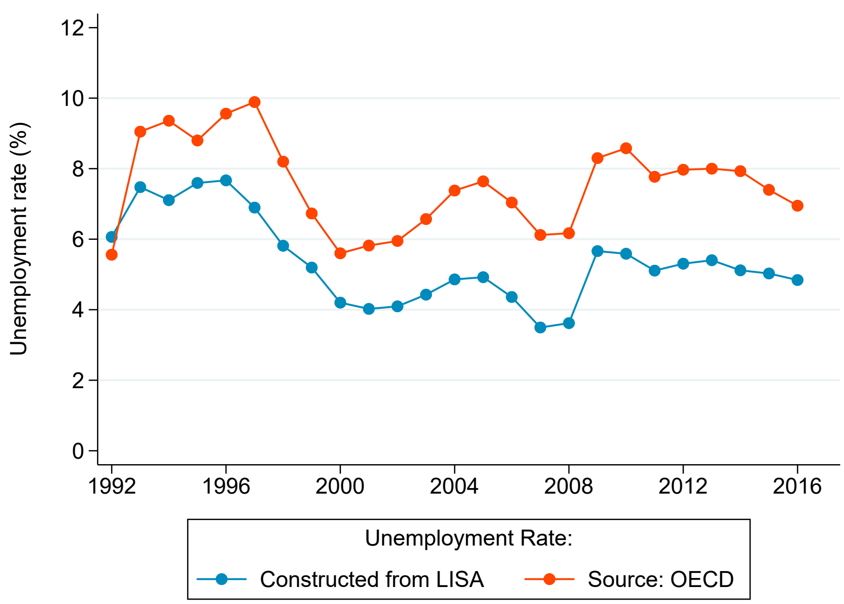


B. 12 Months into Spell



Notes: The figure shows the averages of 1 minus observed and predicted 6-month job-finding rates at different unemployment durations for the hold-out sample for the years 1992-2016. Panel A shows unemployment risk between the 6th and 12th months of unemployment, while Panel B shows unemployment risk between the 12th and 18th months. Predictions are obtained using the corresponding model trained on 2006 data. The grey shaded areas correspond to periods with two consecutive quarters of negative growth in Gross Domestic Product.

Figure A5: Unemployment Rate: LISA vs OECD



Notes: The figure compares the unemployment rate computed from the LISA panel and the official OECD statistics between 1992 and 2016. We use the LISA series throughout the analysis.

Table A10: Relationship between Unemployment and Long-Term Unemployment Risk

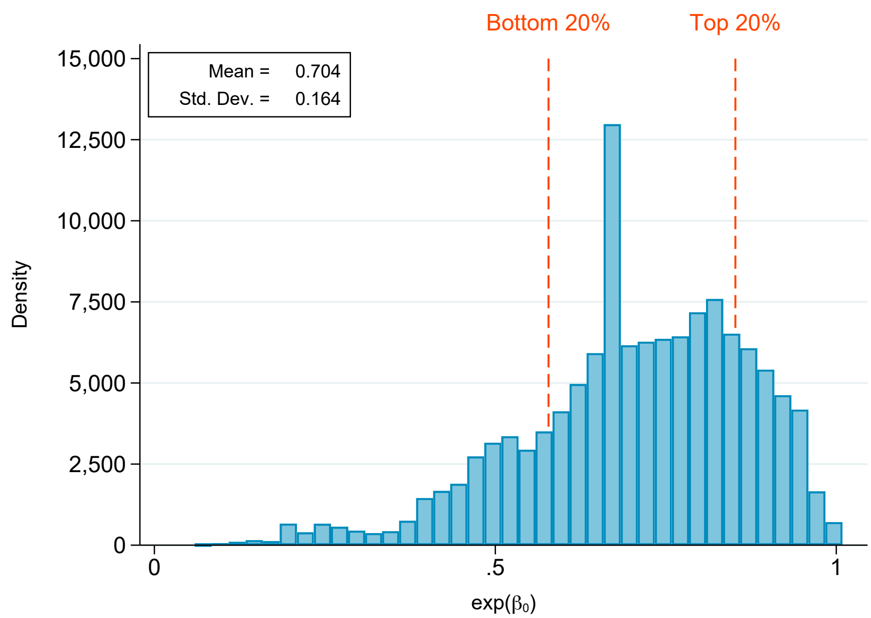
	Predicted log LTU risk (2006)		Observed log LTU risk	
	(1)	(2)	(3)	(4)
	A. 1992-2016			
Log unemployment rate	-0.072	0.119	0.384	0.782
	(0.088)	(0.060)	(0.189)	(0.135)
Time trend		0.012		0.025
		(0.002)		(0.004)
R^2	0.028	0.670	0.152	0.682
Adj. R^2	-0.014	0.640	0.116	0.653
Observations	25	25	25	25
	B. 1995-2016			
Log unemployment rate	-0.068	0.079	0.366	0.688
	(0.108)	(0.050)	(0.229)	(0.088)
Time trend		0.015		0.032
		(0.002)		(0.003)
R^2	0.019	0.817	0.113	0.889
Adj. R^2	-0.030	0.798	0.069	0.877
Observations	22	22	22	22

Notes: The table shows the results of linear regressions of the log of predicted and observed long-term unemployment risk on the log of the aggregate unemployment rate (1-4) and a linear time trend (2 and 4). Panel A uses every year in our sample period, while Panel B restricts to 1995-2016 to avoid early censoring of income and employment histories.

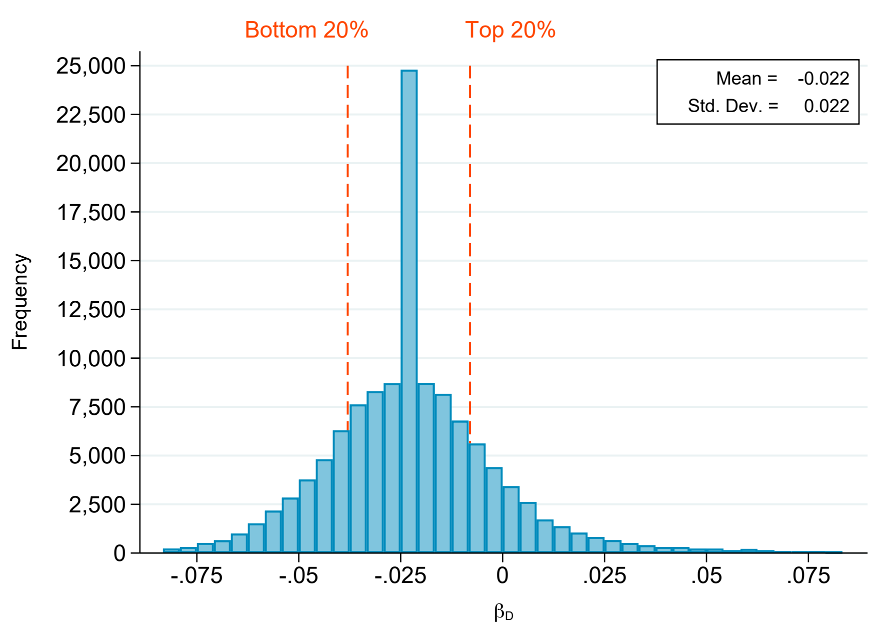
A.10 Heterogeneity in Dynamics over Spell: Additional Results

Figure A6: Distribution of Permanent and Duration-Dependent Component of Job-Finding Risk

A. Histogram of $\exp(\hat{\beta}_0)$



B. Histogram of $\hat{\beta}_D$



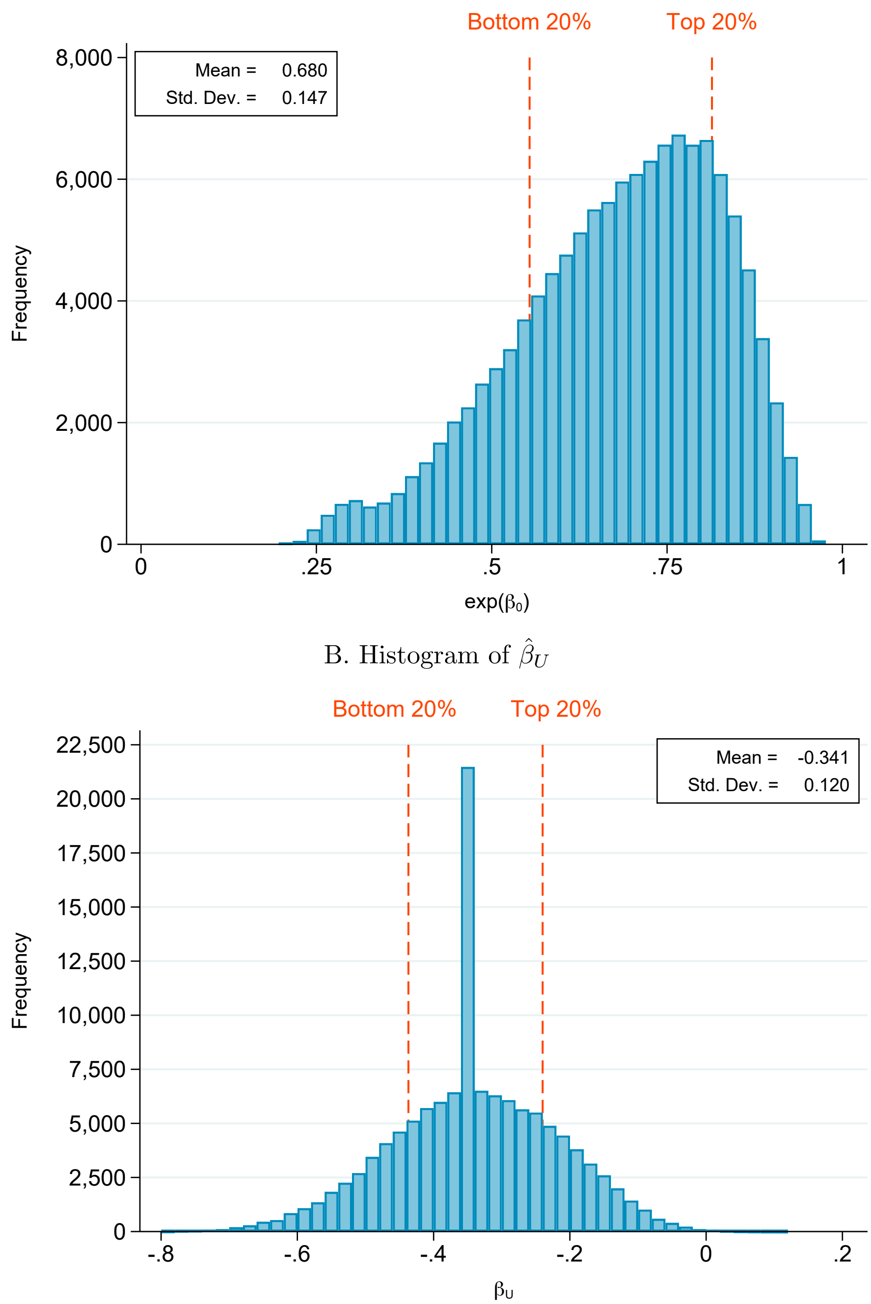
Notes: The figure shows the distribution of the coefficients from the individual-level regressions outlined in equation 7, after applying the shrinkage in equation 8. Panel A shows the histogram of the exponential of the intercept $\exp(\hat{\beta}_0)$, while Panel B shows the histogram of the duration-dependence coefficient $\hat{\beta}_D$.

A.11 Heterogeneity in Dynamics over Business Cycle: Additional Results

Figure A7: Distribution of Permanent and Cyclical Component of Job-Finding Risk

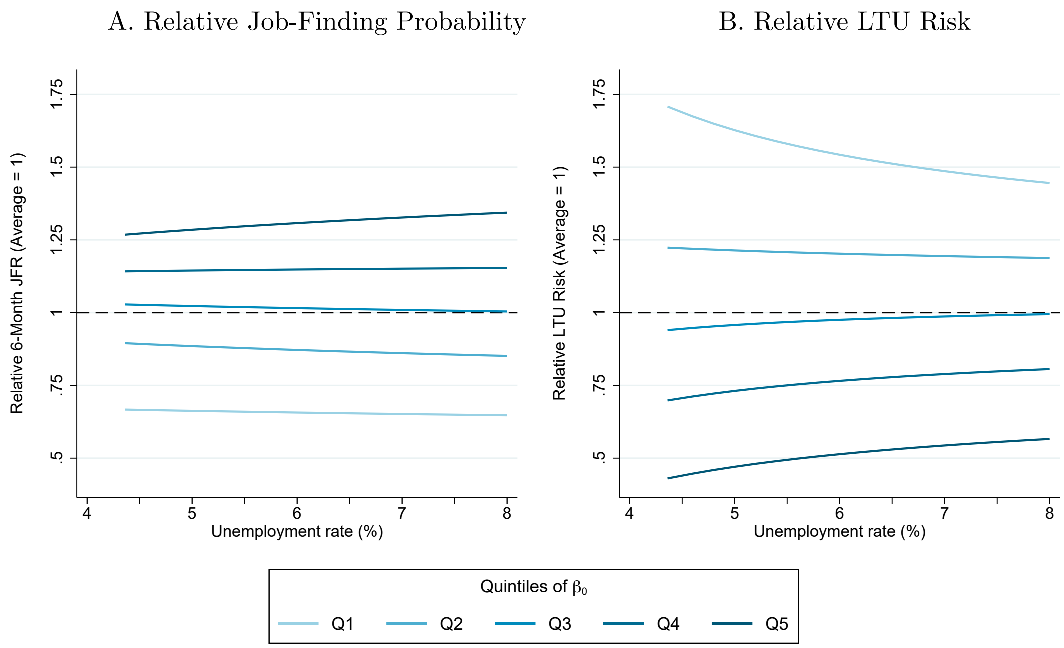
A. Histogram of $\exp(\hat{\beta}_0)$

B. Histogram of $\hat{\beta}_U$



Notes: The figure shows the distribution of the coefficients from the individual-level regressions outlined in equation 9, after applying the shrinkage in equation 8. Panel A shows the histogram of the exponential of the intercept $\exp(\hat{\beta}_0)$, while Panel B shows the histogram of the cyclicality coefficient $\hat{\beta}_U$.

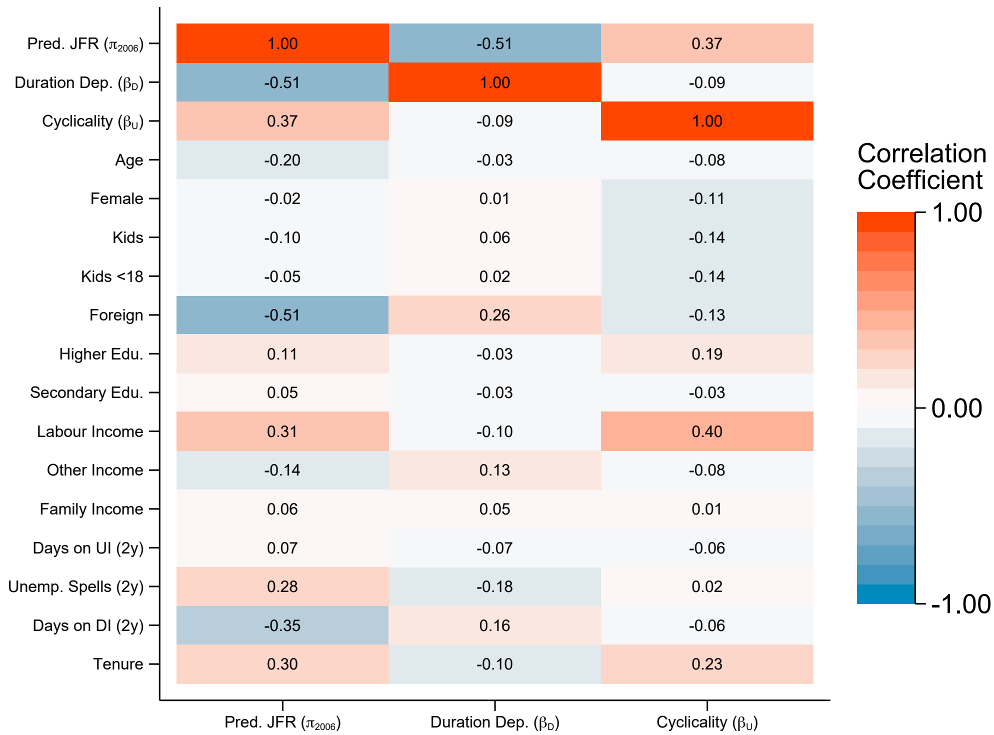
Figure A8: Heterogeneity in Individual Cyclicality: Relative To Average



Notes: Panel A shows the mean predicted individual job-finding rate for the five quintiles of the distribution of the intercept β_0, normalizing the job-finding rate to the mean in 2006 for each unemployment rate. Panel B shows the predicted change in individual LTU risk (defined as the complementary probability), relative to the mean profile, for the five quintiles of the distribution of β_0.

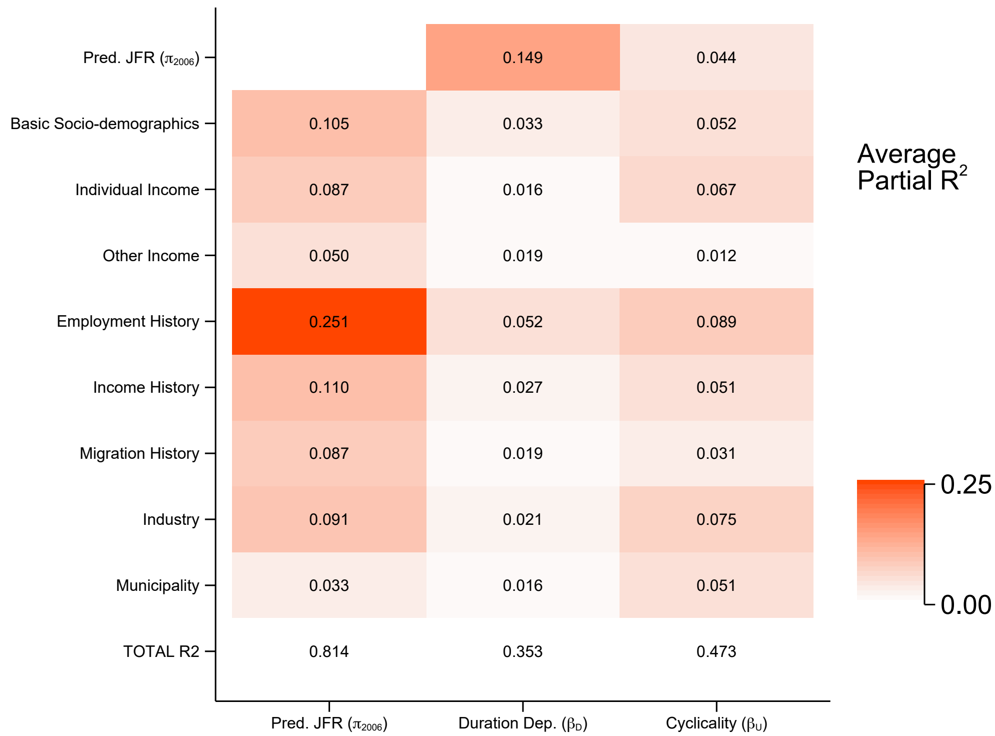
A.12 Heterogeneity and Relationship to Observables: Additional Results

Figure A9: Heterogeneity in Job Finding, Cyclicality and Duration Dependence: Correlation



Notes: This figure reports bivariate correlation coefficients between the predictions and a subset of the variables included in the baseline model. The first column shows correlations between the predicted 6-month job-finding probability at the start of the spell, from the baseline model in 2006, and the variables listed on the y-axis. Columns 2 and 3 show the coefficients for the duration dependence parameter β_D (see Section 4) and the cyclicality parameter β_U (see Section 5), respectively. All coefficients are computed on the 2006 hold-out sample.

Figure A10: Heterogeneity in Job Finding, Cyclicality and Duration Dependence: R^2



Notes: The figure reports Shapley-Owen decompositions of the total R^2 from a linear regression of the predictions on the variable groups included in the baseline model. See Grömping [2007]; Huettner and Sunder [2012] for a full description of the decomposition. The first column shows the Shapley-Owen values for the predicted 6-month job-finding probability at the start of the spell, from the baseline model in 2006. Columns 2 and 3 show the same decomposition for the duration dependence parameter β_D (see Section 4) and the cyclicality parameter β_U (see Section 5), respectively, also including the predicted job-finding probability as a separate variable group in the regressions. All values are computed on the 2006 hold-out sample.

A.13 Targeting: Descriptive Statistics and Robustness

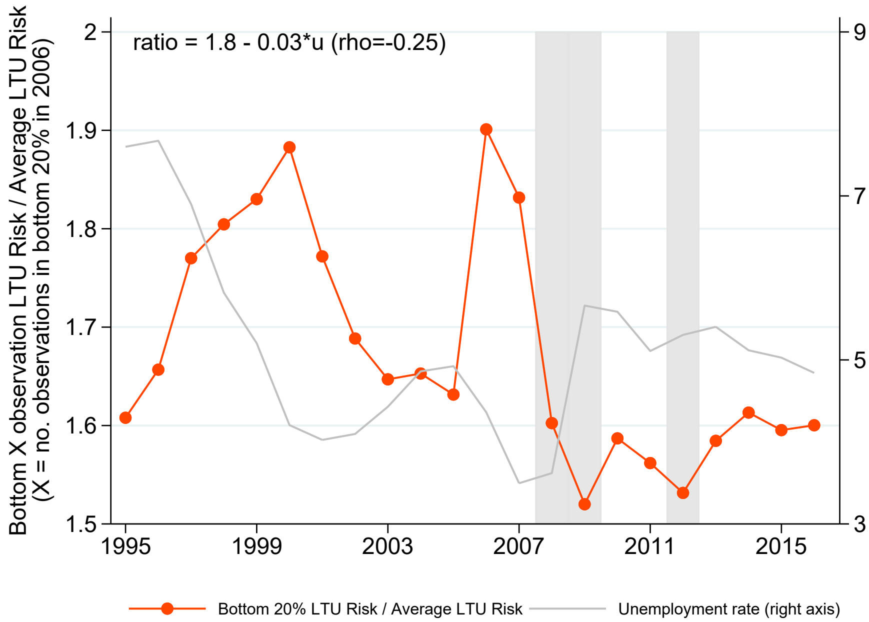
Table A11: ALMP Statistics

Category	% of sample	Prior spell dur.			Length	First 6M	Pred. JFR
		Mean	P25	P75	Mean		Mean
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(A) Vocational training	9.1%	318	89	412	121	46.5%	0.686
(B) Non-vocational training or job-search assistance	9.1%	327	48	391	91	54.1%	0.594
(A) + (B) **ALMPs (narrow)**	15.7%	273	57	343	102	54.9%	0.612
(C) Work experience (may include some training)	12.9%	421	97	588	146	36.1%	0.650
(D) Workfare	2.1%	800	190	1187	262	24.2%	0.697
(C) + (D) Work programs	13.7%	397	93	547	148	37.5%	0.654
(E) Subsidized work for the non-disabled	6.2%	601	190	798	245	24.1%	0.663
(F) Subsidized work for the disabled	1.8%	620	146	830	640	29.6%	0.583
(G) Start-up incentive	1.5%	437	117	533	204	38.5%	0.716
(E) + (F) + (G) Subsidies	9.1%	527	153	714	307	28.4%	0.648
(A) + (B) + (C) + (D) + (E) + (F) + (G) **ALMPs (broad)**	25.6%	226	56	323	139	56.1%	0.627

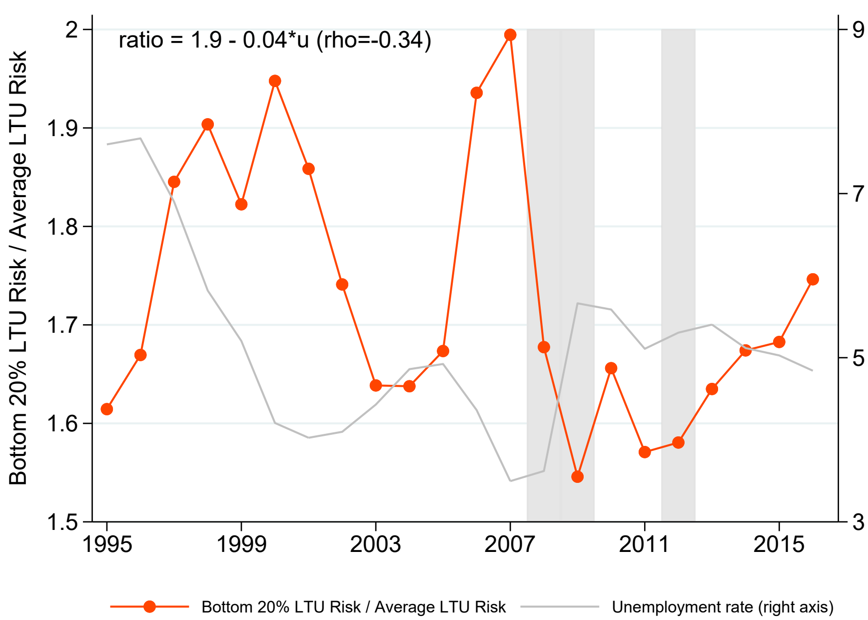
Notes: The table reports descriptive statistics about active labour market policies (ALMPs) in our full 1992-2016 sample. Column 1 shows the percentage of unemployment spells affected by a given policy. Columns 3 to 5 inform about the duration (in days) of the spell before entering the policy, while column 6 reports the length of the policy itself (in days). Column 6 shows the percentage of spells affected by the policy during the first six months of unemployment, relative to the total number of spells affected by the policy. Finally, Column 7 reports the mean predicted job-finding probability at the start of the spell for the fraction of the 2006 hold-out sample that included the policy during the first 6 months of unemployment.

Figure A11: The Value of Targeting over the Business Cycle (Robustness)

A. Fixed Number of Unemployed

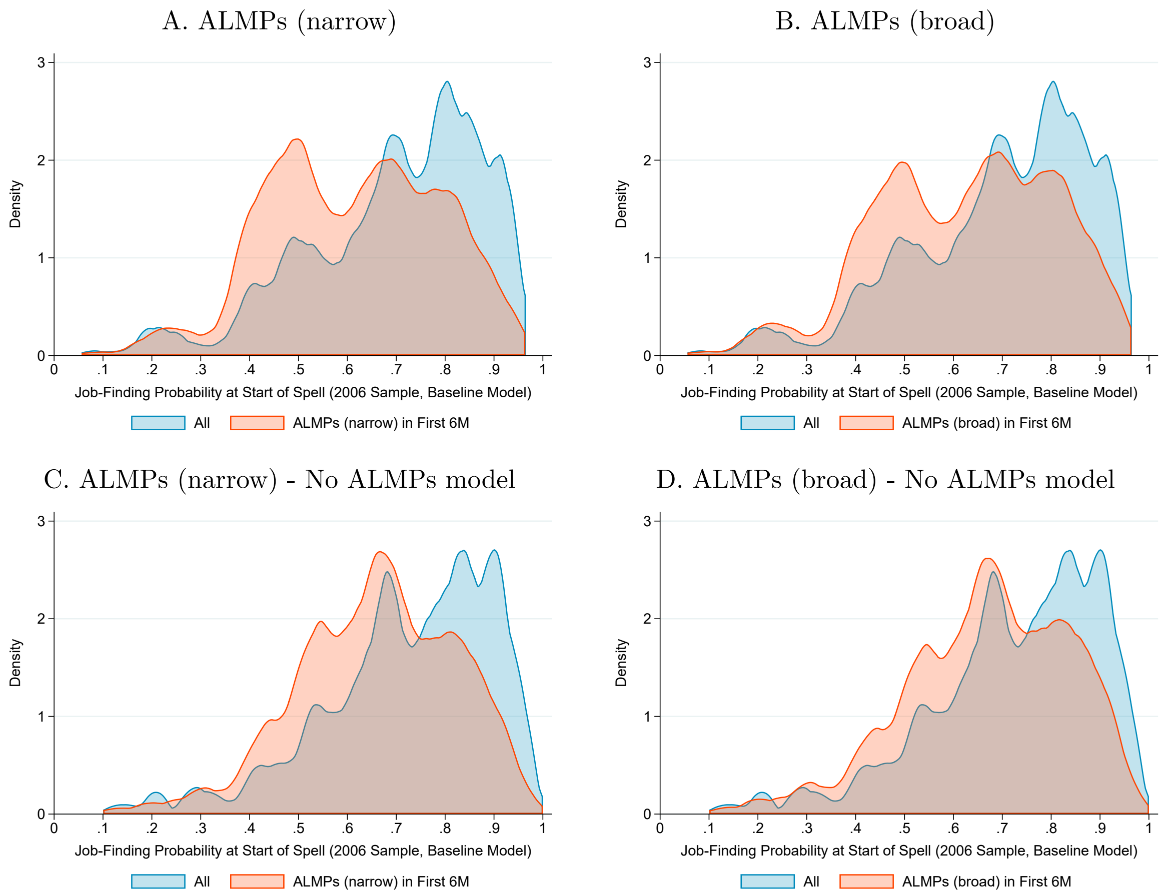


B. Excluding ALMPs



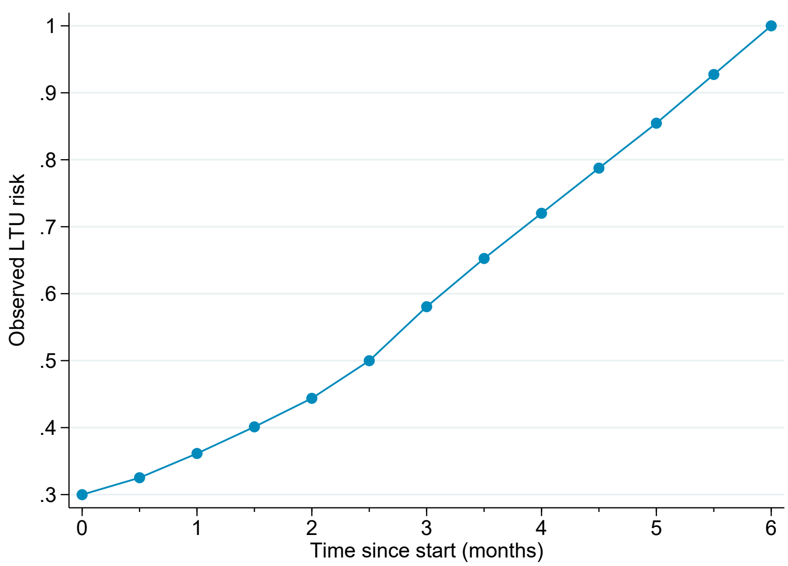
Notes: Panel A shows the value of targeting if we target a fixed number of unemployed (corresponding to the number of unemployed in the bottom 20% in 2006) instead of the bottom 20% of unemployed each year. The lower panel shows that value of targeting if we exclude all those who entered into any ALMP (in the narrow sense) during the first 6 months of their unemployment spell.

Figure A12: Targeting of Training Spells



Notes: This figure shows the distribution of predicted job-finding probabilities as in Figure 2, but separating out the spells that enter ALMPs in the narrow (panels A and C) or broad (panels B and D) sense during the first 6 months of the unemployment spell. Panels A and B show the distribution of predicted job-finding rates from the baseline model, while Panels C and D use a ML model that was trained on a sample that excludes any unemployment spell where the unemployed worker entered ALMPs (in the narrow sense) during the first 6 months of the unemployment spell.

Figure A13: Observed LTU Risk over the First 6 Months of the Spell



Notes: The graph shows observed long-term unemployment risk, defined at the probability of still being unemployed six months after the start of the unemployment spell, conditional on surviving a given amount of time since the start of unemployment. Conditional probabilities have been computed over 15-day increments for the 2006 hold-out sample.

B Proofs and Additional Propositions

B.1 Additional Propositions and Proofs

Proof of Proposition 1. We first note that

$$\begin{aligned}
cov\left(\hat{F}_i, F_i\right) &= E\left(\hat{F}_i F_i\right) - E\left(\hat{F}_i\right) E\left(F_i\right) \\
&= E\left(E\left(\hat{F}_i F_i | T_i\right)\right) - E\left(\hat{F}_i\right) E\left(T_i\right) \\
&= E\left(E\left(\hat{F}_i \times 1 | T_i\right) \Pr\left(F_i = 1 | T_i\right) + E\left(\hat{F}_i \times 0 | T_i\right) \Pr\left(F_i = 0 | T_i\right)\right) - E\left(\hat{F}_i\right) E\left(T_i\right) \\
&= E\left(E\left(\hat{F}_i T_i | T_i\right)\right) - E\left(\hat{F}_i\right) E\left(T_i\right) \\
&= E\left(\hat{F}_i T_i\right) - E\left(\hat{F}_i\right) E\left(T_i\right) \\
&= cov\left(\hat{F}_i, T_i\right).
\end{aligned}$$

Next, we use the assumption that $E\left(\varepsilon_i | X_i\right) = 0$ to show that

$$\begin{aligned}
cov\left(\hat{F}_i, T_i\right) &= cov\left(\hat{F}_i, T(X_i) + \varepsilon_i\right) \\
&= cov\left(\hat{F}_i, T(X_i)\right) + cov\left(\hat{F}_i, \varepsilon_i\right) \\
&= cov\left(\hat{F}_i, T(X_i)\right) + E\left(\hat{F}_i \varepsilon_i\right) - E\left(\hat{F}_i\right) E\left(E\left(\varepsilon_i | X_i\right)\right) \\
&= cov\left(\hat{F}_i, T(X_i)\right) + E\left(\hat{F}_i \varepsilon_i\right) \\
&= cov\left(\hat{F}_i, T(X_i)\right) + E\left(E\left(\hat{F}_i \varepsilon_i | X_i\right)\right) \\
&= cov\left(\hat{F}_i, T(X_i)\right) + E\left(E\left(\hat{F}_i | X_i\right) E\left(\varepsilon_i | X_i\right)\right) \\
&= cov\left(\hat{F}_i, T(X_i)\right).
\end{aligned}$$

Combining the fact that $cov\left(\hat{F}_i, T_i\right) = cov\left(\hat{F}_i, T(X_i)\right)$ and that $cov\left(\hat{F}_i, F_i\right) = cov\left(\hat{F}_i, T_i\right)$, we get that:

$$cov\left(\hat{F}_i, F_i\right) = cov\left(\hat{F}_i, T(X_i)\right).$$

Now we can use the Cauchy-Schwarz inequality,

$$\begin{aligned}
var\left(T(X_i)\right) var\left(\hat{F}_i\right) &\geq cov\left(\hat{F}_i, T(X_i)\right)^2 \\
&= cov\left(\hat{F}_i, F_i\right)^2.
\end{aligned}$$

Hence, we have derived the first lower bound on the variance in job-finding rates,

$$\frac{var\left(T(X_i)\right)}{var\left(F_i\right)} \geq \frac{cov\left(\hat{F}_i, F_i\right)^2}{var\left(\hat{F}_i\right) var\left(F_i\right)} = R^2\left(\hat{F}_i, F_i\right).$$

Given our assumption that $E(\varepsilon_i|X_i) = 0$, we also have that $var(T_i) \geq var(T(X_i))$ and thus

$$R^2\left(\hat{F}_i, F_i\right) \leq \frac{var(T(X_i))}{var(F_i)} \leq \frac{var(T_i)}{var(F_i)}.$$

QED.

Variance of Types with Unbiased Predictors. We can prove the following proposition:

Proposition 3. *If the predictor is unbiased, i.e.* $E(\hat{F}_i|X_i) = T(X_i)$*, then the hold-out sample covariance of the observed realization and the prediction model is an estimate of the variance in observable types as follows:*

$$cov(F_i, \hat{F}_i) = cov(T(X_i), T(X_i)). \tag{A1}$$

Proof. We take from the proof of Proposition 1 that $cov(\hat{F}_i, F_i) = cov(\hat{F}_i, T(X_i))$ and then use the fact that $E(\hat{F}_i|X_i) = T(X_i)$ as follows

$$\begin{aligned}
cov(\hat{F}_i, F_i) &= cov(\hat{F}_i, T(X_i)) \\
&= E\left(\hat{F}_i T(X_i)\right) - E\left(\hat{F}_i\right) E\left(T(X_i)\right) \\
&= E\left(E(\hat{F}_i T(X_i)|X_i)\right) - E\left(E(\hat{F}_i|X_i)\right) E\left(T(X_i)\right) \\
&= E\left(E(\hat{F}_i|X_i) T(X_i)\right) - E\left(E(\hat{F}_i|X_i)\right) E\left(T(X_i)\right) \\
&= E\left(T(X_i)^2\right) - E\left(T(X_i)\right)^2 \\
&= var(T(X_i)).
\end{aligned}$$

QED.

Proof of Proposition 2. We take from the proof of Proposition 1 that, for any δ, $cov_\delta(\hat{F}_i^\delta, F_i^\delta) = cov_\delta(\hat{F}_i^\delta, T^\delta(X_i))$ and by extension $cov_\delta(\hat{F}_i^{\delta'}, F_i^\delta) = cov_\delta(\hat{F}_i^{\delta'}, T^\delta(X_i))$. We then use the fact that $E_\delta(\hat{F}_i^{\delta'}|X_i) = T^{\delta'}(X_i)$ and proceed as follows,

$$\begin{aligned}
cov_\delta(\hat{F}_i^{\delta'}, F_i^\delta) &= cov_\delta(\hat{F}_i^{\delta'}, T^\delta(X_i)) \\
&= E_\delta\left(\hat{F}_i^{\delta'} T^\delta(X_i)\right) - E_\delta\left(\hat{F}_i^{\delta'}\right) E_\delta\left(T^\delta(X_i)\right) \\
&= E_\delta\left(E_\delta(\hat{F}_i^{\delta'} T^\delta(X_i)|X_i)\right) - E_\delta\left(E_\delta(\hat{F}_i^{\delta'}|X_i)\right) E_\delta\left(T^\delta(X_i)\right) \\
&= E_\delta\left(E_\delta(\hat{F}_i^{\delta'}|X_i) T^\delta(X_i)\right) - E_\delta\left(E_\delta(\hat{F}_i^{\delta'}|X_i)\right) E_\delta\left(T^\delta(X_i)\right) \\
&= E_\delta\left(T^{\delta'}(X_i) T^\delta(X_i)\right) - E_\delta\left(T^{\delta'}(X_i)\right) E_\delta\left(T^\delta(X_i)\right) \\
&= cov_\delta(T^{\delta'}(X_i), T^\delta(X_i)).
\end{aligned}$$

QED.

Proof of Corollary 1. First, we follow Mueller, Spinnewijn and Topa [2021] and decompose the observation decline in job finding between two adjacent periods δ and δ' as follows

$$
\begin{aligned}
E_\delta\left(T_i^\delta\right) - E_{\delta'}\left(T_i^{\delta'}\right) &= E_\delta\left[T_i^\delta - T_i^{\delta'}\right] + E_\delta\left(T_i^{\delta'}\right) - E_{\delta'}\left(T_i^{\delta'}\right) \\
&= E_\delta\left[T_i^\delta - T_i^{\delta'}\right] + \int T_i^{\delta'} dF^\delta(T_i^\delta) - \int T_i^{\delta'} dF^{\delta'}(T_i^{\delta'}) \\
&= E_\delta\left[T_i^\delta - T_i^{\delta'}\right] + \int T_i^{\delta'} dF^\delta(T_i^\delta) - \frac{\int T^{\delta'}(1-T_i^\delta) dF^\delta(T_i^\delta)}{1-E_\delta(T_i^\delta)} \\
&= E_\delta\left[T_i^\delta - T_i^{\delta'}\right] + \frac{(1-E_\delta(T_i^\delta))E_\delta(T_i^{\delta'})}{1-E_\delta(T_i^\delta)} - \frac{\int T^{\delta'}(1-T_i^\delta) dF^\delta(T_i^\delta)}{1-E_\delta(T_i^\delta)} \\
&= E_\delta\left[T_i^\delta - T_i^{\delta'}\right] + \frac{(1-E_\delta(T_i^\delta))E_\delta(T_i^{\delta'})}{1-E_\delta(T_i^\delta)} - \frac{E_\delta(T^{\delta'}) - E_\delta(T^{\delta'}T_i^\delta)}{1-E_\delta(T_i^\delta)} \\
&= E_\delta\left[T_i^\delta - T_i^{\delta'}\right] + \frac{E_\delta(T^{\delta'}T_i^\delta) - E_\delta(T^{\delta'})E_\delta(T^\delta)}{1-E_\delta(T_i^\delta)} \\
&= E_\delta\left[T_i^\delta - T_i^{\delta'}\right] + \frac{cov_\delta(T_i^\delta, T_i^{\delta'})}{1-E_\delta(T_i^\delta)},
\end{aligned}
$$

where we used the fact that $dF^{\delta'}(T_i^{\delta'}) = \frac{(1-T_i^\delta)dF^\delta(T_i^\delta)}{\int(1-T_i^\delta)dF^\delta(T_i^\delta)}$. The equation above can be re-arranged to

$$
\begin{aligned}
E_\delta\left(T_i^\delta - T_i^{\delta'}\right) &= E_\delta\left(T_i^\delta\right) - E_{\delta'}\left(T_i^{\delta'}\right) - \frac{cov_\delta(T_i^\delta, T_i^{\delta'})}{1-E_\delta(T_i^\delta)} \\
&= E_\delta\left(T_i^\delta\right) - E_{\delta'}\left(T_i^{\delta'}\right) - \frac{cov_\delta(T^\delta(X_i)+\varepsilon_i^\delta, T^{\delta'}(X_i)+\varepsilon_i^{\delta'})}{1-E_\delta(T_i^\delta)} \\
&= E_\delta\left(T_i^\delta\right) - E_{\delta'}\left(T_i^{\delta'}\right) - \frac{cov_\delta(T^\delta(X_i)+\varepsilon_i^\delta, T^{\delta'}(X_i))}{1-E_\delta(T_i^\delta)} - \frac{cov_\delta(T^\delta(X_i)+\varepsilon_i^\delta, \varepsilon_i^{\delta'})}{1-E_\delta(T_i^\delta)} \\
&= E_\delta\left(T_i^\delta\right) - E_{\delta'}\left(T_i^{\delta'}\right) - \frac{cov_\delta(T^\delta(X_i), T^{\delta'}(X_i))}{1-E_\delta(T_i^\delta)} - \frac{cov_\delta(T^\delta(X_i)+\varepsilon_i^\delta, \varepsilon_i^{\delta'})}{1-E_\delta(T_i^\delta)} \\
&= E_\delta\left(T_i^\delta\right) - E_{\delta'}\left(T_i^{\delta'}\right) - \frac{cov_\delta(T^\delta(X_i), T^{\delta'}(X_i))}{1-E_\delta(T_i^\delta)} - \frac{cov_\delta(\varepsilon_i^\delta, \varepsilon_i^{\delta'})}{1-E_\delta(T_i^\delta)} - \frac{cov_\delta(T^\delta(X_i), \varepsilon_i^{\delta'})}{1-E_\delta(T_i^\delta)}
\end{aligned}
$$

The last covariance term $cov_\delta(T^\delta(X_i), \varepsilon_i^{\delta'})$ equals 0 by the assumption that any unobserved heterogeneity is orthogonal to the obervables, $E(\varepsilon_i|X_i) = 0$. This indeed implies that unobserved heterogeneity and observable heterogeneity across adjacent periods are orthogonal for a given set of individuals.

If $cov_\delta(\varepsilon_i^\delta, \varepsilon_i^{\delta'}) \geq 0$, then the equation above

$$
E_\delta\left(T_i^\delta - T_i^{\delta'}\right) \leq E_\delta\left(T_i^\delta\right) - E_{\delta'}\left(T_i^{\delta'}\right) - \frac{cov_\delta(T^\delta(X_i), T^{\delta'}(X_i))}{1-E_\delta(T_i^\delta)}.
$$

Next, we use the fact that, in the hold-out sample, $E_\delta\left(F_i^\delta\right) = E_\delta\left(T_i^\delta\right)$, $E_{\delta'}\left(F_i^{\delta'}\right) = E_{\delta'}\left(T_i^{\delta'}\right)$, and $cov_\delta(\hat{F}_i^{\delta'}, F_i^\delta) = cov_\delta(T^{\delta'}(X_i), T^\delta(X_i))$ (from Proposition 2) and thus get

$$E_\delta\left(T_i^\delta - T_i^{\delta'}\right) \leq E_\delta\left(F_i^\delta\right) - E_{\delta'}\left(F_i^{\delta'}\right) - \frac{cov_\delta\left(F_i^\delta, \hat{F}_i^{\delta'}\right)}{1 - E_\delta\left(F_i^\delta\right)}.$$

QED.

Identification with Data on Multiple Unemployment Spells Per Person. We prove the following proposition:

Proposition 4. *The covariance of actual job finding across two unemployment spells of the same individual at time δ and δ' is equal to the covariance in the underlying probabilities across the two unemployment spells:*

$$cov_{\delta,\delta'}\left(F_i^\delta, F_i^{\delta'}\right) = cov_{\delta,\delta'}\left(T_i^\delta, T_i^{\delta'}\right).$$

Proof.

$$\begin{aligned}
cov_{\delta,\delta'}\left(F_i^\delta, F_i^{\delta'}\right) &= E_{\delta,\delta'}\left[F_i^\delta F_i^{\delta'}\right] - E_{\delta,\delta'}\left(F_i^\delta\right) E_{\delta,\delta'}\left(F_i^{\delta'}\right) \\
&= E_{\delta,\delta'}\left[E_{\delta,\delta'}\left(F_i^\delta F_i^{\delta'} | T_i^\delta, T_i^{\delta'}\right)\right] - E_{\delta,\delta'}\left(T_i^\delta\right) E_{\delta,\delta'}\left(T_i^{\delta'}\right) \\
&= E_{\delta,\delta'}\Big[E_{\delta,\delta'}\left(1 \times 1 | T_i^\delta, T_i^{\delta'}\right) \Pr\left(F_i^\delta = 1 \ \& \ F_i^{\delta'} = 1 | T_i^\delta, T_i^{\delta'}\right) \\
&\quad + E\left(1 \times 0 + 0 \times 1 + 0 \times 0 | T_i^\delta, T_i^{\delta'}\right)\left(1 - \Pr\left(F_i^\delta = 1 \ \& \ F_i^{\delta'} = 1 | T_i^\delta, T_i^{\delta'}\right)\right)\Big] \\
&\quad - E_{\delta,\delta'}\left(T_i^\delta\right) E_{\delta,\delta'}\left(T_i^{\delta'}\right) \\
&= E_{\delta,\delta'}\left[\Pr\left(F_i^\delta = 1 \ \& \ F_i^{\delta'} = 1 | T_i^\delta, T_i^{\delta'}\right)\right] - E_{\delta,\delta'}\left(T_i^\delta\right) E_{\delta,\delta'}\left(T_i^{\delta'}\right) \\
&= E_{\delta,\delta'}\left[T_i^\delta T_i^{\delta'}\right] - E_{\delta,\delta'}\left(T_i^\delta\right) E_{\delta,\delta'}\left(T_i^{\delta'}\right) \\
&= cov_{\delta,\delta'}\left(T_i^\delta, T_i^{\delta'}\right).
\end{aligned}$$

QED.

Corollary 2. *In the stylized model in Section 2, where $T_i^\delta = T_i + h(\delta) + \nu_i^\delta$, the covariance of actual job finding across two unemployment spells of the same individual at time δ and δ' is equal to the variance in the fixed heterogeneity across the two spells, T_i, but a lower bound for the total heterogeneity, i.e.:*

$$cov_{\delta,\delta'}\left(F_i^\delta, F_i^{\delta'}\right) = var_{\delta,\delta'}\left(T_i\right) \leq var_{\delta,\delta'}(T_i^\delta) = var_{\delta,\delta'}(T_i^{\delta'}).$$

Proof. We take from the proof of Proposition 4 above the fact that $cov_{\delta,\delta'}\left(F_i^\delta, F_i^{\delta'}\right) = cov_{\delta,\delta'}\left(T_i^\delta, T_i^{\delta'}\right)$. Then:

$$
\begin{aligned}
cov_{\delta,\delta'}\left(F_i^{\delta}, F_i^{\delta'}\right) &= cov_{\delta,\delta'}\left(T_i^{\delta}, T_i^{\delta'}\right) \\
&= cov_{\delta,\delta'}\left(T_i + h\left(\delta\right) + \nu_i^{\delta}, T_i + h\left(\delta'\right) + \nu_i^{\delta'}\right) \\
&= cov_{\delta,\delta'}\left(T_i, T_i\right) \\
&= var_{\delta,\delta'}\left(T_i\right).
\end{aligned}
$$

It is obvious that $var_{\delta,\delta'}\left(T_i\right) \leq var_{\delta,\delta'}\left(T_i\right) + var_{\delta,\delta'}\left(\nu_i^{\delta}\right) = var_{\delta,\delta'}(T_i^{\delta})$, and that $var_{\delta,\delta'}\left(T_i\right) \leq var_{\delta,\delta'}\left(T_i\right) + var_{\delta,\delta'}\left(\nu_i^{\delta'}\right) = var_{\delta,\delta'}(T_i^{\delta'})$. QED.

C Additional Details on Prediction Model

In this Appendix, we describe the binary prediction algorithm that we use to obtain the job-finding probabilities, and report its accuracy across different subgroups.

C.1 Prediction Algorithm

The algorithm we use to predict the probability that an individual finds a job in the next 6 months is a standard machine learning method for binary classification, an ensemble learner that consists in our case of a random forest model, gradient boosted regression trees and LASSO model. To avoid overfitting, we train and calibrate the prediction algorithm on a training sample, for which we use 51.4% of the overall sample. We then use this trained prediction algorithm to obtain predictions for a hold-out sample, which consists of the remaining unemployment spells. All the analyses and statistics in the paper are developed use only this hold-out sample.

The prediction method we use follows four steps, which closely resemble the steps used in Einav et al. [2018]. First, we follow standard practice in machine learning by tuning key parameters that govern the prediction models by 3-fold cross-validation. Second, we train the three resulting prediction models separately. Third, we combine the three obtained predictions into one using a linear combination that we calibrate in the data. Finally, we calibrate the resulting final ensemble predictions using a linear spline. We describe each of the four steps in more detail here.

Parameter Tuning As the three machine learning models that we use have parameters that are at the discretion of the researcher, we follow standard practice and tune these parameters using 3-fold cross validation. More specifically, we tune the following parameters using 10 percent of the sample: minimal node size (mid.node.size), number of variables used at each node (mtry) for the random forest model, learning rate (eta) for the boosted regression trees, and the shrinkage parameter (lambda) for the LASSO.[1] For each of these parameters, we optimize among 5 to 7 alternatives. We tune these parameters using 3-fold cross validation, where we are optimizing the area under the receiver operating characteristic curve (AUC).[2] Thus, for each of the parameter values we want to test, the model is trained on 2 folds (subsets of the training sample), and then the performance is measured in the 3rd fold. The parameter values for which the AUC in the 2006 hold-out sample is highest for each prediction algorithm are: mtry = 50, min.node.size = 12, eta = 0.5, lambda = 0.01.

Estimating the Models Using these tuned parameter values, all models are estimated using 30% of the sample.

Obtaining Ensemble Predictor We combine the predictions from the random forest, gradient boosting regression trees, and LASSO into one ensemble prediction. Following Einav et al. [2018], we construct the ensemble prediction to be the linear combination $p_{ensemble} = \hat{\beta}_{rf}\hat{p}_{rf} + \hat{\beta}_{gb}\hat{p}_{gb} + \hat{\beta}_{lasso}\hat{p}_{lasso}$, where $\hat{p}_x$ is the prediction from algorithm x and $\hat{\beta}_x$ is the associated weight.

[1] We use the package CARET in R that provides a standardized way to tune parameters. The prediction models we use are RANGER (random forest), XGBLINEAR (boosted regression trees), and GLMNET (LASSO).

[2] This is a common metric used in the machine learning literature to measure the performance of a prediction model.

We obtain estimates for the weights from a constrained linear regression (with no constant and the weights summing to one) of the dummy for job finding on the three individual predicted probabilities. For this step, we use 6% of the sample. We find associated weights for the baseline model in 2006 that are $\hat{\beta}_{rf} = 1.03$, $\hat{\beta}_{gb} = 0.04$ and $\hat{\beta}_{lasso} = -0.07$. The gradient-boosted regression trees seems to perform less well than the other prediction models.

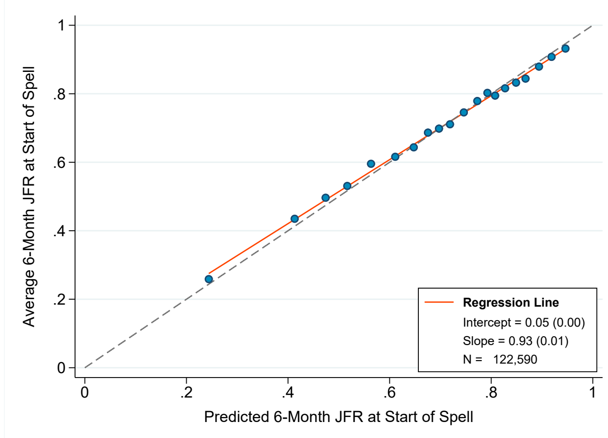
Calibrating Probabilities Finally, the raw probability predictions we get from the ensemble step are calibrated to the actual observed probabilities by estimating a linear spline. This calibration is done using 5.4% of the sample, again not used in any of the previous steps. 250 equal sized bins are created based on the ranked predicted probability. In every bin the mean probability is calibrated to the observed mean probability for these observations. The piece-wise linear spline that follows from linearly interpolating all intermediate points serves as the last step in the prediction mechanism.

C.2 Additional Discussion of Prediction Model

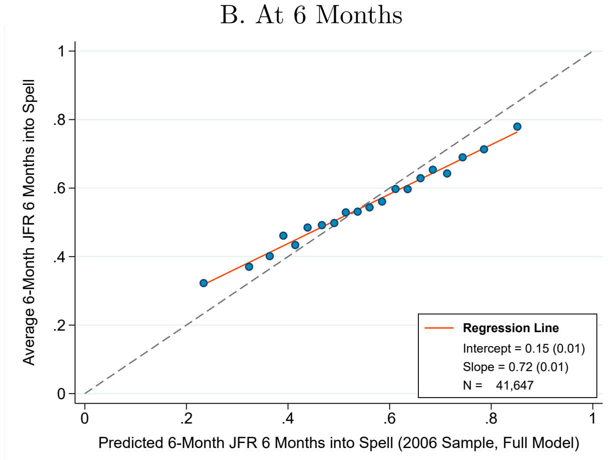
While Panel A of Figure 1 shows a calibration plot for the entire sample, Figures A15 to A19 show a calibration plot for certain subgroups of the sample. If we construct 144 groups by income decile, gender, citizenship, days on DI and days on UI, we see from Panel A, B and C of Figure A20 that average predicted probabilities within the groups remain well calibrated. This makes us comfortable that the observed differences in predicted long-term unemployment risks across different groups are not due to differential prediction accuracy of our ensemble predictor.

Figure A14: Comparing Predictions to Outcomes

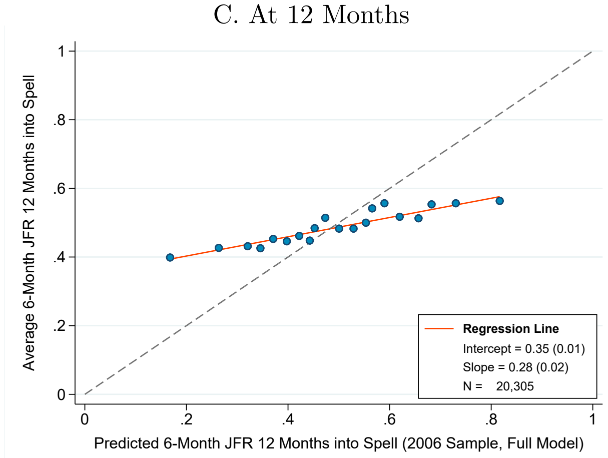
A. At 0 Months



B. At 6 Months

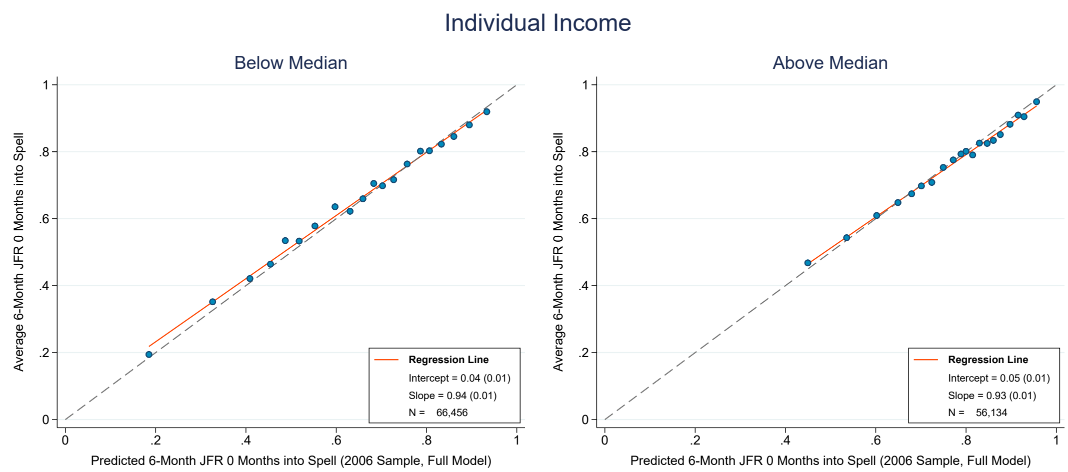


C. At 12 Months



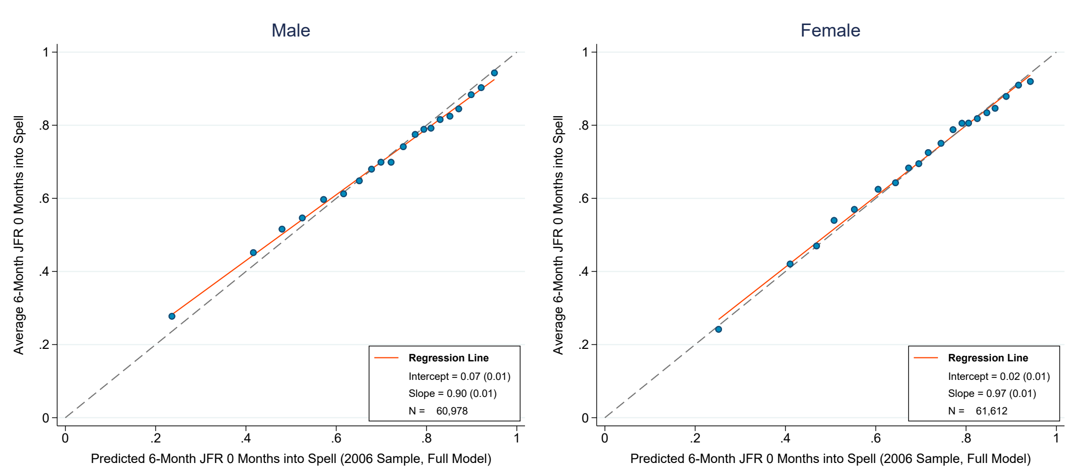
Notes: The figure presents binned scatter plots of observed and predicted job-finding rates, as in Figure 1, for various unemployment durations. Panel A simply reproduces Panel A of 1, while Panels B and C show the predictions 6 and 12 months into the unemployment spell, respectively. All results correspond to the 2006 hold-out sample.

Figure A15: Comparing Predictions to Outcomes: by Income



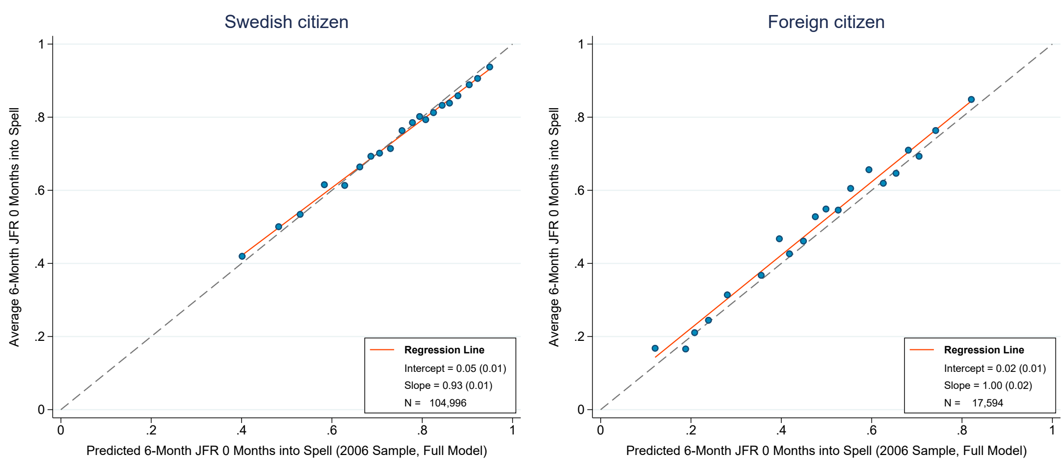
Notes: The figure presents binned scatter plots of observed and predicted job-finding rates at the start of the spell, as in Figure 1, but splitting the 2006 hold-out sample into two bins by individual labour income.

Figure A16: Comparing Predictions to Outcomes: by Gender



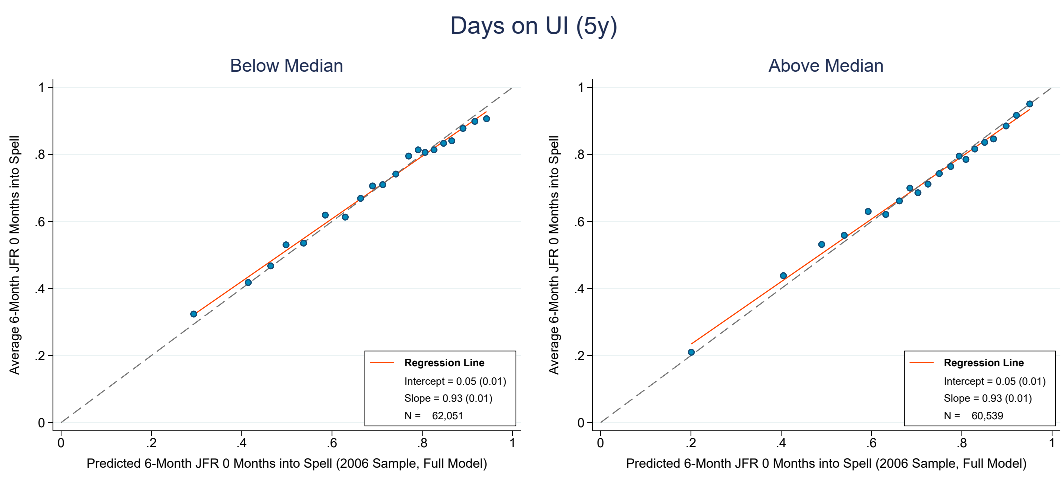
Notes: The figure presents binned scatter plots of observed and predicted job-finding rates at the start of the spell, as in Figure 1, but splitting the 2006 hold-out sample into two bins by gender.

Figure A17: Comparing Predictions to Outcomes: by Citizenship



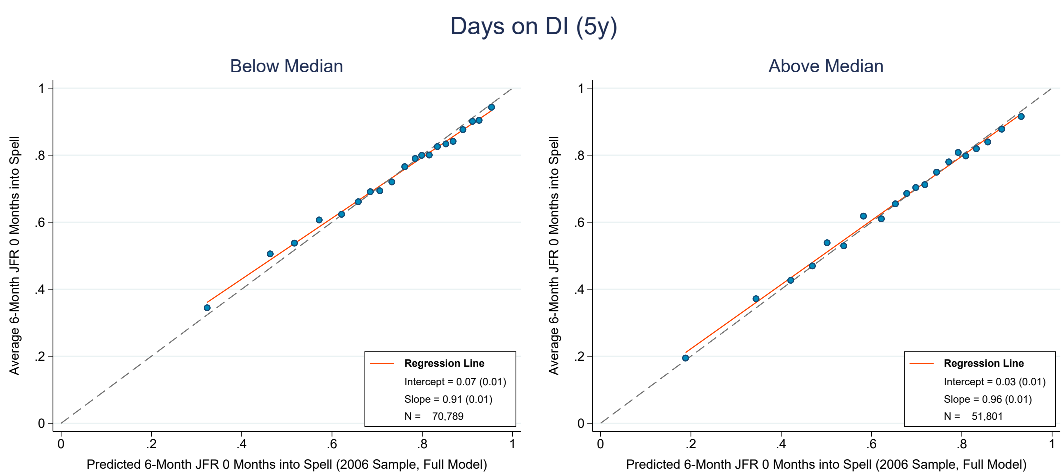
Notes: The figure presents binned scatter plots of observed and predicted job-finding rates at the start of the spell, as in Figure 1, but splitting the 2006 hold-out sample into two bins by citizenship.

Figure A18: Comparing Predictions to Outcomes: by Days on UI



Notes: The figure presents binned scatter plots of observed and predicted job-finding rates at the start of the spell, as in Figure 1, but splitting the 2006 hold-out sample into two bins by days on UI during the 5 years preceding the unemployment spell.

Figure A19: Comparing Predictions to Outcomes: by Days on DI

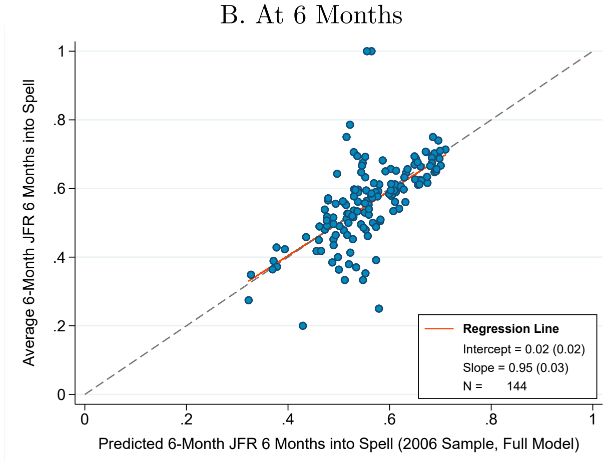


Notes: The figure presents binned scatter plots of observed and predicted job-finding rates at the start of the spell, as in Figure 1, but splitting the 2006 hold-out sample into two bins by days on DI during the 5 years preceding the unemployment spell.

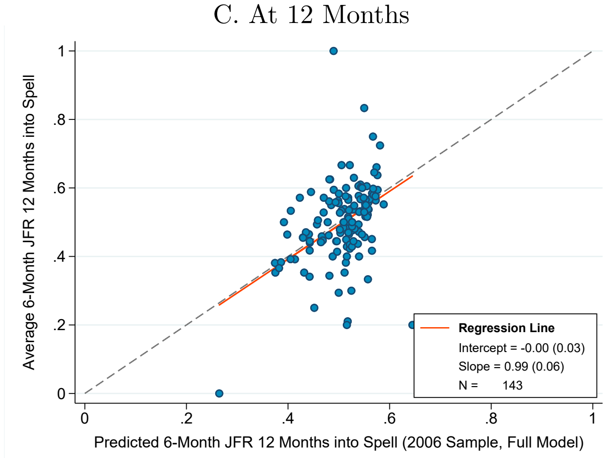
Figure A20: Comparing Predictions to Outcomes: by 144 groups

A. At 0 Months

B. At 6 Months



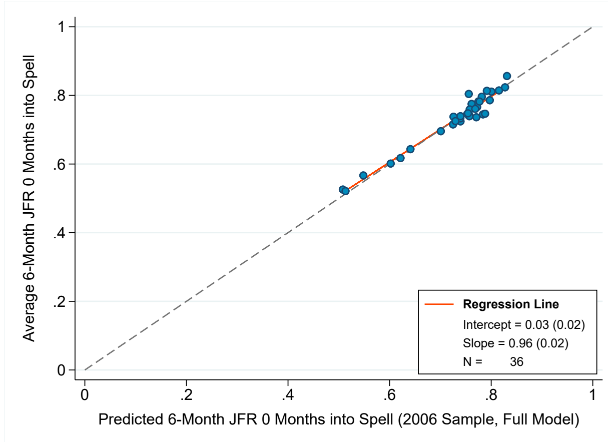
C. At 12 Months



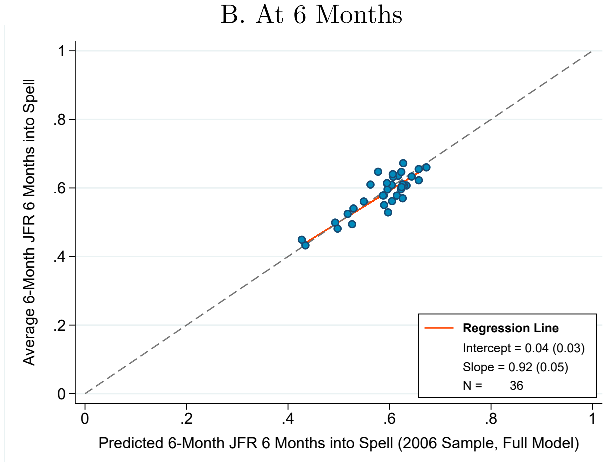
Notes: The figure presents binned scatter plots of observed and predicted job-finding rates for various unemployment durations. Here we construct 144 bins by deciles of labour income, gender, citizenship, days on UI and days on DI, and we report average observed and predicted job-finding rates for each bin. The regression output corresponds to a regression of bin averages. Panel A uses the baseline predictions at the start of the spell, while Panels B and C show the predictions 6 and 12 months into the unemployment spell, respectively. All results correspond to the 2006 hold-out sample.

Figure A21: Comparing Predictions to Outcomes: by 36 groups

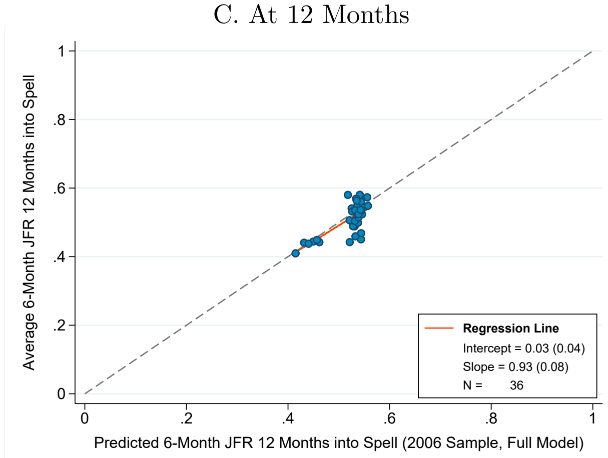
A. At 0 Months



B. At 6 Months



C. At 12 Months



Notes: The figure presents binned scatter plots of observed and predicted job-finding rates for various unemployment durations. Here we construct 36 bins by deciles of labour income, gender and citizenship, and we report average observed and predicted job-finding rates for each bin. The regression output corresponds to a regression of bin averages. Panel A uses the baseline predictions at the start of the spell, while Panels B and C show the predictions 6 and 12 months into the unemployment spell, respectively. All results correspond to the 2006 hold-out sample.