A Structural Analysis of Mental Health and Labor Market Trajectories*

Grégory Jolivet † Fabien Postel-Vinay‡

University of Bristol UCL and IFS

July 29, 2020

Abstract

We analyze the joint life-cycle dynamics of labor market and mental health outcomes. We allow for two-way interactions between work and mental health. We model selection into jobs on a labor market with search frictions, accounting for the level of exposure to stress in each job using data on occupational health contents. We estimate our model on British data from *Understanding Society* combined with information from O*NET. We estimate the impact of job characteristics on health dynamics and of the effects of health and job stress contents on career choices. We use our model to quantify the effects of job loss or health shocks that propagate over the life cycle through both health and work channels. We also estimate the (large) values workers attach to health, employment or non-stressful jobs. Lastly, we investigate the consequences on health, employment and inequality of trend changes in the distribution of job health contents.

JEL classification: I12, I14, J62, J64.

*We would like to thank Stéphane Bonhomme, Donna Gilleskie, Eric French, Amanda Michaud, Joe Mullins, Renata Narita, Andrew Oswald, Helen Simpson, Anthony Terriau, Marcos Vera-Hernandez and Frank Windmeijer for very useful comments at various stages in the production of this paper. We are also grateful to seminar and workshop participants at Aarhus (Sandbjerg Manor), Aix-Marseille (SaM), CREST, Edinburgh, Essex, Georgetown, Louvain-La-Neuve, Lyon, Nottingham, Royal Holloway, UCL, and Warwick for helpful feedback. The usual disclaimer applies.

†Address: Department of Economics, University of Bristol, Priory Road Complex, Bristol BS8 1TU, UK. E-mail: `gregory.jolivet@bristol.ac.uk`.

‡Address: Department of Economics, University College London, Drayton House, 30 Gordon Street, London WC1H 0AX, UK. E-mail: `f.postel-vinay@ucl.ac.uk`.

1 Introduction

Tackling the personal and economic costs of mental ill health is making its way up the list of priorities of social scientists and policy makers alike.[1] As a preliminary step toward the design of efficient policies addressing mental ill health, this paper aims to contribute to a better understanding of the link between individual labor market trajectories (employment, wages, occupations) and mental health outcomes.

The interaction between work and health is a two-way one: ill health affects labor supply, and conversely working in a stressful job likely affects future mental health. This two-way interaction has been acknowledged in the health economics literature (see Currie and Madrian, 1999 for a survey) and has recently been factored into structural models of health and labor supply (discussed in Section 2).

We add to that literature by modeling that two-way interaction jointly with key drivers of labor market careers such as the cost of labor supply, search frictions and mobility on the job ladder, and by fully estimating our model using longitudinal worker data linked to new measures of the health contents of occupations. Moreover, we focus on mental health. This allows us to inform recent debates on the mental health consequences of recessions, the increase in work-related stress, or the changing nature of work through polarization (as documented by Autor and Dorn, 2013 and, in the health dimension, by Kaplan and Schulhofer-Wohl, 2018). Using our model for treatment evaluation, we speak to important questions such as the impact of job loss on mental health, the effect of a mental health shock on careers, or the cost of working in a stressful job.

The main ingredients of our model are as follows. Jobs are characterized by a wage, working hours, and a "health content" — i.e. a measure of how much of a strain the job puts on workers' mental health. Individuals' mental health evolves stochastically over time depending on their current job characteristics. Individuals self-select in and out of employment as well as across jobs subject to search frictions. Job characteristics affect individual utility (through the wage or through the disutility of working) but also future health. Conversely, health affects the decision to work through the disutility of labor as working in a stressful job is more difficult when in poor health. Hence the feedback loop between health and work is built into our model.

[1] The World Health Organization has launched a 5-year special initiative for mental health in 2019 (WHO, 2019). The OECD calls for a stronger policy response to mental ill-health, outlining its economic costs (OECD, 2015). See also the recent article by Layard (2017).

We conduct our empirical analysis on a sample of men aged 25 to 55 on the UK labor market (we later discuss why we focus on this country, gender and age group). We collate information from two different data sets: individual-level health and labor market data from the UK Household Longitudinal Survey (UKHLS) and data on the health contents of occupations from the O*NET project. UKHLS is a large panel data set with detailed information on individual labor market status, earnings, and labor market transitions. Crucially, UKHLS also conveys information on individual mental health. That information is based on answers to the widely used 12-item Short-Form (SF-12) health questionnaire, which are conveniently aggregated into summary measures of mental (or physical) health. As for occupation data, the O*NET project contains a large set of descriptors conveying information about the mental health contents of occupations. We run a principal component analysis on specific descriptors to construct an indicator of the stress contents of occupations in O*NET. That indicator can then be linked to UKHLS data using standardized occupation codes.

We develop a multi-stage procedure to estimate our structural model. First, the discrete distribution of individual heterogeneity driving labor and health outcomes is estimated building on the recent clustering approach of Bonhomme et al. (2019a,b). The second stage pertains to health dynamics and maximizes the likelihood of health transitions conditionally on individual and job characteristics. In the third stage we estimate the remaining parameters relating to wage offers, search frictions and preferences, by matching simulated moments to their empirical counterparts. Importantly, we prove identification of the model parameters, conditional on the distribution of individual heterogeneity.[2] Even though the model allows us to follow individual trajectories along two dimensions (work and health), the computational cost of our multi-stage estimation procedure remains very manageable. Hence our approach would easily lend itself to useful extensions for further policy analysis. These will be discussed in the conclusion.

Our first set of results follows directly from the estimation of the model. We find that being older, earning a higher wage or working in a less stressful job have a positive causal impact on future mental health. These effects are small in magnitude for year-on-year health changes, but can lead to rather different health trajectories when accumulated over several years. We further estimate individual preference parameters showing that the job stress content, poor

[2] In the absence of a formal proof of identification for the distribution of individual heterogeneity, we perform a (tough) numerical test of internal consistency by comparing the clusters obtained on simulated trajectories from the estimated model with those obtained from the data. The very high overlap between these two distributions leads us to conclude that our approach to estimating individual heterogeneity is, at least, internally consistent.

mental health and age all increase the cost of labor supply.

We use the estimated structure of our model to study three types of shocks: a job loss shock, a health shock and an increase in job stress content. All treatments are administered at age 30 and their effects are allowed to vary with individual heterogeneity, health or labor market status. We find that, while the health effect of a one-time mental health shock dissipates after 3 to 4 years on average, the effects of such a shock on employment and income can last longer. Moreover, a job loss shock has large and lasting negative effects not only on labor market outcomes, consistently with the recent labor economics literature (see Davis and Von Wachter, 2011, or Jarosch, 2015), but also on health.

The two-way interaction between health and work is key to the propagation of both health and labor market shocks. For instance, we show that, while an adverse health shock has little impact on the labor supply decisions of workers in less stressful jobs, it can induce workers employed in more difficult jobs to quit and enter the health and career dynamics of an unemployed worker, making them more likely to stay in poor health and finding certain jobs too costly to accept. Our treatment analysis allows for, and quantifies, this type of interactions.

Alongside these treatment evaluations, we can also estimate the (expected) utility cost of being in poor health, of becoming unemployed, or of an increase in job stress content. Consistently with the recent macro-labor literature, we find that the cost of job loss is very high. The value loss due to a severe health shock is also substantial: for workers in a median-wage and average-stress job, it is about a third of the cost of losing their job. Moreover, we find that workers in stressful jobs are willing to give up more than 10% of their wage to avoid an adverse health shock. Job health content is also important to workers: 30-year old workers are willing to forego 15% to 22% of their wage to go from a high-stress to a medium-stress job.

The labor market polarization literature (see e,g, Autor and Dorn, 2013) has documented structural changes in the skill contents of occupations, away from routine tasks and toward high-skill jobs with strong cognitive content. A recent paper by Kaplan and Schulhofer-Wohl (2018) explores the consequences of these changes on workers' self-perceived stress and overall job satisfaction. In this paper, we investigate the health consequences of structural labor market changes by simulating workers' health outcomes for different counterfactual job offer distributions, ranging from high proportions of low-stress jobs to a majority of high-stress jobs. We show that workers starting their careers on labor markets offering more stressful jobs have a higher risk of being in poor mental health mid career, although the magnitude of that effect is quite small. We conclude that recent trends in the distribution of job health contents are, in themselves, unlikely to cause a substantial deterioration of mental health in the working population. However, we also show that a higher prevalence of stressful jobs increases lifetime inequality mainly because of its differential impact across the health distribution, as the rationing of low-stress jobs hurts workers in ill health and benefits workers in good health.

While our approach provides a detailed description of health and labor market trajectories, it has certain limitations. First, health affects an individual's flow utility only through the disutility of labor: we cannot separately identify the effects of health on the flow utility of employed and unemployed workers. We can thus quantify the extent to which poor health makes it more costly to work or affects a worker's labor market decisions but we do not capture the utility cost of being in poor health while not working. Our predictions on the welfare costs of poor health should thus be understood as lower bounds. Second, we perform our analysis on UK data where the National Health Service offers universal coverage, and abstract from any health insurance choice issues (see Currie and Madrian, 1999, or Gruber and Madrian, 2004). Third, the effect of job characteristics on health are modeled in a causal reduced-form fashion: we do not delve into individual decisions to invest in medical treatment or healthier living (Gilleskie, 1998 has a structural analysis of medical treatment decisions). Lastly, our analysis focuses on men aged 25 to 55. We discuss important extensions of our approach to older workers and to women in the conclusion as they are next on our research agenda.

The paper is organized as follows. Section 2 discusses the literature. Section 3 presents the data, including our new measure of the mental health content of jobs. Section 4 presents our model and Section 5 details our multi-step identification and estimation procedure. Estimation results and model fit are analyzed in Section 6. Sections 7 and 8 present various counterfactual analyses. Section 9 concludes and discusses potential extensions. Details on the data, estimation procedure and identification proofs are gathered in Appendices A-D.

2 Some related literature

This paper is an attempt to bridge the gap between the structural labor and health literatures by explicitly modeling the two-way interaction between work and health in a quantitative dynamic model with labor supply, labor market frictions and on-the-job search. Structural models have been used in many ambitious empirical evaluations of the effects of labor market shocks (job loss, recessions, changes in task contents) on inequality, individual trajectories and welfare. Recent contributions include Lise and Robin (2017) who analyze aggregate shocks, Jarosch (2015) and Burdett et al. (2018) who study the effect of job loss or Low et al. (2010) who look at productivity and job destruction shocks. The causal effects of those shocks on health have been documented by both the economic (Sullivan and Von Wachter, 2009; Davis and Von Wachter, 2011) and medical (Fryers et al., 2003) literatures, but attempts at fully incorporating them into structural models of individual careers are still relatively scarce.

A branch of the structural literature investigates the effect of health on retirement decisions (French, 2005 and French and Jones, 2011) or on long-term labor earnings and welfare (de Nardi et al., 2017). Apart from very few exceptions, that literature tends to stay silent on the feedback effect of labor market outcomes on health. Exceptions include Jacobs and Piyapromdee (2016), who seek to explain the reverse retirement behavior. To that end, they extend French and Jones (2011) to include a burnout-recovery process, whereby workers' stock of accumulated stress evolves depending on labor force status (full-time employed, part-time employed, or retired) and can impact labor supply. More recently, Salvati (2020) develops a rich structural life-cycle model of (physical) health, labor supply, and savings of women at older ages, explicitly allowing for a two-way interaction between health and labor supply and exploiting exogenous changes in women's state pension age (in the UK) for identification. Aside from the difference in focus, both these papers differ from ours in that neither features labor market frictions or heterogeneity across occupations in job health contents.

The two-way interaction between work and health has been considered in structural models of health and medical treatment decisions (e.g. Papageorge, 2016). These models are geared to the description of health investments rather than career dynamics: due to their design and purpose, they do not include features such as search frictions or job ladder transitions that not only drive labor market turnover and inequality but are also likely to play a role in the propagation of health shocks.

Finally, a few recent structural contributions on health shocks and labor supply allow for a reciprocal link between health and work. Capatina et al. (2020) propose a sophisticated life-cycle model to analyze the contribution of health shocks to earnings inequality and labor market outcomes, while allowing for human capital accumulation and a rich description of individual health dynamics. Their model is calibrated on US data, using a compound measure of health (rather than mental health). In a similar vein, Harris (2019) and Tran (2017) study, respectively, the effect of body weight and mental health on labor market choices while reciprocally allowing occupations to impact weight and mental health. Lastly, Michaud and Wiczer (2018a) calibrate a dynamic labor supply model to study the effect of macroeconomic shocks on workers who can choose between occupations with different disability risks.

Our paper differs from these important contributions in several key dimensions.[3] First, we allow for heterogeneity in the health contents of jobs and introduce an original occupation-level measure of exposure to stress. Our model thus explicitly features a disamenity (stress on the job) that affects both workers' labor supply and their mental health. Second, our model includes labor market frictions and a job ladder mechanism driving worker turnover. The combination of frictions and heterogeneity in job stress contents is important to evaluate the consequences of health and labor market shocks. For example, workers who suffer an adverse health shock causing them to quit their job then have to resume climbing the job ladder from the bottom and, because of search frictions, may have to accept more stressful or less-well paid jobs before going back to a job that is less damaging for their health. We believe that our model is the first to capture this propagation mechanism of health shocks that operates through labor market frictions. We also believe that, thanks to our explicit modeling of the stress contents of jobs, our analysis is the first to evaluate jointly the health and labor market consequences of structural changes documented by the job polarization literature.[4]

3 Data and descriptive statistics

3.1 Labor force and health data: UKHLS

Our main data source is the UK Household Longitudinal Survey (UKHLS), a.k.a. *Understanding Society*. UKHLS is a yearly household panel started in 2009 as the follow up to the British Household Panel Survey (BHPS).[5] It contains yearly observations of respondents' labor market status, wage and occupation (if any), age as well as high-frequency observations of any labor

[3]While this section focuses on what our paper has that existing papers do not, we obviously acknowledge that those papers have features that are absent from ours, due to their different focus. For instance, Michaud and Wiczer (2018a) model disability insurance and retirement decisions (see also Michaud and Wiczer, 2018b, for a related paper). Capatina et al. (2020) also model retirement and, due to their application in the US context, account for employer-provided health insurance, which is not a first-order issue in our analysis on UK data.

[4]On a more technical note, we should also mention that while many studies cited above rely on the calibration of an elaborate model, the results in this paper follow from the full structural estimation of our model, supported by a formal proof of identification.

[5]Interviews for each wave of UKHLS take place over a period of 24 months, but these 24 month periods overlap to ensure that each individual is interviewed once a year. So, for example, wave 1 interviews took place in 2009 and 2010, wave 2 interviews took place in 2010 and 2011, and wave 9 (the most recent at the time of writing) was conducted in 2017 and 2018.

market transitions respondents may have experienced between interviews.

We use the nine waves of data available at the time of writing and focus on men aged 25 to 55. We leave women out of this analysis as we do not model fertility decisions, which are likely to affect both labor market and health trajectories. We stop following individuals after 55 because we do not model retirement decisions. We drop individuals who are never employed during our observation period and we stop following individuals after they become self employed. We exclude observations based on proxy interviews as health and some employment questions are not asked in proxy interviews (we keep all of an individual's non-proxy observations). We further delete a small number of observations with inconsistent information on labor market transitions. We will refer to the data set consisting of only the first observation for each individual as the "initial cross section".

Importantly, UKHLS provides information on individuals' health every year along two dimensions, mental and physical. We focus on mental health (our approach could easily be applied to physical health). Mental health is measured by a continuous summary variable, from 0 to 100, based upon the 12-item Short-Form survey (SF-12), on which we provide more details in Appendix A. For tractability and interpretation, we derive four discrete health states from the continuous health indicator, which we label as 'Good', 'Average', 'Poor' and 'Severe'. As far as we know, there is no consensus on where to set in SF-12 indicator cutoffs to define specific health states. Several studies from the medical literature (for instance Ware et al., 1996 and Salyers et al., 2000) have assessed the validity of the SF-12 mental health indicator against other indicators of depression, anxiety or more severe mental disorders and showed that it performs well in screening different diagnoses. In particular, the study by Gill et al. (2007) shows that thresholds of 36 and 50 are adequate to identify severe psychological conditions and common mental disorders (depression or anxiety), respectively. We use these two threshold to define the 'Severe' (score below 36) and 'Poor' (between 36 and 50) health states.[6] We could not find references for a cutoff between an 'Average' and a 'Good' state so we set it at the 80% quantile (57.2) of the score distribution in the initial cross-section. The distribution of the continuous mental health indicator and the three cutoffs are shown in Appendix A.

Our model is written at a yearly frequency because health is only measured once a year. Identification of the health transition process at a higher frequency would require unwarranted

[6]Scores of 36 and 50 are respectively at the 7^{th} and 36^{th} percentile of the mental health score distribution in the initial cross section.

formal restrictions: we prefer to leave the health process relatively unconstrained and thus work with a yearly frequency. This level of time aggregation prevents our model from replicating observations with more than two job transitions in a year. There are 251 such observations, out of 31,211 observations overall (0.8%) and out of 3,509 observations (7.2%) with at least one transition. We remove those 251 observations.

Our final work sample has 30,960 observations (relating to 8,018 individuals). Each observation consists of the individual's labor market status (non employed, employed part time or employed full time), their (monthly) earnings, their occupation (more on this in the next section), their age, their mental health, their labor market transitions (between full time and part time and/or between employers and/or in and out of employment) since last interview and, if they are not employed, the reason why their previous job ended (thus allowing us to separate quits from job losses). Importantly, we will account for time-invariant individual heterogeneity in the estimation using clustering methods, which we present in Section 5.1.

3.2 Occupation data: O*NET

The health content of jobs is measured based on data from the O*NET program. O*NET, a.k.a. the *Occupational Information Network*, is a database describing occupations.[7] It comes as a list of around 300 descriptors, with ratings of importance, level, relevance or extent, for almost 1,000 different occupations. O*NET data are compiled in the US. Since, to our knowledge, no comparable data set exists for the UK, we have to work on the assumption that the contents of occupations, as described by O*NET, are identical in the UK and US.

O*NET descriptors are organized into ten broad categories: *Skills*, *Abilities*, *Knowledge*, *Experience/Education Levels Required*, *Job Interests*, *Work Activities*, *Work Context*, *Work Values*, *Work Styles*, and *Tasks*. We retain the 16 descriptors from the category *Work Styles* and the 13 descriptors from the *Structural Job Characteristics* section of the *Work Context* category. These provide the descriptors that are most germane to the stress content of the job; descriptors contained in the other files are less directly interpretable in terms of mental health contents. This leaves us with the following 29 descriptors:[8]

[7] O*NET is developed by the North Carolina Department of Commerce and sponsored by the US Department of Labor. More information is available on `https://www.onetcenter.org`, or on the related Department of Labor site `https://www.doleta.gov/programs/onet/eta_default.cfm`.

[8] While O*NET uses 29 different rating systems in total, only two are relevant to the work styles and work context files that we use. Descriptors from the *Work Styles* files are rated in terms of "importance". The "importance" rating measures how important an activity or ability is in exercising the occupation being described. Descriptors in the the *Work Context* file are based upon a "context" rating, which, according to the O*NET

Work styles: *Achievement/Effort*, *Adaptability/Flexibility*, *Analytical Thinking*, *Attention to Detail*, *Concern for Others*, *Cooperation*, *Dependability*, *Independence*, *Initiative*, *Innovation*, *Integrity*, *Leadership*, *Persistence*, *Self Control*, *Social Orientation*, *Stress Tolerance*.

Work Context (Structural Job Characteristics): *Consequence of Error*, *Degree of Automation*, *Duration of Typical Work Week*, *Freedom to Make Decisions*, *Frequency of Decision Making*, *Impact of Decisions on Co-workers or Company Results*, *Importance of Being Exact or Accurate*, *Importance of Repeating Same Tasks*, *Level of Competition*, *Pace Determined by Speed of Equipment*, *Structured versus Unstructured Work*, *Time Pressure*, *Work Schedules*.

O*NET data have been updated at least once a year since 2003. However, not all occupations or descriptors are updated every year: this is done on a rotating basis and it takes about five years for all occupation/descriptor pairs to be updated. We process the O*NET data to construct a panel of occupations covering the same time period as the UKHLS (2009-2017). One challenge with O*NET data is that is compiles information from different sources, mainly either from "analysts" or from incumbent workers. Over time, the data is increasingly based on answers given by incumbent workers. This raises the issue that a given occupation's stress tolerance may be assessed by an analyst one year and a worker the following year. Changes over time may thus reflect genuine changes in occupation stress contents or changes in perspective when substituting, e.g., a worker for an analyst. To address this, we take out an "incumbent worker" fixed effect from every O*NET descriptor and then take averages over time.

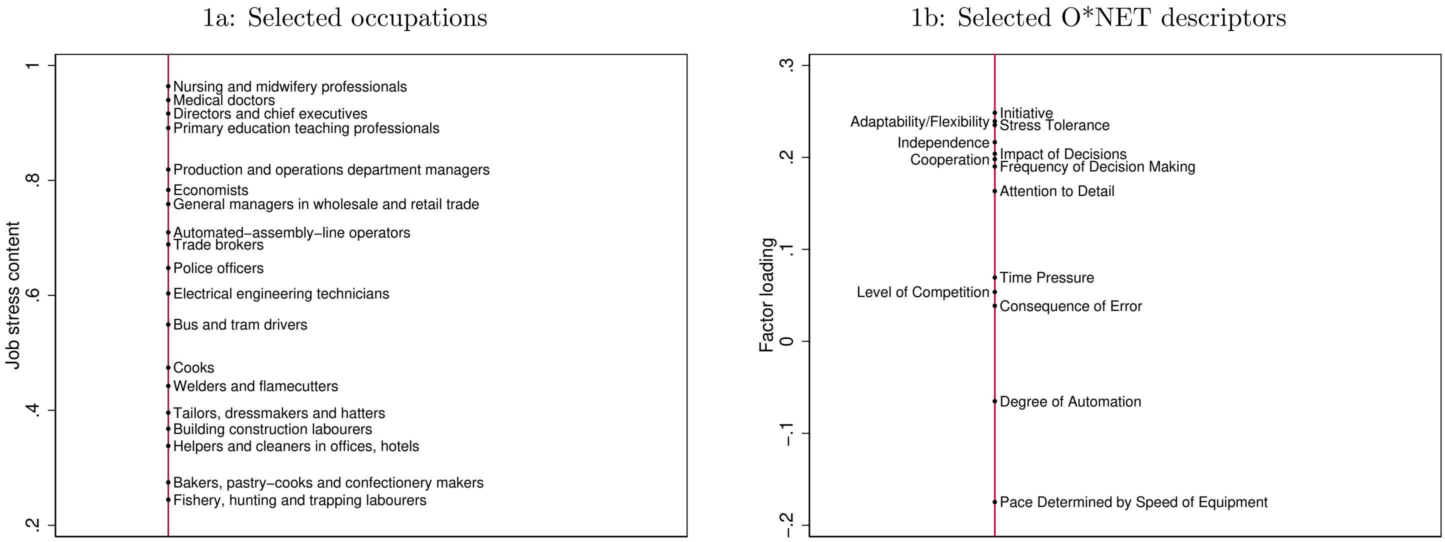
That leaves us with a cross section of occupations with 29 adjusted and averaged descriptors. We run a principal component analysis of these 29 descriptors (in the population of occupations) and keep the first component, to which we apply an affine transform to confine it to $[0, 1]$. We then merge the resulting table into the UKHLS sample described in the previous subsection.

To illustrate our job health content indicator, we show in Figure 1a the value of this variable for a selected sample of occupations. Medical, executive or teaching jobs are classified as high-stress while tailors, bakers or cleaners tend to be low-stress occupations. To further illustrate the job stress indicator, Figure 1b shows the factor loadings on a few selected O*NET descriptors (recall that the job stress variable is the first principal component of our set of descriptors).

documentation, "includes a variety of scales with some unique and specific work context variables". More details are available on `https://www.onetonline.org/help/online/scales`. We rescale all of the descriptors linearly to take values inside the $[0, 1]$ interval.

"Stress tolerance", "adaptability/flexibility" or "impact of decisions" have relatively large positive loadings while the "degree of automation" or "pace determined by speed of equipment" both have negative loadings. Figure 1 thus suggests that our job stress measure is more related to the cognitive and interpersonal dimensions of jobs than to their manual skill requirements.

Figure 1: Job health content - Illustrations



Finally, for computational reasons, we discretize the support of the job stress content variable into five equal-size bins (the five bin mid-points are 0.39, 0.51, 0.64 and 0.77 and 0.89). From now on, this variable will thus be discrete with a five-point support.

3.3 Descriptive statistics

The work sample has 30,960 observations, relating to 8,018 men aged 25 to 55 during our observation period. In the initial cross section, 90% of workers are employed full time, 3% employed part time and 7% are not employed.

Table 1 shows information on the joint distribution of monthly wages and job contents in the initial cross-section of employed workers. The left panel shows the extent of wage dispersion in the data, both within and between full-time and part-time jobs, with $Q_{90/10}$ ratios of 2.8 and 3.9 respectively. The right panel of Table 1 shows the distribution of the job mental health content variable. The higher the value, the more stressful the job. The five values of job health contents are in the top row, their shares in the initial cross section of employed workers is on the second row and the last row shows the average wage among workers in jobs with that level of stress, pointing at a positive relationship in the data between wages and job health contents.

Table 1: Distribution of wages and job health content

	Monthly wage quantiles					Job health content					
	Q_{10}	Q_{25}	Q_{50}	Q_{75}	Q_{90}	*Value*	0.39	0.51	0.64	0.77	0.89
Full time	1,082	1,325	1,746	2,305	3,030	*Share*	0.14	0.22	0.24	0.34	0.07
Part time	346	463	647	866	1,363	*Mean wage*	1,396	1,489	1,910	2,238	2,427

Note: Distributions in the initial cross section.

To investigate correlations, we run an ordered logit regression of the categorical health variable on age, wage, and job content in the initial cross section of employed workers. The estimated coefficients (not shown here) are all significant at the 1% level and show that health is positively correlated with age and wages while it is negatively correlated with the job health content.[9] While these correlations seem to have intuitive signs, they are not causal because of selection into jobs according to age and health. One of the objectives of our structural approach is to estimate causal channels between these variables.

We now examine the life-cycle profiles of the variables of interest, starting with health in Figure 2a. The clearest pattern in this graph is the increase in the proportion of workers in the 'Good' health state with age, especially at later ages. We also observe a decreasing age profile of the probability of being in the worst health state ('Severe') from age 25 to 40 and, to a lesser extent, of the proportion of workers in the 'Poor' health state from age 30 to 50.

In our data, older workers seem to be in better mental health. A legitimate concern would be that the mental health age profiles in Figure 2a might depend on sample selection or on the use of the SF-12 indicator. We show in Appendix A that the descriptive patterns are robust to using either the raw data sample or other questions pertaining to mental health or well being.

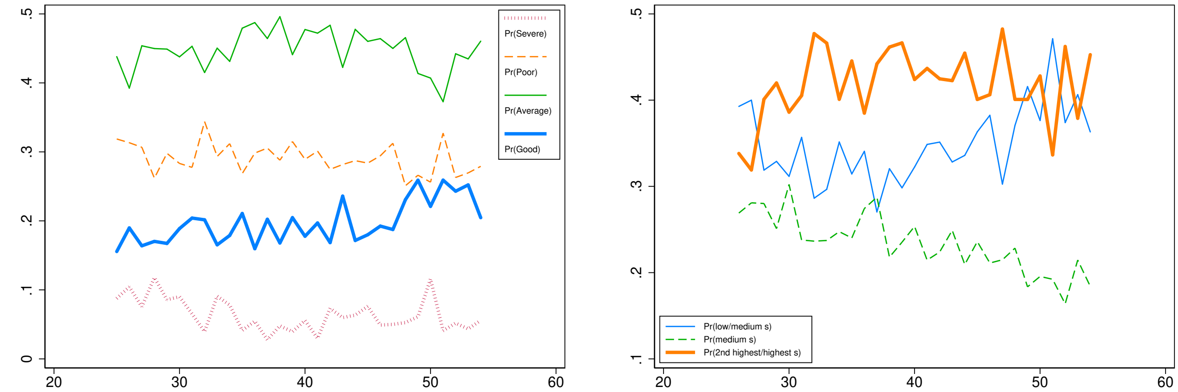
Figure 2b shows the proportions of individuals in the least stressful (first two values of the job health content), medium-stress (third value) or most stressful jobs (two highest values) by age (see the values in Table 1). Even though those profiles are noisy, they suggest that workers are most likely to be employed in the most stressful jobs around mid-career, while the prevalence of low-stress jobs is highest amongst younger and especially older workers.

We then examine labor market turnover by age. Figure 3a shows the proportion of employed workers of a given age who experience a job-to-job transition and the proportion of non employed workers who start a new job (note that those two series are plotted against different vertical axes, for readability). For the former, job-to-job mobility, we see a clear decreasing pattern, where the proportion of job changers goes from around 15% at age 25 to around 5% after 50.

[9] The estimated coefficients of age, wage and job content are respectively 0.0090, 0.234 and −0.641, with no p-value larger than 0.001.

Figure 2: Health and job health content by age

2a: Health 2b: Job health content



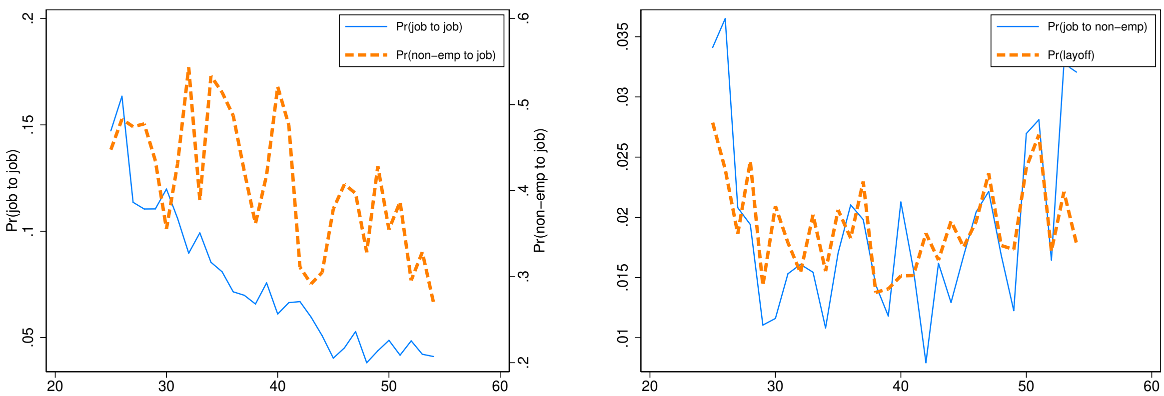
Note: Figure 2a shows the proportion of individuals in each of the four health state by age in the initial cross section. Figure 2b shows the proportion of individuals in low-/high-stress jobs by age in the initial cross section.

This descriptive profile is consistent with a typical job ladder model, where more experienced workers are less likely to find a better job than the one they already have. The non-employment exit rate shows a qualitatively similar age profile, which could come from dynamic selection of workers at the beginning of their career but also from an effect of age on the determinants of the participation decision. Our empirical analysis will be able to disentangle these two channels.

Transition rates out of jobs are shown in Figure 3b. The proportion of employed workers who go to non employment is U-shaped in age. Interestingly, our data make it possible to tell whether job exits were due to a layoff or a quit. Figure 3b shows that the age profile of layoffs is less convex than that of job-to-unemployment transitions. The difference will be coming from quits, thus motivating the extensive participation margin that we will include in our model.

Figure 3: Labor turnover by age

3a: Transitions into new jobs 3b: Transitions out of jobs



4 The Model

4.1 Environment

We write a discrete-time partial-equilibrium model of the labor market with search on and off the job, an extensive employment margin and health dynamics. Workers are characterized by a triple (x, t, h) where $x \in \mathbb{R}$ is a constant attribute (individual heterogeneity), t is the worker's (discrete) age, and h is their time- (or age-)varying health status. Jobs are endowed with an attribute $\mathbf{y} = (w, \ell, s) \in \mathbb{R}^3$, where w is the monthly wage attached to that job, ℓ is the number of working hours in the job, and s is a measure of the (mental) health content of the job. We use $\mathbf{y} = \emptyset$ as a placeholder for job characteristics for non employed workers. Moreover, we restrict working hours to $\ell \in \mathcal{L} = \{0, 1/2, 1\}$ where $\ell = 0$ means that the worker is not employed, $\ell = 1/2$ means part-time work, and $\ell = 1$ means full-time work.

The cost to the worker of supplying ℓ is $c(\ell, s, t, h) + \varepsilon$, independent of x or w, where ε is a labor supply shock, i.i.d. across workers and time with cdf $H(\cdot)$, taking values in $\mathbb{R}_+$. Nonemployed workers generate a flow income of b and incur no explicit labor supply cost.

Each period unemployed and employed workers receive job offers with probability λ_0 and λ_1, respectively. An offer consists of a job type $\mathbf{y}^o = (w^o, \ell^o, s^o)$ drawn from a sampling distribution F, independent of initial employment status. The terms of the offer (wage, working hours, health contents) are non-negotiable, and the recipient's only choice is whether to accept or reject it.

A job's health content s stays constant over time. Working hours are subject to occasional random changes following a first-order Markov process with transition probabilities $\Pr\{\ell'|\ell\}$ which is independent of worker and job attributes. The wage offer is drawn conditionally on the worker's characteristics and on the job's health content and working time. The wage attached to a job then remains constant *conditional on working hours*, i.e. it only changes when working hours change. Thus, the characteristics of an ongoing job change from $\mathbf{y} = (w, \ell, s)$ to $\mathbf{y}' = (w', \ell', s)$ with probability $\Pr\{\ell'|\ell\}$: the health content of the job always stays constant, and the change in wage is entirely driven by the change in working hours. When an employer seeks to impose a change of working hours (e.g. from full time to part time), and a corresponding change in wage, the worker can either accept the change or quit.

Employed workers also face a per-period job destruction probability of δ, and always have the option of quitting into unemployment.

As we discussed in section 3.1, health status is only observed once a year in our data, which constrains us to set our model period to one year. While the vast majority of workers experience no more than one job transition within a typical year, some move either from job to unemployment back into a job, or from unemployment to job back into unemployment between two consecutive annual interviews. Section 3.1 showed that multiple transitions within a year are rare, but non-negligible as a fraction of all labor market transitions. To accommodate those observations, we introduce two additional reallocation shocks. With probability $\widetilde{\lambda}$, workers who were just hit by a job destruction shock δ draw an offer $\mathbf{y}^o$ before the end of the current period and start the following period as employed if they accept that job. Symmetrically, initially unemployed workers who draw a job offer (probability λ_0) face a probability $\widetilde{\delta}$ of that job being terminated before the end of the period. Those features of the model have no particular theoretical content and are meant to handle time aggregation.

Health evolves over a finite set of values following a first-order Markov process with transition probabilities $\Pr\{h'|\mathbf{y}, x, t, h\}$ which depend on the worker's age and human capital, as well as the wage, health content and hours of the job the worker is currently employed at. These probabilities are $\Pr\{h'|\emptyset, x, t, h\}$ for non employed workers. In the application, we restrict health to take on one of four values: $h \in \mathcal{H} = \{\mathrm{G}, \mathrm{A}, \mathrm{P}, \mathrm{S}\}$ for Good, Average, Poor and Severe.

Finally, workers have a working life of T periods, at the end of which they retire and enjoy an exogenous terminal value.

4.2 Dynamics

Workers start each period with knowledge of their current age (t), health status (h), labor supply shock (ε), job status and job attributes ($\mathbf{y} = (w, \ell, s)$ if employed, or b if unemployed). Then, the period unfolds in a way that depends on the worker's employment status.

Unemployed workers. For initially unemployed workers, the sequence of events is as follows:

1. The worker receives his unemployment income b.
2. At the end of the period, he draws his next-period health status $h' \sim \Pr\{h'|\emptyset, x, t, h\}$, labor supply shock $\varepsilon' \sim H$, and his age is incremented.
3. With probability λ_0, he receives a job offer $\mathbf{y}^o \sim F$. With probability $\widetilde{\delta}$, this job offer gets hit by a destruction shock straightaway.[10]

[10] An implicit assumption in our within-period timing is that, whenever two transitions occur within a period, both occur at the end of the period, the intermediate spell being of negligible length. As a consequence, the attributes of that intermediate spell have no impact on current-period values or next-period health. We take this

4. The worker chooses the better option between unemployment (always available) and the job offer (if he received a job offer which was not immediately destroyed).

Employed workers. If the worker is employed at a job of type $\mathbf{y} = (w, \ell, s)$ at the start of the period, the sequence of events is the following:

1. The worker incurs a labor supply cost $c(\ell, s, t, h) + \varepsilon$, and receives the wage w.
2. At the end of the period, he draws his next-period health status $h' \sim \Pr\{h'|\mathbf{y}, x, t, h\}$, labor supply shock $\varepsilon' \sim H$, and working hours $\ell' \sim \Pr\{\ell'|\ell\}$. Age is incremented.
3. With probability δ, his current job is destroyed. In this case, the worker draws a new job offer $\mathbf{y}^o \sim F$ with probability $\widetilde{\lambda}$ before the start of the new period.
4. If his current job has not been destroyed, the worker receives an outside job offer $\mathbf{y}^o \sim F$ with probability λ_1.
5. The worker chooses the best option between unemployment (always available), his old job (unless that job was destroyed by a δ-shock), and the job offer (if he has received one).

Value functions. In what follows, we denote the current characteristics of a worker's job by $\mathbf{y} = (w, \ell, s)$, and the future characteristics of the same job by $\mathbf{y}' = (w', \ell', s)$. We denote the characteristics of a job offer by $\mathbf{y}^o$, and the value of unemployment to a worker with characteristics (x, t, h) by $U(x, t, h)$. The value to a worker with characteristics (x, t, h) and current labor supply shock ε of being employed in a job with attributes $\mathbf{y}$ is $V(x, t, h, \mathbf{y}) - \varepsilon$, where V is the worker's value gross of the labor supply shock. The dynamics of the model described above imply the following value functions for unemployed workers:

$$\begin{aligned} U(x,t,h) = b + (1+r)^{-1} \sum_{h' \in \mathcal{H}} \Pr\{h'|\emptyset, x, t, h\} \Bigg[& \left(1 - \lambda_0 \left(1 - \widetilde{\delta}\right)\right) U(x, t+1, h') \\ & + \lambda_0 \left(1 - \widetilde{\delta}\right) \iint \max\left\{U(x, t+1, h'); V(x, t+1, h', \mathbf{y}^o) - \varepsilon'\right\} dH(\varepsilon') dF(\mathbf{y}^o) \Bigg] \end{aligned} \quad (1)$$

On the right-hand side of (1), the first term is unemployment flow income. From the second term, we reach the end of the period so discounting is applied and the new health and labor supply shocks are drawn. With probability $1 - \lambda_0 \left(1 - \widetilde{\delta}\right)$, the worker either fails to receive a job offer or receives one which is immediately destroyed, and in both cases remains unemployed

shortcut because our data lack information on short spells that start and end between two consecutive interviews. In particular, when a worker moves from unemployment to job to unemployment, the data convey no information about the characteristics (wage, hours, occupation...) of that intervening job spell.

at the start of $t+1$ with value $U(x,t+1,h')$. With probability $\lambda_0\left(1-\widetilde{\delta}\right)$, the worker receives an offer that is not destroyed during the period, he may choose to accept it or stay unemployed, thus receiving the greater of the value of unemployment and that of his job offer.

For employed workers, the value function is:

$$
\begin{aligned}
V(x,t,h,\mathbf{y})-\varepsilon &= w-c(\ell,s,t,h)-\varepsilon \\
&+(1+r)^{-1}\sum_{h'\in\mathcal{H}}\sum_{\ell'\in\mathcal{L}}\Pr\left\{h'|\mathbf{y},x,t,h\right\}\Pr\left\{\ell'|\ell\right\}\Bigg[\delta\left(1-\widetilde{\lambda}\right)U\left(x,t+1,h'\right) \\
&+\delta\widetilde{\lambda}\iint\max\left\{U(x,t+1,h');V(x,t+1,h',\mathbf{y}^o)-\varepsilon'\right\}dH(\varepsilon')dF(\mathbf{y}^o) \\
+(1-\delta)\lambda_1&\iint\max\left\{U(x,t+1,h');V(x,t+1,h',\mathbf{y}')-\varepsilon';V(x,t+1,h',\mathbf{y}^o)-\varepsilon'\right\}dH(\varepsilon')dF(\mathbf{y}^o) \\
&+(1-\delta)(1-\lambda_1)\int\max\left\{U(x,t+1,h');V(x,t+1,h',\mathbf{y}')-\varepsilon'\right\}dH(\varepsilon')\Bigg] \qquad (2)
\end{aligned}
$$

The right-hand side of (2) has the instantaneous utility of being employed (wage minus disutility of labor) on the first line. The discounted, end-of-period continuation value starts on the second line, where a new health status, new labor supply shock and new working hours in the current job are drawn. With probability $\delta\left(1-\widetilde{\lambda}\right)$, the current job is destroyed and no offer is drawn before the new period starts so the worker becomes unemployed (second line). On the third line, the current job is destroyed and the worker received an outside offer so he can take this new job or become unemployed. If the current job is not destroyed (probability $1-\delta$), the worker may receive an outside offer with probability λ_1 and he can then choose between his current job, the outside offer or unemployment (fourth line). On the last line, the worker did not draw an outside offer so he can choose to stay in his current job or become unemployed.

5 Estimation and identification

5.1 First estimation step: individual heterogeneity

Worker types x take values on a discrete set $\{1,\cdots,K\}$. We build on the approach of Bonhomme et al. (2019a), (BLM thereafter, see also Bonhomme et al., 2019b) who estimate the classes of individual heterogeneity by k-means clustering based on relevant moments of outcome variables, where the moments are computed at the individual level. Let individuals be denoted by $i\in\{1,\cdots,n\}$ and let $\mathbf{m}_i$ be a given vector of M moments of individual i's outcomes. We estimate a discrete distribution of individual heterogeneity, with a fixed, finite number K of classes. A partition assigns a group $x_i \in \{1, \cdots, K\}$ to each individual i. The classification step of BLM consists in finding the partition that solves:

$$\min_{\{x_i\} \in \{1,\cdots,K\}^N} \quad \sum_{i=1}^{N} \|\mathbf{m}_i - \widetilde{\mathbf{m}}(x_i)\|^2 , \tag{3}$$

where $\|\cdot\|$ is the Euclidian norm and $\widetilde{\mathbf{m}}(k)$ is the mean of vector $\mathbf{m}$ in group k. Once the classification step is completed, the estimation step (presented in the next subsections) can be carried out, treating individual heterogeneity as observed group dummy variables.

The BLM approach hinges on the existence of an injective map between the individual unobserved heterogeneity and the vector of moments $\mathbf{m}_i$, essentially meaning that the moments used for the criterion in (3) are informative about individual heterogeneity. Consistency of the BLM approach assumes that individual heterogeneity is identified, something we cannot prove formally. We will however use simulations to assess the ability of our approach to recover the estimated distribution of individual heterogeneity (see Section 6.4). In our application, the two main sources of individual heterogeneity are productivity and health. Accordingly, we use observed wages and health status to compute the individual moments $\mathbf{m}$. More precisely, the two moments we use are the average wage (set to 0 if unemployed) and the average of a health indicator equal to 1 for 'Severe', 2 for 'Poor', 3 for 'Average' and 4 for 'Good'.

Our particular application of the BLM clustering technique raises a specific data issue which requires a modification of (3). We observe an individual's job characteristics and health only for up to nine years and our panel data has many different cohorts. Hence we observe some individuals' wage and health in their early thirties while for other individuals these outcomes will be observed when they are in their fifties. Using moments of the wage and/or health over the observation period to form the individual moment vector $\mathbf{m_i}$ in (3) will thus lead to clusters that capture a mix of individual heterogeneity and age effects. To overcome this, rather than solving (3), we consider the following problem:

$$\min_{\{x_i\} \in \{1,\cdots,K\}^N, \ \mathbf{g}:\{1,\cdots,K\}\times\mathbb{R}\to\mathbb{R}^M} \quad \sum_{i=1}^{N} \left\|\mathbf{m}_i - \mathbf{g}\left(x_i, T_i^0\right)\right\|^2 , \tag{4}$$

where T_i^0 is the age of individual i at the beginning of the observation period (thus identifying his cohort) and $\mathbf{g}(1, \cdot), \cdots, \mathbf{g}(K, \cdot)$ are parametric functions of the individual's cohort, accounting for within-class differences in moment outcomes due to age. In our specification, we will write each $\mathbf{g}(x,\cdot)$ function as a (2[nd]-order) polynomial of T:

$$\mathbf{g}(x,T)=\mathbf{g}_{x,0}+\mathbf{g}_{x,1}\cdot T+\mathbf{g}_{x,2}\cdot T^2,\quad x\in\{1,\cdots,K\}, \tag{5}$$

where, for any $x\in\{1,\cdots,K\}$, $\mathbf{g}_{x,0}$, $\mathbf{g}_{x,1}$ and $\mathbf{g}_{x,2}$ are vectors of M parameters estimated by solving (4).[11] Estimating the $\mathbf{g}$ functions in the classification step prevents us from using standard k-means algorithms as in BLM. Further details are provided in Appendix B.

5.2 Second estimation step: health dynamics.

We specify health transition probabilities as follows:

$$\begin{aligned}
\Pr\left\{h'=\mathrm{G}|\mathbf{y},x,t,h\right\}&=1-\Lambda\left(\tau_{\mathrm{G}}+\tau\left(\mathbf{y},x,t,h\right)\right)\\
\Pr\left\{h'=\mathrm{A}|\mathbf{y},x,t,h\right\}&=\Lambda\left(\tau_{\mathrm{G}}+\tau\left(\mathbf{y},x,t,h\right)\right)-\Lambda\left(\tau_{\mathrm{A}}+\tau\left(\mathbf{y},x,t,h\right)\right)\\
\Pr\left\{h'=\mathrm{P}|\mathbf{y},x,t,h\right\}&=\Lambda\left(\tau_{\mathrm{A}}+\tau\left(\mathbf{y},x,t,h\right)\right)-\Lambda\left(\tau_{\mathrm{P}}+\tau\left(\mathbf{y},x,t,h\right)\right)\\
\Pr\left\{h'=\mathrm{S}|\mathbf{y},x,t,h\right\}&=\Lambda\left(\tau_{\mathrm{P}}+\tau\left(\mathbf{y},x,t,h\right)\right)
\end{aligned} \tag{6}$$

where Λ is the logistic cdf, τ_G,τ_A,τ_P are parameters and the threshold function τ is given by:

$$\tau\left(\mathbf{y},x,t,h\right)=\eta^{(t)}\cdot t+\eta^{(t,\ell)}\cdot t\cdot\mathbf{1}\{\ell>0\}+\eta^{(w)}\cdot w+\eta^{(s)}\cdot s+\mathbf{I}_\ell\cdot\eta^{(\ell)}+\mathbf{I}_h\cdot\eta^{(h)}+\mathbf{I}_x\cdot\eta^{(x)}. \tag{7}$$

Throughout the paper we use $\mathbf{I}_h$ to denote a vector of dummy variables for all health values, and use similar notations for heterogeneity classes ($\mathbf{I}_x$) or working times ($\mathbf{I}_\ell$). The first two terms on the right-hand side of (7) reflect the effect of age t on health dynamics, allowing for this effect to depend on employment ($\ell>0$). The third and fourth terms on the right-hand side of (6) pertain to the wage and job health content. The next two terms respectively capture the effect of the employment status and the effect of past health. The last term is the effect of individual heterogeneity x. Whenever a worker is unemployed ($\mathbf{y}=\emptyset$), we set $\ell=w=s=0$.

As (6) shows, health transitions are governed by a first-order Markov process which only involves observable variables $(\mathbf{y},x,t)$ and an i.i.d. shock. Thus (6) can be estimated by maximizing the likelihood of observed health transitions conditionally on $(\mathbf{y},x,t)$. Given the parametric

[11]Deriving the first-order conditions of this problem shows that if $\mathbf{g}_{x,1}=\mathbf{g}_{x,2}=\mathbf{0}_M$ for all x then the optimal $\mathbf{g}_{x,0}$ is the average of individual moments in class x, and the solution corresponds to that of (3).

(logistic) assumption made on the i.i.d. shock, this boils down to an ordered logit regression.[12]

5.3 Third estimation step: job characteristics and utility.

Specification. We start with the specification of the distribution of job offer characteristics. The offer distribution of $\mathbf{y}$ is modeled as follows: s and ℓ are drawn independently from their respective marginal distributions. The sampling distribution of s has a finite five-point support $s_1 < \cdots < s_5$, with $\Pr\{s_j\} = s_j^\alpha / \sum_{k=1}^5 s_k^\alpha$. Working time ℓ takes on two possible values: $\ell = 1$ for full-time work and $\ell = 1/2$ for part-time work (recall that $\ell = 0$ when a worker is unemployed). The Markov process governing working hours in ongoing matches is characterized by two transition probabilities, for which we use the intuitive notation $\Pr\{\text{FT}|\text{PT}\}$ and $\Pr\{\text{FT}|\text{FT}\}$. The working time associated with any new job offer is part-time with probability $\Pr_0\{\text{PT}\}$. Then, conditional on worker characteristics, s and ℓ, the wage offer is:

$$\log w = \beta^{(0)} + \mathbf{I}_x \cdot \beta^{(x)} + (\mathbf{I}_x * s) \cdot \beta^{(s)} + \beta^{(\text{pt})} \cdot \mathbf{1}\{\ell = \text{PT}\} + \sigma^{(\omega)} \cdot \omega \tag{8}$$

where $\omega \sim \mathcal{N}(0,1)$ independently of worker characteristics, s or ℓ. The term $\mathbf{I}_x * s$ means that we interact individual heterogeneity with job stress. The parameters $\beta^{(s)}$ (one for each heterogeneity class) thus capture compensating wage differentials for stressful work in job offers.

We next present the specification of preferences. The distribution of the labor supply shock is specified as log-normal: $\ln \varepsilon \sim \mathcal{N}\left(-\frac{\sigma^{(\varepsilon)2}}{2}, \sigma^{(\varepsilon)}\right)$. This specification implies that the mean of ε is normalized at 1 as said mean is not separately identified from the scale of unemployment income b. Recall that the labor supply shock is assumed i.i.d. across periods (persistence in observed labor supply is produced in the model by search frictions, health and infrequent 'forced' changes in working hours within ongoing jobs). The labor supply shock process is therefore characterized by a single parameter, $\sigma^{(\varepsilon)}$. The cost of labor supply is otherwise specified as:

$$c(\ell, s, t, h) = \mathbf{I}_h \cdot \kappa^{(h)} \times \mathbf{I}_\ell \cdot \kappa^{(\ell)} \times \left(1 + \kappa^{(s)}(s - s_1)\right) \times \left(1 + \kappa^{(t)}(t - 25)\right). \tag{9}$$

This specification allows for interactions between each of the four inputs so that, for example, the cost of working in a stressful job varies with age as well as health or working hours.[13]

[12] We also have to assign a value of $\mathbf{y}$ to workers experiencing a labor market transition (i.e. a change of labor market spell) within the year. We use the $\mathbf{y}$ of the job where the worker spent more time during the year.

[13] The factor specific to health (resp. to job health content, resp. to age) is equal to 1 when the health state is 'Good' (resp. when in the least stressful type of job s_1, resp. at the youngest age in our sample, 25 years old.)

Lastly, we fix the discount rate r at 10 percent per annum.

Parameters estimated directly from the data. Several parameters pertaining to job characteristics, job dynamics and preferences have direct data counterparts, and can therefore be estimated directly. We start with coefficients in the wage equation. Equation (8) predicts that log wage growth amongst individuals who stay in the same job, but go from full time to part time between two consecutive periods is equal to $\beta^{(\mathrm{pt})}$, thus offering a simple estimate of $\beta^{(\mathrm{pt})}$ as the mean log wage growth in that population.

The coefficients $\beta^{(x)}$ and $\beta^{(s)}$ in (8) can also be estimated directly. Define the log wage net of part-time effects as $\ln w_{\mathrm{net}} = \ln w - \beta^{(\mathrm{pt})} \cdot \mathbf{1}\{\ell = \mathrm{PT}\}$. Next define $\overline{w}_{\mathrm{net}}(x, s)$ as the maximum observed wage amongst all matches involving a type-x worker in a type-s job. In a large data set, this equals $\beta^{(0)} + \mathbf{I}_x \cdot \beta^{(x)} + (\mathbf{I}_x * s) \cdot \beta^{(s)} + \sigma^{(\omega)} \cdot \overline{\omega}$, where $\overline{\omega}$ is the upper support of ω. An OLS regression of $\overline{w}_{\mathrm{net}}(x, s)$ on worker type dummies $\mathbf{I}_x$ and their interaction with job content s in a cross-section of employed workers thus provides consistent estimates of $\beta^{(x)}$ and $\beta^{(s)}$.

Next, the within-job full time/part time transition probabilities, $\Pr\{\mathrm{FT}|\mathrm{FT}\}$ and $\Pr\{\mathrm{FT}|\mathrm{PT}\}$, can be estimated from observed within-job changes in hours. For any $(\ell, \ell') \in \{PT, FT\}^2$, the probability $\Pr\{\ell' \mid x, t, h', \mathbf{y} = (w, s, \ell)\}$ of going from hours ℓ to ℓ' while staying on the same job, conditional on worker and job characteristics $(x, t, h', \mathbf{y})$ an be estimated directly from the data.[14] In Appendix C, we show that:

$$\frac{\Pr\{\ell' = \mathrm{PT} \mid x, t, h', \mathbf{y} = (w, s, \mathrm{FT})\}}{\Pr\{\ell' = \mathrm{PT} \mid x, t, h', \mathbf{y} = (w, s, \mathrm{PT})\}} = \frac{1 - \Pr(\mathrm{FT}|\mathrm{FT})}{1 - \Pr(\mathrm{FT}|\mathrm{PT})} \tag{10}$$

Likewise, using the conditional probabilities of observing a worker staying in the same full-time job or going from part- to full-time in the same job, we obtain the ratio $\Pr\{\mathrm{FT}|\mathrm{FT}\}/\Pr\{\mathrm{FT}|\mathrm{PT}\}$. Solving the resulting linear system of two equations, we obtain $\Pr\{\mathrm{FT}|\mathrm{FT}\}$ and $\Pr\{\mathrm{FT}|\mathrm{PT}\}$.

The exogenous job destruction rate δ is estimated directly as the probability that a worker employed in a given year either experiences a job-to-unemployment-to-job transition or is non employed the following year, and says that the reason that their previous job ended was one of "made redundant", "dismissed/sacked", "temporary job ended" or "other reason".[15]

The first reallocation shock $\widetilde{\delta}$ can also be directly estimated from the probabilities of making an unemployment-to-job transition (call it $\Pr(u2j)$) or of making an unemployment-to-job-to-

[14]Note that h' denotes next-period health, i.e. health at age $t+1$ (an observable variable).

[15]This leaves out the following other possible reasons: "left for better job", "took retirement", "health reasons", "left to have baby", "look after family", "look after other person", and "moved area".

unemployment transition (call it $\Pr(u2j2u)$). The former probability, $\Pr(u2j)$ is equal to λ_0 (draw a job offer) times $1-\widetilde{\delta}$ (that job is not terminated within the year) times the probability of accepting the job offer. $\Pr(u2j2u)$ is the product of the probability of receiving an offer (λ_0), of accepting a job offer and of the job offer being destroyed before the start of the next period $\left(\widetilde{\delta}\right)$. Hence, $\widetilde{\delta}$ can be identified from the ratio $\Pr(u2j2u)/[\Pr(u2j2u)+\Pr(u2j)]$, which is readily available from the data.

The other reallocation shock, $\widetilde{\lambda}$ cannot be estimated directly however the ratio $\widetilde{\lambda}/\lambda_0$ can. Denoting as $\Pr(u2j)$ (resp. $\Pr(j2u2j)$) the probability of making an unemployment-to-job transition (resp. a job-to-unemployment-to-job transition), it is easy to show that the ratio $\widetilde{\lambda}/\lambda_0$ equals $\frac{1-\widetilde{\delta}}{\widetilde{\delta}} \cdot \frac{\Pr(j2u2j)}{\Pr(u2j)}$, which is known.[16] Hence, we just need to identify λ_0 and $\widetilde{\lambda}$ follows.

Indirect inference. At this point we are left with the following vector of parameters to estimate: the job offer arrival rates λ_0 and λ_1, the standard deviation $\sigma^{(\omega)}$ of the log-wage offer residual, the distribution parameter α of health content in job offers, the standard deviation of the labor supply shock $\sigma^{(\varepsilon)}$, the share of part time job offers $\Pr_0(\ell' = \mathrm{PT})$, and the effect of age $\left(\kappa^{(t)}\right)$, job content $\left(\kappa^{(s)}\right)$, full/part-time $\left(\kappa^{(\ell)}_{\mathrm{FT}} \text{ and } \kappa^{(\ell)}_{\mathrm{PT}}\right)$ and health $\left(\kappa^{(h)}_G, \kappa^{(h)}_A, \kappa^{(h)}_P, \kappa^{(h)}_S\right)$ on the disutility of labor (where $\kappa^{(h)}_G$ is normalized to 1).

We estimate these parameters using indirect inference. We simulate a panel of just over 160,000 workers (20 times the size of our initial cross-section) over nine years (eight full years, plus the initial observation) to replicate our data sample. The initial state of each worker (age, health, employment state, job type, wage) is taken directly from the initial cross-section in the data. We then arrange observations, both in simulated and actual data, into 10 three-year worker age classes — age classes thus contain workers that are 25-27, 28-30, etc. up to 55 years old at the time of observation —, and compute averages by age group as moments to match. This indirect inference step is the computationally costly part of our estimation and motivates many of the simplifying assumptions of the model (for instance the Markovian dynamic process of health) and discretization of some state variables (health and job health content).

The set of matched moments includes age-specific averages of relevant variables. It consists of mean log wage and mean job health content by age class (2×10 moments); unemployment and part-time rates by age class (2×10 moments); mean health of unemployed, part-time and

[16] To be precise, the probabilities $\Pr(j2u2j)$ and $\Pr(u2j)$ first need to be estimated conditionally on x, t and h to allow for compositional differences between employed and unemployed workers with respect to individual heterogeneity, age and health.

full-time workers by age class (3×10 moments, where health encoded as $1 =$'Severe', $2 =$'Poor', $3 =$'Average', $4 =$'Good'); quit, unemployment-to-job and job-to-job transition rates by age class (3×10 moments).

In addition, we match the coefficients obtained from the following pooled OLS regressions, to capture covariances between variables of interest: log wage on worker type, health, and part-time dummies, age, job content, and a constant (10 moments); squared residual of the previous wage regression on worker type, health and unemployment-to-job transition dummies, and a constant (8 moments); yearly log wage change on job content both at the start and at the end of the period, a job-to-job transition dummy, and a constant (4 moments); employment dummy on worker type and health dummies, age, and a constant (8 moments); current job type s on worker type and health dummies, age, and a constant (8 moments). Finally, to inform the job-type sampling distribution, we target the distribution of job types s accepted by workers upon exiting unemployment (4 additional moments).

Standard errors. We estimate standard errors by bootstrap. Our multi-step estimation procedure is applied to 200 bootstrap samples and we take the standard errors from the resulting distribution of parameter estimates. All our estimation steps, from the clustering procedure described in section 5.1 to the indirect inference step presented above are run on each bootstrap samples. Hence, the standard errors show in the following results sections account for all the variation arising at each stage of our estimation procedure.

5.4 Identification

As discussed in Section 5.1, we do not formally prove identification of the unobserved heterogeneity classes, but we will conduct a numerical test in Section 6.4 to gauge the internal consistency of our estimation approach. The second step, presented in Section 5.2, is a straightforward ordered logit regression using observed regressors only and allows us to identify and estimate all the parameters in the health dynamic process. In the last step, as we explain in Section 5.3, several parameters have direct data counterparts and are thus identified by construction. However identification is less straightforward for the final set of parameters which are estimated by indirect inference. These parameters pertain to the arrival rates and distribution of job offers and to individual preferences (namely, the disutility of work).

An important feature of our analysis is that we are able to formally prove identification of those parameters from our data under weak assumptions, and conditional on knowledge of the distribution of individual heterogeneity (and the variance of the labor supply shock). The technical proof is in Appendix C and consists of three steps. Essentially, we face a selection model where workers sort themselves between unemployment and different types of jobs. The first step identifies the outcome equation (and consequently the selection equation). The second step exploits the Bellman equations (1)-(2) to show that the instantaneous utility functions follow from the value functions in the selection equation by using a mapping identified from the data. The last step then combines the first two results to identify the preference parameters.

6 Estimation results

6.1 Model estimates: individual heterogeneity.

The first step of our estimation procedure produces a partition of the sample of workers into four heterogeneity classes (see Section 5.1). In Table 2 we describe these classes by showing class-level averages of various relevant variables.

Class 4 is a group of high-skill/high-wage workers who tend to be in average health and work in more stressful jobs relative to other groups. At the other end of the spectrum, Classes 1 and 2 earn similarly low wages and work in moderately stressful jobs. The differences between Classes 1 and 2 is that Class 1 has much poorer health and a lower employment rate. Finally, Class 3 is in between Class 4 and Classes 1 and 2 in terms of wages, and job stress contents, and tends to be in mediocre health (better than Class 1 but worse than Class 2 or 4).

Table 2: Individual heterogeneity classes

	Share	**Health** (%)				**Employed**	**Average...**		
Class	(%)	**Severe**	**Poor**	**Average**	**Good**	(%)	**Wage**	**Content**	**Age**
1	26.0	21.2	57.5	18.5	2.7	88.1	1,444	0.61	39
2	34.6	0.7	7.5	51.7	40.1	93.3	1,510	0.60	40
3	27.3	2.4	32.8	54.0	10.8	95.1	2,132	0.67	39
4	12.1	1.3	22.3	60.3	16.1	98.0	3,390	0.73	36

In short, our first estimation step has produced four groups of workers along essentially two axes: high-productivity/healthy workers (Class 4), average-productivity/unhealthy workers (Class 3), low-productivity/healthy workers (Class 2) and low-productivity/unhealthy workers (Class 1). Lastly, we should note that the four classes are relatively close in age. This indicates that the age adjustment we brought to the BLM approach prevents age cohorts from driving the partitioning of workers into the four classes.

6.2 Model estimates: health dynamics.

The second estimation step maximizes the likelihood of health transitions given the model structure (6) and the logistic assumption on health shocks. Parameter estimates of our reduced-form specification (6) of the health dynamic process are reported in Table 3. While the magnitudes of those estimates are not straightforward to interpret, their signs mostly are.

Table 3: Health dynamics - Ordered logit estimates

$\eta^{(t)}$	0.0153	(0.008)	$\eta^{(P)}$	0.8688	(0.063)	$\eta^{(FT)}$	-0.2237	(0.112)
$\eta^{(t,\ell)}$	0.0048	(0.009)	$\eta^{(A)}$	1.3877	(0.077)	$E_x(\tau_P)$	-1.7822	(0.188)
$\eta^{(w)}$	0.0615	(0.036)	$\eta^{(G)}$	1.8858	(0.115)	$E_x(\tau_A)$	1.1405	(0.195)
$\eta^{(s)}$	-0.3286	(0.111)	$\eta^{(PT)}$	-0.1092	(0.141)	$E_x(\tau_G)$	3.9497	(0.179)

Note: Ordered logit estimates of equations (6)-(7).

The estimated coefficients of the wage and of the job content in the threshold (7) are respectively significantly negative and positive (only at the 9% significance level for the wage), indicating that better pay increases the probability of health improving across years while the converse is true for the job stress level. Also, age tends to help improving (mental) health — a result paralleling the descriptive fact shown on Figure 2. Finally, working full time has a ceteris paribus adverse impact on mental health compared to not working. When thinking about this result, however, one has to bear in mind that the state of 'not working' is associated with a wage of zero, which in itself is detrimental to health. The overall impact of being unemployed (as opposed to being employed, say, full time at a certain wage) is therefore ambiguous, which partly motivates the next set of results.

To better appreciate the magnitude of the impact of the various determinants of health dynamics, we show in Table 4 a selection of transition probabilities based upon the estimated model. We compute the yearly probabilities of going to the severe or poor health state from any of the four health states for a given employment status, age, wage and job content, where each probability is averaged over the distribution of individual heterogeneity x in the initial cross section (estimated in the first step).

The top left panel shows how health transition probabilities vary for 40-year old men between working full time at a median-wage/medium stress job and being unemployed. We note that unemployment lowers the probability to leave the bad (severe or poor) health state and increases the rate at which an individual in good or average health enters the poor or severe health state

Table 4: Estimated probabilities of going to the severe or poor health state

Employment status					
40 years old					
Med. wage/stress			*Unemployed*		
	S	P		S	P
S	0.16	0.48	S	0.17	0.49
P	0.08	0.40	P	0.09	0.41
A	0.05	0.34	A	0.06	0.35
G	0.03	0.27	G	0.04	0.28

Age					
Median wage, medium stress					
30 years old			*50 years old*		
	S	P		S	P
S	0.19	0.49	S	0.14	0.47
P	0.10	0.42	P	0.07	0.38
A	0.06	0.36	A	0.04	0.31
G	0.04	0.30	G	0.03	0.25

Wage					
40 years old, medium stress					
1^{st} *wage decile*			9^{th} *wage decile*		
	S	P		S	P
S	0.17	0.48	S	0.16	0.48
P	0.09	0.40	P	0.08	0.40
A	0.05	0.34	A	0.05	0.33
G	0.03	0.27	G	0.03	0.27

Job health content					
40 years old, median wage					
low-stress job			*high-stress job*		
	S	P		S	P
S	0.15	0.48	S	0.17	0.49
P	0.08	0.39	P	0.09	0.41
A	0.05	0.33	A	0.06	0.35
G	0.03	0.26	G	0.04	0.28

Notes: Origin health states in rows, destination states in columns.
S: Severe - P: Poor - A: Average - G: Good. Employed means full-time employment.
Transition probabilities averaged over individual heterogeneity.

(the difference in transition rates is small, around 1 percentage point).

The top right panel shows how health dynamics changes with age (for someone employed at the median wage and middling job content). Older workers are more likely to recover from ill health and are less likely to experience health transitions in the other direction. The magnitude of the age effect seems very moderate but can become substantial over several years.

The bottom left panel shows health transition probabilities for 40 year old men as a function of the wage. It reveals that the wage has little impact on the probability of going to the severe or poor health state: going from the 1^{st} to the 9^{th} wage decile while keeping everything else constant only decreases the risk of being in severe or poor health the following year by no more than one percentage point.

However, the bottom right panel shows that the job stress content has a visible impact on health transitions: working in the most stressful type of job makes it less likely to leave the severe or poor health states and more likely to enter those bad health states for workers currently in average or good health. Therefore, even though the wage does not appear to have a noticeable impact on health dynamics once unobserved heterogeneity is accounted for, the reverse causality channel from labor market outcomes to mental health is still active through the effect of employment and of the stress content of jobs.

6.3 Model estimates: utility and job characteristics

The remaining parameter estimates pertain to the arrival rates and sampling distribution of job offers (Table 5), and to worker preferences, including the labor supply cost function (Table 6). First, recall that several parameters in the wage offer equation (8) are allowed to vary with individual heterogeneity x. We report in Table 5 the average and standard deviation of the wage intercept $\beta^{(x)}$ and of the wage effect of the job health content $\beta^{(s)}$. The average impact of job stress content on wage offers is significantly positive, $\mathbf{E}_x\left(\beta^{(s)}\right) = 0.416$: more stressful job offers tend to be compensated by higher wages. While there is variation in this compensation coefficient across individuals (see the support of $\beta^{(s)}$ in the notes for Table 5), it is positive for all four heterogeneity classes.

Another noticeable result in Table 5 is that the exponent of the sampling distribution of job health contents $\alpha = 2.48$ implies that the sampling distribution is skewed toward high-stress jobs, more so than the observed cross-section distribution of health contents in the population of employed workers (Table 1). This indicates some degree of self-selection of workers away from higher-stress jobs.

Table 5: Job offer characteristics and arrival rates

Job offers						Arrival rates					
$\mathbf{E}_x\left(\beta^{(x)}\right)$	6.393	(0.054)	$\beta^{(\mathrm{pt})}$	-0.285	(0.072)	λ_0	0.683	(0.015)	$\widetilde{\delta}$	0.045	(0.010)
$\mathrm{std}_x\left(\beta^{(x)}\right)$	0.255	(0.033)	$\sigma^{(\omega)}$	0.658	(0.048)	λ_1	0.609	(0.025)	Pr(PT—offer)	0.027	(0.001)
$\mathbf{E}_x\left(\beta^{(s)}\right)$	0.416	(0.048)	α	2.484	(0.062)	δ	0.018	(0.001)	Pr(FT—PT)	0.882	(0.026)
$\mathrm{std}_x\left(\beta^{(s)}\right)$	0.118	(0.043)				$\widetilde{\lambda}$	0.234	(0.028)	Pr(PT—FT)	0.003	(0.0004)

Notes: Parameters $\beta^{(x)}$ and $\beta^{(s)}$ vary with individual heterogeneity x.
Estimated support points for $\beta^{(x)}$: 6.09, 6.31, 6.57 and 6.89. Estimated support points for $\beta^{(s)}$: 0.60, 0.33, 0.32 and 0.48.
$\mathbf{E}_x$ (resp. std_x) is the expectation (resp. standard deviation) across workers. Estimated standard errors in parentheses.

The utility parameter estimates in Table 6 show that, as expected, worse health, higher job stress levels, age and full time work increase the disutility of labor. For ease of interpretation, we can replace the estimated values from Table 6 into the labor cost formula (9) for different values of health, age and job content. The resulting cost estimates are shown in Table 7.

We note that the disutility of labor can almost double from age 30 to 50. Controlling for age and health, the disutility of labor increases substantially when going up one level of job stress (from s_n to s_{n+1}). Lastly, we note that being in poor health increases the cost of labor supply dramatically relative to being in good health, regardless of age. These large effects of poor health and job stress on the cost of labor supply make health an important determinant of individual labor market decisions. The counterfactual analysis conducted in the next section

Table 6: Utility estimates

Disutility of labor						Supply shock		
$\kappa^{(t)}$	0.044	(0.005)	$\kappa_S^{(h)}$	43.42	(0.98)	$\sigma^{(\varepsilon)}$	2.16	(0.075)
$\kappa^{(s)}$	49.58	(3.67)	$\kappa_P^{(h)}$	25.23	(0.84)			
$\kappa_{\mathrm{FT}}^{(\ell)}$	0.966	(0.029)	$\kappa_A^{(h)}$	13.20	(0.65)	**Unemp. income**		
$\kappa_{\mathrm{PT}}^{(\ell)}$	0.577	(0.019)	$\kappa_G^{(h)}$	1.00	(norm.)	b	30.11	(1.02)

Notes: Estimated standard errors in parenthesis.

Table 7: Labor supply cost estimates (GBP/month)

	Poor health					Good health				
	s_1	s_2	s_3	s_4	s_5	s_1	s_2	s_3	s_4	s_5
age 30	29	207	385	564	742	1	8	15	22	29
age 50	50	361	672	983	1,294	2	14	27	39	51

Notes: Units are GBP/month. Job stress grows from s_1 (lowest) to s_5 (highest).

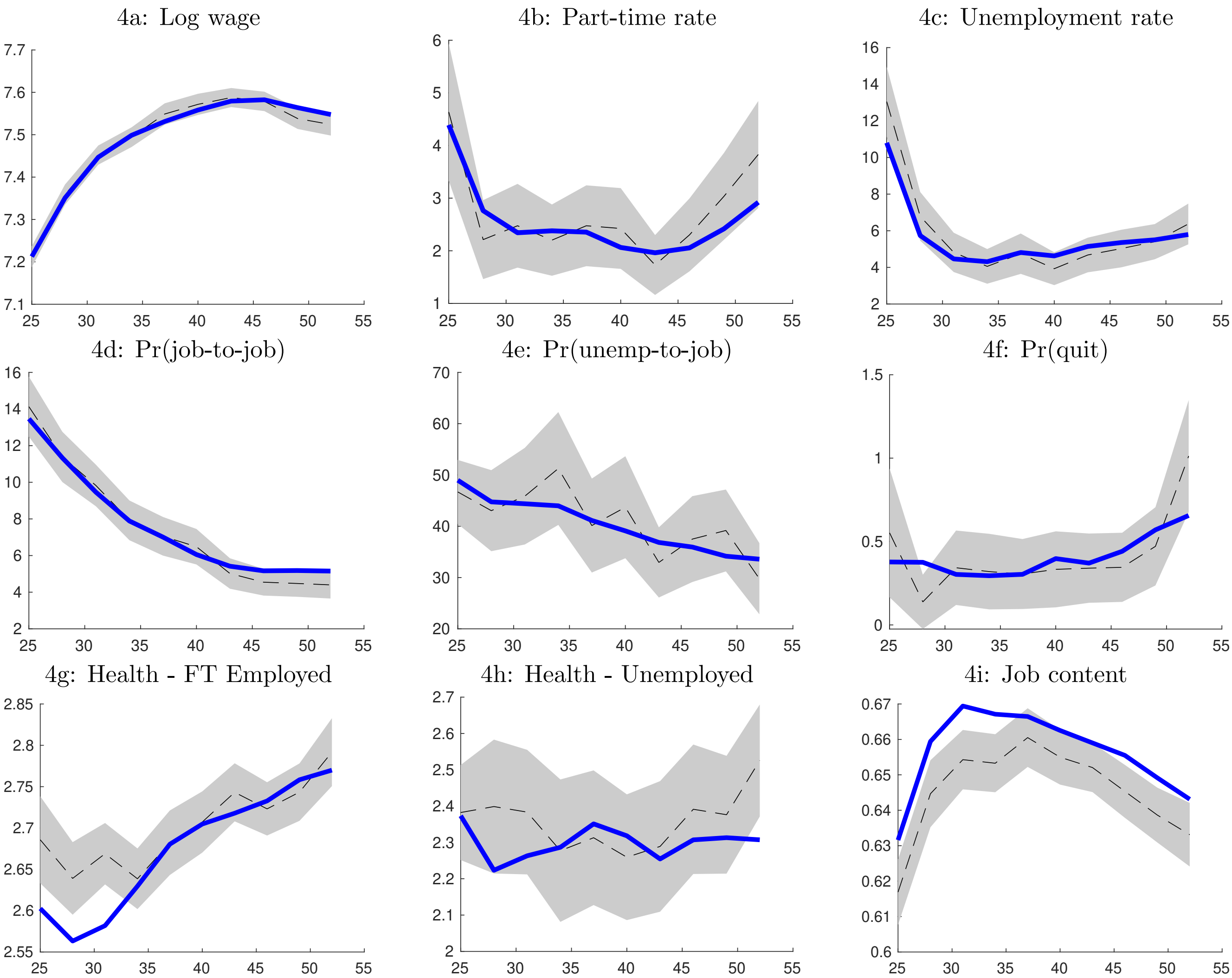
will further investigate this interaction.

6.4 Model fit

Targeted moments. As explained in Section 5.3, we match two types of moments in the last step of the estimation. The first set of moments is a series of age profiles for health, job characteristics and labor turnover. The second set of moments follows from auxiliary regressions. Figure 4 presents the fit of the model for the first type of moments: model-predicted age profiles are shown together with their empirical counterparts and associated 95% confidence bands.

Overall our model offers a very good fit to the age profiles of the wage, unemployment and part-time rates, and health conditional on employment status (despite a slight tendency to under-predict the health of full-time workers under the age of 35). We do not show health for part-time workers to save space but the fit is as good as for the other two employment statuses. The graphs in the middle row of Figure 4 show that the model can also fit the profiles of labor turnover, whether between employment and unemployment or across jobs. The U shaped pattern of quits is captured by our model. The last graph shows that we tend to over-predict the average job stress content. This may be due to an overly restrictive specification of the job offer distribution (only one parameter governs job health content offers) and could be fixed by allowing for a more flexible parametric distribution. However, looking at the magnitude of the difference between the data and model prediction, the model only overstates s by less than 0.02 (less than 3% of the mean stress content). We thus deemed the computational cost of

Figure 4: Model fit - Age profiles



Notes: Age on horizontal axis. Solid line: Model. Dashed line: Data. Shaded area: 95% confidence bands around empirical moments. The health outcome is an average of a variable equal to 1, 2, 3 or 4 if health is S, P, A or G respectively.

introducing more parameters to be unjustified as the model offers an accurate enough prediction of job stress content by age.

The fit to auxiliary regression parameters is shown in Appendix D, Figure 12, to save space in the main text. All regression coefficients are reasonably well captured by the model, which is thus able to fit a large set of moments relating to health, job outcomes or labor turnover with a relatively small number of parameters (only 15 in the indirect inference step of the estimation).

Unobserved heterogeneity. To conclude this section on model fit, we investigate the internal consistency of our estimated distribution of unobserved heterogeneity. As discussed in section 5.1, we cannot formally prove that the set of individual outcomes used for clustering identifies the heterogeneity classes. What we can show, however, is that our clustering is internally consistent in that it can recover classes used as inputs in the model. To that end we simulate individual health and labor market trajectories using the estimated model with the classes produced in the first estimation step from the real data (see Section 6.1), then run our clustering algorithm on the simulated data. Each 'simulated' class should overlap strongly with only one of the 'empirical' classes that were used as inputs.

Table 8 shows the mapping between the two partitions. Each number is the probability that an individual from the 'empirical' class (rows) be assigned to a 'simulated' class (columns).[17] The results show that we are not far from a bijective mapping between the two partitions, with a minimal 'matching rate' of 95% (and up to almost 99%). We view these results as strongly encouraging for the validity of our estimation of unobserved individual heterogeneity.

Table 8: Fit of the re-estimated heterogeneity classes

	% in simulated class			
Empirical class	1	2	3	4
1	0.7	1.5	0.0	97.7
2	98.8	0.8	0.2	0.3
3	2.0	95.4	1.1	1.5
4	0.2	4.6	95.2	0.0

7 The effects of health and employment shocks

7.1 Job loss

The value of employment at age 30. We begin our counterfactual analysis by quantifying the utility cost of job loss. We do so by computing the difference in expected utility between working and being unemployed. Formally, the value of being employed in a type-$\mathbf{y}$ job relative to being unemployed for a type-x worker of age t and with health h can be written as:

$$\Delta^{(\ell)}V(t,\mathbf{y},h) = \mathbf{E}_{x,\varepsilon}\left[V(x,t,h,\mathbf{y}) - \varepsilon - U(x,t,h)\right], \tag{11}$$

where V is defined by (2), U is defined by (1) and we average the difference over the distribution of individual heterogeneity x and labor supply shock ε. We report the value of employment $\Delta^{(\ell)}V(t,\mathbf{y},h)$ for several health states, wage deciles and job stress contents in Table 9, together with the relative welfare loss due to unemployment (in parentheses).

[17] Note that class labels do not need to match i.e. the 'empirical' class 1 needs to match one of the four 'simulated' classes but not necessarily the 'simulated' class 1.

Table 9: Value of a full-time job (*vs* unemployment) at age 30

	Medium-stress job			High-stress job		
	Wage decile			Wage decile		
Initial health	2^{nd}	5^{th}	8^{th}	2^{nd}	5^{th}	8^{th}
Severe	24,078 (16%)	49,265 (27%)	89,718 (40%)	5,204 (4%)	24,100 (15%)	58,109 (30%)
Poor	27,744 (17%)	52,915 (28%)	93,356 (40%)	12,545 (9%)	31,511 (19%)	65,565 (32%)
Average	30,122 (19%)	55,285 (29%)	95,718 (41%)	17,303 (11%)	36,308 (21%)	70,387 (34%)
Good	32,506 (20%)	57,660 (30%)	98,087 (41%)	22,071 (14%)	41,111 (23%)	75,212 (35%)

Note: Value differences $\Delta^{(\ell)}V(t, \mathbf{y}, h)$ as defined in (11), *measured in GBP.*
Relative differences in parentheses: $\Delta^{(\ell)}V(t, \mathbf{y}, h)/\mathbf{E}_{x,\varepsilon}\left[V(x, t, h, \mathbf{y}) - \varepsilon\right]$.

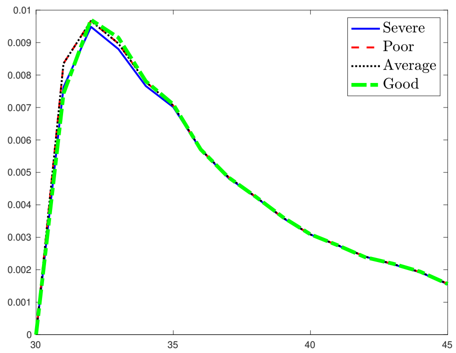
We estimate the value of employment at age 30 to be substantial, ranging from £5,000 for a low-wage highly stressful job (when in severe health) to almost £100,000 for a high-wage and medium stress job (when in good health). These values are in the order of four months to four years of the average wage in jobs of the corresponding stress level (see Table 1). The relative welfare loss due to unemployment is below 10% only for workers in bad health employed in a low-wage and high-stress job. Becoming unemployed in any of the other cases we consider causes a two-digit relative welfare loss which can reach 40% for high-wage workers.

The effect of job loss on labor market and health trajectories. Next, we take a more dynamic perspective and evaluate the effect of job loss at age 30 on future paths of individual careers and health. First, we simulate the health and career trajectories of a representative[18] worker who at age 30 is working full time at the median wage and a middling stress level (s_3, the third out of five points of support); these workers are our 'control group'. Second, we simulate the same trajectories for a representative worker who is unemployed at age 30 (the 'treatment' group). We define the effect of job loss on future labor market and health paths as the difference between the age profiles of the treatment and control groups.

Figure 5 shows the impact of job loss on future health outcomes. Specifically, it shows the difference between treatment and control groups in the probability of being either in the severe or poor health state (which we collectively refer to as 'bad health'), conditional on initial health. We first note that the initial health state does not affect the impact of unemployment on health as the four trajectories are very close. Importantly, recall that this effect is measured as a first difference (along the employment margin). Hence Figure 5 does not tell us that the level of the probability to be in bad health next year for unemployed workers is the same irrespective

[18]We simulate trajectories for each worker type x, then take the average over the estimated distribution of x.

Figure 5: Health effects of job loss at age 30



Note: The vertical axis shows the difference in the probability of being in severe or poor health, given initial health at age 30, between an unemployed worker and a full-time worker at the median wage/medium job stress content. The horizontal axis shows age (in years).

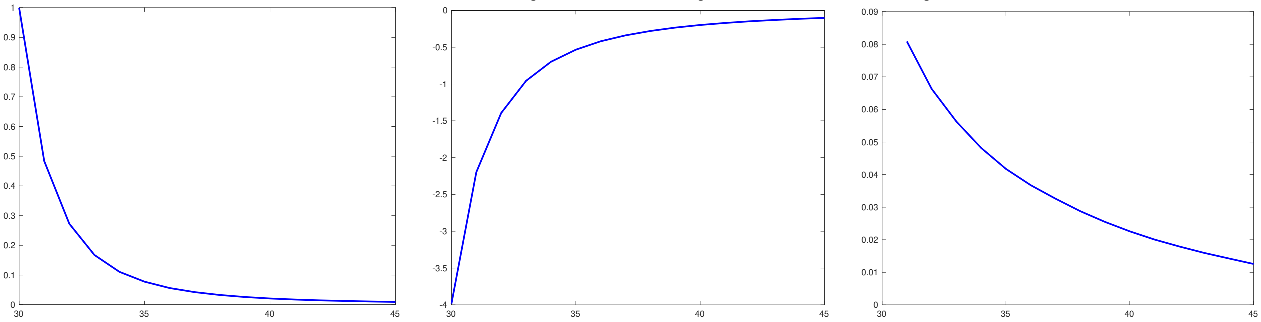
of their health this year. Instead it shows that the increase in the risk of bad health due to unemployment (*vs* employment) is not affected by the worker's health at the time of job loss. Figure 6 will delve further into the interaction between employment and health.

Figure 5 shows that losing a median-wage/medium-stress job at age 30 (and becoming unemployed) causes a very persistent, albeit small increase in the probability of being in poor or severe health. This effect peaks at almost 1% two years after job loss and decreases slowly afterwards. The moderate but noticeable increase in the risk of bad health is consistent with our estimation results for the health dynamic process (see the top left panel of Table 4).

The persistence of the effect of job loss on health may arise from differences in labor market trajectories between the treatment and control groups, which cause different subsequent health outcomes. To assess that, we turn to labor market outcomes in Figure 6. In the first row, we focus on the average treatment effect of going from a median wage/medium stress full-time job to unemployment at age 30 for a worker in average health. Figure 6a shows that the unemployment rate remains higher for the treated group after five years. The income loss, in Figure 6b, shows even greater persistence, consistently with the recent macro-labor literature. Indeed, ten years after the shock, income (wage or unemployment income) is still 23 log points lower for workers who lost their job at age 30 than for those who kept it. This is partly explained by a slow return to employment and by the job ladder mechanism, whereby workers gradually select into better jobs through on-the-job search. Interestingly, we note in Figure 6c that workers who lost their job at a medium stress level return to work in more stressful jobs than those who initially stayed employed. The combined effect of income loss, persistent unemployment and more stressful job contribute to the persistent health effect shown in Figure 5.

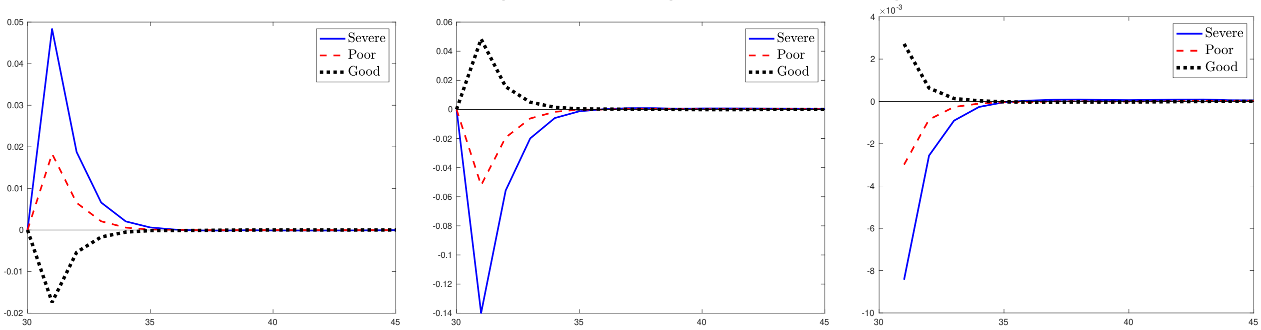
Figure 6: Labor market effects of job loss at age 30

Conditionnal on being in the Average health state at age 30



6a: Unemployment rate 6b: Income (log) 6c: Job health content

Relative to being in the Average health state at age 30



6d: Unemployment rate 6e: Income (log) 6f: Job health content

Note: Age on horizontal axis. The vertical axis shows the labor market outcome difference between an unemployed worker and a full-time worker at the median wage/medium job stress content.

Figures 6a-6c showed the average treatment effect of job loss on labor market outcomes conditionally on being in average health at age 30. To further explore the two-way relationship between health and work, we show in 6d-6f how these effects vary with the worker's initial health, taking average health as the reference. We can then see that the short-term effect of job loss is worse when the worker's health is poorer. For example, we see in Figure 6d that being in severe rather than average health adds up to five percentage points to the increase in unemployment risk experienced by treated workers during the five years following job loss. Consistently, the drop in income is up to 14 percentage points larger for workers initially in severe health relative to average health. Interestingly, the effect on job stress content is lower for individuals in bad health. This can be explained by the fact that stressful jobs are more costly to workers in bad health so unemployed workers will be less likely to accept stressful job offers when their health is worse, making unemployment spells longer and job stress upon unemployment exit lower, consistently with both Figures 6d and 6f.

7.2 Health shocks

The value of health for employed workers. We now quantify the utility cost of health shocks, proceeding as we did to quantify the value of employment by computing the values of being in different health states while controlling for individual and job attributes. We define the relative value for an individual of age t and working in job $\mathbf{y} \neq \emptyset$, of being in health state h compared to state h' as $\Delta^{(h)}V(t, \mathbf{y}, h, h') = \mathbf{E}_x [V(x, t, h, \mathbf{y}) - V(x, t, h', \mathbf{y})]$, where V is defined by (2). The value difference $\Delta^{(h)}V$ combines differences in instantaneous utility and differences in continuation values, as health affects future career dynamics. The former can be written as the difference in the disutility of labor: $\Delta^{(h)}c(\ell, s, t, h, h') = c(\ell, s, t, h) - c(\ell, s, t, h')$, where c is defined by (9). Note that since the wage enters the utility function additively and health is a discrete variable in our model, one can interpret $\Delta^{(h)}c$ as a structural marginal willingness to pay i.e. the wage increase required to equalize the instantaneous utility of two workers with different health statuses and otherwise similar characteristics.

While this measure of preferences would indeed be a model primitive (and thus invariant to policy), we may be more interested in the wage compensation needed to equalize value functions across different health states, since welfare is measured by those values in our dynamic model. We then define $MWP^{(h)} = MWP^{(h)}(t, \mathbf{y}, h, h') = \mathbf{E}_x \left[MWP_x^{(h)}(x, \mathbf{y}, \ell, h, h')\right]$, where $MWP_x^{(h)}$ is implicitly defined by $V(x, t, h, \mathbf{y}) = V\left((x, t, h', \left(w \cdot e^{MWP_x^{(h)}(x,t,\mathbf{y},h,h')}, \ell, s\right)\right)$. In words, $MWP^{(h)}$ is the average (in the population) log-wage increase required to equalize the value functions of two workers with different health status and otherwise similar characteristics.

Results are shown in Table 10 for 30 year-old full-time workers in a median-wage job with either a middling or a high stress content. The values shown are the difference between the average and the severe health states. The value attached to being in average (*vs* severe) health is large, over £13,000. In Table 9, we evaluated the value of a median wage, high-stress job (*vs* unemployment) at around £36,000 for a worker in average health. The cost of severe health shocks is therefore substantial, in the order of over a third of the cost of job loss.

Comparing the first and second row of Table 10, we note that only two thirds of the value loss is incurred at the time of the health shock. The remaining third is in the continuation value and is therefore incurred in subsequent years through two channels. The first is persistence in bad health and its related labor disutility cost. As we will see in Figure 7, this effect is short lived. The second channel is triggered when the health shock occurs simultaneously with a labor supply shock, inducing workers to quit into unemployment. Quitting their job harms

Table 10: Value of being in the average *vs* severe health state at age 30

	Employed full-time at the median wage in a...	
	medium-stress job	high-stress job
$\Delta^{(h)}c$	5,540	10,668
$\Delta^{(h)}V$	7,060	13,140
$MWP^{(h)}$	0.062	0.133

Notes: Differences in costs $\Delta^{(h)}c$ and values $\Delta^{(h)}V$ are in GBP.
Medium stress refers to the 3^{rd} (out of 5) job health content.
High stress refers to the highest job health content.

their future income and employment prospects, thus leading to a large drop in continuation value relative to workers who were not hit by the severe health shock.

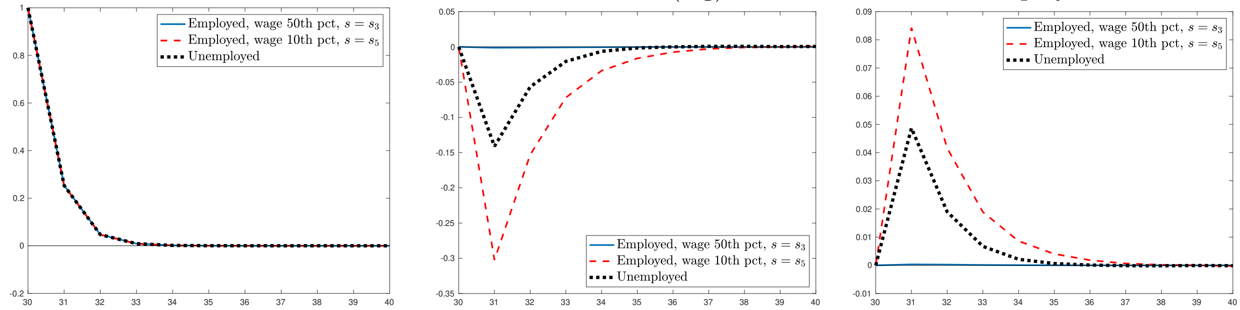
Hence, even though our model does not explicitly account for the (flow) utility cost of being in poor health while unemployed, our estimates of the value of health for workers highlight the importance of taking health shocks into account when studying labor market trajectories. This is further supported by our estimates of workers' marginal willingness to pay for health (see the last row of Table 10): 30 year old workers in average health and employed full time in a median-wage medium-stress (resp. high-stress) job would give up 6% (resp. 13%) of their wage to avoid being hit by a severe mental health shock.

The effect of a severe health shock on labor market and health trajectories. We next evaluate the dynamic effects of a severe health shock at age 30 on future health and labor market paths. The 'control group' now consists of a representative individual in labor market state $\mathbf{y}$ (employment, wage and stress content) at age 30 who is in the average health state. The 'treatment group' consists of the same individual starting out in the severe health state. We use our estimated model to simulate the health, income and unemployment trajectories for these two individuals and show the difference between treated and control groups in Figure 7. Three cases are considered: one where the individual is initially working full time at the median wage in a medium-stress (s_3) job (solid line), one where the individual is initially working full time at a low wage (1^{st} decile) in the most stressful (s_5) job type (dashed line) and one where the individual is initially unemployed (dotted line).

First, we note that it takes about three years for differences in the probability to be in bad (i.e. severe or poor) health to disappear, regardless of the initial employment status. Then, the interaction between health and work which is at the core of our model is further illustrated in Figures 7b-7c. We see that a severe health shock creates income losses that can persist for up

Figure 7: Effects of being in severe vs average health at age 30

7a: Prob. of Severe or Poor health
7b: Income (log)
7c: Unemployment rate



Notes: Difference between the outcomes of individuals in severe vs average health at age 30
Age on horizontal axis.

to six years for workers who are unemployed or working in a stressful and poorly paid job. This is largely explained by the extensive margin of labor supply. Figure 7c shows that workers hit by the health shock are less likely to be employed in the short term. The severe health shock increases the disutility of working, especially in a stressful job, leading some employed workers to quit, especially those in low-paying jobs, and preventing unemployed workers from accepting job offers they would take if they were in better health. Workers who quit their job following the health shock then see their health and labor market outcomes start down the unemployment dynamics shown in the previous subsection, with very persistent income losses and potentially long unemployment duration, especially for workers in bad health (as is the case for treated workers in our exercise).

Summing up, an important result of this analysis is that even though the effect of a severe health shock on individual health has mostly disappeared after four to five years, such a shock can generate persistent income losses and higher unemployment risks for workers who are either unemployed or employed in stressful or poorly paid jobs at the time of the shock. This result complements the ones shown in the previous subsection, providing further evidence of feedback effects between health and labor market outcomes.

7.3 Stressful jobs

The willingness to pay for a non-stressful job. We now look at the workers' valuation of job health contents. Similarly to what we did when analyzing the effects of health or job loss shocks, we define the value for an individual of age t, health h, of working in job $\mathbf{y} = (w, s, \ell)$ relative to job $\mathbf{y}' = (w, s', \ell)$ as $\Delta^{(s)}V(t, h, \mathbf{y}, \mathbf{y}') = \mathbf{E}_x [V(x, t, h, \mathbf{y}) - V(x, t, h, \mathbf{y}')]$, where V is defined by (2) and is averaged over the distribution of individual heterogeneity x.

We can also define $MWP^{(s)} = MWP^{(s)}(t, h, \mathbf{y}, \mathbf{y}') = \mathbf{E}_x \left[MWP_x^{(s)}(x, t, h, \mathbf{y}, \mathbf{y}') \right]$, where $MWP_x^{(s)}$ is implicitly defined by $V(x, t, h, \mathbf{y}) = V\left(x, t, h, \left(w \cdot e^{MWP_x^{(s)}(x,t,h,\mathbf{y},\mathbf{y}')}, s', \ell\right)\right)$. In words, $MWP^{(s)}$ is the average log-wage increase required to equalize the value functions of two workers with different job health contents and otherwise similar characteristics.

We report the value of a medium- *vs* high-stress job in the first row Table 11 for 30-year old full-time workers at the median wage and in each of the four health states. The job health content is very important to workers: going from a medium to a high-stress job decreases the worker's value by a larger amount than a severe health shock (see Table 10). This is partly due to the utility cost of working in a more stressful job but also to the fact that an increase in job stress is a very persistent shock in a labor market with search frictions.

The second row of Table 11 shows the marginal willingness to pay for going from a high- to medium-stress job. These results show that workers have stronger preferences for less stressful jobs when they are in worse health: workers in the severe health state are willing to give up 22% of their wage to reduce the stress content of their job from high to medium. Hence the welfare cost of more stressful jobs is far from negligible, especially for workers in poorer health.

Table 11: Value of a medium- vs high-stress job at age 30

	Health			
	Severe	Poor	Average	Good
$\Delta^{(s)}V$	25,204	21,443	19,015	16,588
$MWP^{(s)}$	0.221	0.192	0.173	0.152

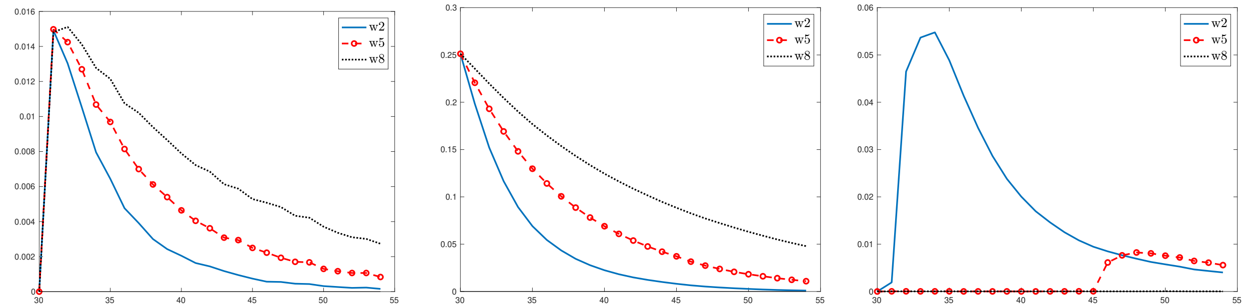
Note: value differences $\Delta^{(s)}V$ *are in GBP.*

The effect of working in a stressful job. We finally turn to the dynamic effects of an increase in job stress content on future health and career trajectories. We consider a representative 30-year old full-time worker at a specific wage level and in the average health state. In the control group, this worker is in a medium-stress job. In the treated group, this worker is in the most stressful type of job. We simulate the health and job health content trajectories for these two groups and show the differences between the corresponding age profiles of the treatment and control groups in Figure 8.

Figure 8a shows that working in a more stressful job raises the probability to be in bad (severe or poor) health. While the magnitude is not large (less than 1.5 percentage points), the effect is persistent so that over time the proportion of workers whose health has deteriorated

Figure 8: Effects of going from a medium to most stressful job type at age 30

8a: Prob. of Severe or Poor health | 8b: Job health content | 8c: Unemployment rate



Notes: In all graphs, workers are in 'average' health at age 30. Age on horizontal axis.
At age 30, workers are employed full time at the 2^{nd}, 5^{th} or 8^{th} wage percentile (resp. w2, w5 and w8).

can become substantial. Interestingly, the persistence of the job stress shock depends positively on the initial wage: 20 years after the shock, the probability of being in bad health is still 0.5% point above that of the control group for workers who were initially at the 8^{th} wage decile, while it is equal to that of the control group for workers who were at the median (or 2^{nd} decile wage).

The heterogeneity in persistence is due to worker selection across jobs, as illustrated by Figure 8b. Compared to the control group (workers in a medium-level stress job), workers in a high-stress job at age 30 will on average spend a long time in more stressful jobs. The effect decreases with age but is quite persistent. As mentioned above, this comes from search frictions. Moreover, the persistence varies positively with the initial wage. Workers who earn a low wage at age 30 and are hit by a high job stress shock tend to go back to less stressful jobs more quickly than workers who initially earned higher wages. This explains the heterogenous responses of health dynamics observed in Figure 8a. Indeed, a high wage may compensate for the disutility of working in a stressful job so workers feel less of an urge to change for a less stressful (and potentially less well paid) job. This may have a negative effect on their future health but this effect is small, discounted as part of the continuation value and the higher wage will partly alleviate the disutility of working while in poor health.

More stressful jobs also affect the extensive margin of labor supply, as illustrated in Figure 8c. When the wage is not too low (at the median or higher quantiles) we see that working in a medium or high-stress job makes no difference to quits or unemployment, at least in the first half of one's career. However, a combination of low wage and high stress does increase the risk of unemployment (see the solid line in Figure 8c). This is because the wage no longer compensates for the higher disutility of working in a high-stress job, leaving workers close to the participation margin and thus likely to quit following an adverse health or labor supply shock. The increase in quits will set workers on a different income path: the average treatment effect on income (not shown here) goes down to -23% after four years. An interesting feature of Figure 8c is that the persistence in job stress leads to a small increase in job quits even at the median wage as the worker gets older. Indeed, labor disutility increases with job stress and also with age, so while workers in high-stress jobs may not be close to the participation margin when in their thirties (provided the wage is high enough), there will come a point when they find it more difficult to stay in these stressful jobs.[19]

8 The changing nature of work

While a large numbers of papers have documented structural changes in the labor market over the last three decades in terms of the skill content of jobs (see Autor and Dorn, 2013 among many others), much less attention has been given to the consequences of these changes on workers' health, particularly through a change in the composition of jobs in terms of stress contents. Recently, though, Kaplan and Schulhofer-Wohl (2018) have highlighted trends in the (dis)utility of work in the US since the 1950s based upon data on workers' perception of their own contentment with their job in a reduced-form analysis. They document a trend in workers' perception of work as being increasingly stressful and less meaningful, especially amongst men.

We can use our structural model to investigate the consequences of such changes on workers' health and labor market outcomes. To this end, we now examine the model's response to changes in the sampling distribution of job health contents. We consider individuals aged 25 in the initial cross section of our data set, use their observed initial labor market and health outcomes as starting values and simulate their trajectories with our estimated model where the distribution of job health content is set either to its estimated value or to a series of counterfactual values.

More precisely, recall that the sampling probability mass function of s is specified in Section 5.3 as $\Pr\{s_j\} = s_j^\alpha \big/ \sum_{k=1}^5 s_k^\alpha$. Hence, the higher α, the larger the proportion of high-stress job offers. We consider a range of values for the parameter α, from -2 (high share of low-stress jobs) to 4 (high share of high-stress jobs), including 0 (uniform distribution) and the estimated value $\widehat{\alpha} = 2.48$. To facilitate interpretation, we identify each counterfactual with the odds of drawing a job offer with the highest stress content relative to a medium-stress job offer (these

[19] Note that while Figures 8c-d seem to suggest a discontinuity at age 45 in labor disutility there is no such discontinuity in our model. We also simulated the treatment effect for workers at the 3^{rd} or 4^{th} wage decile and the increase in quits occurred earlier. The sharp increase observed for the median wage group in Figures 8c-d is thus caused by a 'genuine' increase in quits due to higher labor supply costs for older workers.

odds are a strictly increasing function of the parameter α).

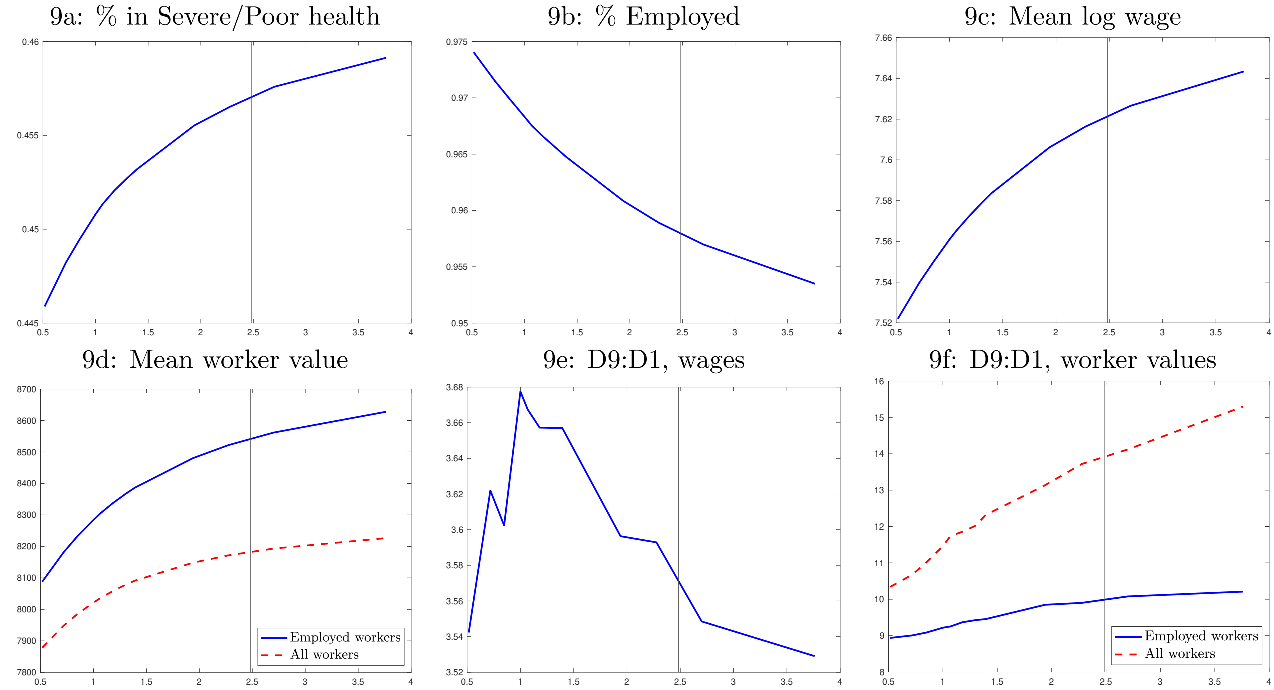
For each job offer distribution considered, we obtain at each age a distribution of individual labor market and health outcomes and can then compute age-specific moments of these distributions. First, we show in Figure 9a the probability of being in bad (i.e. severe or poor) health at age 45 for all counterfactual distributions. The distribution estimated in our data is shown by a vertical line at the corresponding value of the high/medium stress job offer odds.

Figure 9a shows that the share of workers in bad health increases as the distribution of job offers shifts towards higher proportions of stressful jobs. The magnitude of the effect is not large: the share of individuals in bad health increases by roughly one percentage point when going from an odds ratio of 0.5 (half as many high-stress as medium-stress job offers) to an odds ratio of 2 (twice as many high-stress as medium-stress job offers). Hence, even though the increase in the proportion of stressful jobs has an adverse effect on workers' mental health, our counterfactual exercise shows that this channel can cause a deterioration of mental health over a couple of decades, but is unlikely in itself to create a quantitatively large change in the distribution of mental health.

Changing the distribution of health content among job offers also affects labor market outcomes. We see in Figure 9b that the employment rate at age 45 drops by about 2 percentage points when the high to medium stress odds ratio increases from 0.5 to 2. This is intuitive as workers are less likely to accept job offers when they are drawn from a distribution placing more weight on an attribute (job stress) they dislike.

Figure 9c shows that the mean wage among workers who are employed at age 45 increases with the prevalence of high-stress jobs (by around 10 log points when spanning the whole range of counterfactual distributions). In our partial equilibrium analysis, the increase in mean wage results from the combined impact of at least two forces. First, the increase in the likelihood of being offered a high-stress job *for a given wage draw* makes a larger fraction of low-paying jobs unappealing to workers. The increase in the prevalence of high-stress jobs thus weeds out some low-paying jobs, raising the mean wage. Second, compensating differentials also put upward pressure on wages. Recall that wage offers depend on the job stress content through the parameter $\beta^{(s)}$ in equation (8), which we estimated positive and statistically significant (Table 5). Shifting the distribution of job offers toward more stressful jobs thus mechanically causes an increase in wage offers. Interestingly, we see that this increase is not enough to stop the employment rate from falling (Figure 9b), even though the mean wage increases.

Figure 9: Effect of counterfactual job stress distributions on health at age 45



Notes: In all graphs, the horizontal axis displays the odds of drawing a job offer with the highest vs. third highest stress content. The vertical line indicates the value of these odds in the estimated model. The vertical axis shows one of the following outcomes at age 45: the probability of being in severe or poor health (9a), the probability of being employed (9b), the mean log wage (9c), mean worker lifetime values (9d), the D9:D1 (ninth to first decile) ratio of wages (9e), D9:D1 ratios of worker lifetime values (9f)

Figure 9d shows two different profiles of mean worker lifetime values: the mean among employed workers (solid line), and the population-wide mean, including unemployed workers. Both of these profiles are upward sloping. In other words, neither the increase in the incidence of poor health, nor the drop in employment rates are sufficient to counteract the impact of higher wages on worker values as the odds of sampling a high-stress job increase.

The final two Figures, 9e and 9f, summarize the impact of the increased prevalence of high-stress jobs on inequality, both in wages (Figure 9e) and in worker lifetime values (Figure 9f). In both cases, inequality is measured by the D9:D1 (ninth to first decile) ratio. Figure 9e suggests that raising the odds of sampling high-stress jobs had quantitatively very little impact on wage inequality: the D9:D1 ratio profile for wages is non-monotonic, and its range is quantitatively very narrow. This profile results from a complex combination of forces: the change in the dispersion of s in the sampling distribution (which tends to decrease as α increases), the 'weeding out' of low-wage jobs, and the behavioral response of workers to the higher likelihood of sampling high-stress jobs (workers tend to select themselves away from high-stress jobs). Now, interestingly, the solid line on Figure 9f shows that, despite the lack of response of wage inequality, inequality in *lifetime values* among employed workers increases as high-stress jobs become more prevalent.[20] This increase in the dispersion of lifetime values must be driven by the distribution of health among employed workers. Due to the complementarity of job type and health in labor disutility, workers in ill health are those who are the most hurt by a shift of the offer distribution towards higher-stress jobs. Conversely, workers in good health are in a better position to withstand higher levels of job stress, and benefit from the compensating wage increases. The gap in value between workers in good health and those in ill health thus widens.

9 Conclusion

We have constructed a structural model of the joint dynamics of individual careers and mental health trajectories. The model accounts for two-way interactions between work and health and for key features of the labor market, such as search frictions, on-the-job search, and selection of workers into jobs based on wages and the health contents of jobs. Taking the model to British data, we estimate the dynamic response of health to various job attributes, as well as worker selection into jobs as a function of health and job characteristics. We then use our estimated model to quantify the impact of job loss, health shocks, or an increase in job stress on individual health and labor market outcomes, highlighting the way in which the health (resp. employment) channel amplifies the adverse effects of job loss (resp. of a health shock) early on in the career. Motivated by recent evidence of a trend toward more stressful jobs, we further simulate structural changes in the distribution of available jobs and show, amongst other things, that an increase in the prevalence of stressful job offers increases lifetime inequality, hurting workers in ill health and benefitting workers in good health.

Our model can be extended to address specific issues. For example, we could shift the focus to older workers and add more structure on the extensive employment margin (modelling pensions and savings) to study how an individual's employment, wage but also occupation history affects their late-career health and retirement decisions. This would extend the work of French (2005), French and Jones, 2011, and Salvati (2020), and could prove useful for the design of pension systems that include occupational health contents as a factor impacting eligibility.[21]

Another important extension is to study the interaction between health and work for women.

[20] That increase is obviously magnified when inequality is measured in the whole population (dashed line of Figure 9f), as the share of unemployed workers, who have very low values, increases (Figure 9b).

[21] For instance, such a reform was implemented in France in 2015 ("*compte pénibilité*"). The number of years spent in 'strenuous' jobs can bring forward the age of full pension eligibility.

We focus on men in the present paper because a serious analysis of the health and careers of women would have to model fertility decisions and their impact on both health and career outcomes. One of our counterfactual results showed that men in low-wage, stressful jobs are especially vulnerable to health shocks as they can trigger an unemployment spell, with lasting effects on income and employment, which then slow down health recovery. We think these mechanisms may be even more salient for women as fertility may act as an additional propagation channel for health or employment shocks. The approach used here could thus inform the joint design of unemployment, parental leave and health insurance policies.

Finally, the model in this paper is decidedly partial equilibrium: both the job offer probabilities and the wage offer function are taken as exogenous. This is a clear limitation, particularly for the 'long-run' counterfactual analysis of structural changes in Section 8. Predicting the response of wage offers, particularly compensating differentials, to a technological shift toward more stressful jobs in an equilibrium model is far from straightforward. To be credible, such a prediction would have to be based upon very rich data conveying information on firms (in addition to the health and work variables we are already using) over a time window of several decades, to capture structural changes over time. Until such data becomes available, the results from our counterfactual analysis offer a first attempt at informing policy on the comovements of job health contents, individual health and labor market outcomes. Note that our estimates of the health process in Table 4 ascribed a relatively minor role to wages. It is therefore not unreasonable to expect that our results on the response of mental health to changes in the sampling distribution of job stress contents will remain broadly valid in an equilibrium analysis where compensating wage differentials are endogenous.

References

Autor, D. and D. Dorn (2013): "A Distributional Framework for Matched Employer Employee Data," *American Economic Review*, 103, 1553–97.

Bonhomme, S., T. Lamadon, and E. Manresa (2019a): "Discretizing Unobserved Heterogeneity," mimeo, University of Chicago.

——— (2019b): "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 87, 699–38.

Bonhomme, S. and E. Manresa (2015): "Grouped Patterns of Heterogeneity in Panel Data," *Econometrica*, 83, 1147–84.

Brusco, M. and D. Steinley (2007): "A Comparison of Heuristic Procedures for Minimum Within-Cluster Sums of Squares Partitioning," *Psychometrika*, 72, 583–600.

Burdett, K., C. Carillo-Tudela, and M. Coles (2018): "The Cost of Job Loss," mimeo, University of Essex.

Capatina, E., M. Keane, and S. Maruyama (2020): "Health Shocks and the Evolution of Earnings over the Life-Cycle," mimeo, University of New South Wales.

Currie, J. and B. Madrian (1999): "Health, Health Insurance and the Labor Market," in *Handbook of Labor Economics*, ed. by O. Ashenfelter and D. Card, Elsevier Science, chap. 50, 3309–416.

Davis, S. and T. Von Wachter (2011): "Recessions and the Cost of Job Loss," *Brookings Papers on Economic Activity*, 43, 1–72.

de Nardi, M., S. Pashchenko, and P. Porapakkarm (2017): "The Lifetime Costs of Bad Health," NBER working paper no. 23963, National Bureau of Economic Research.

French, E. (2005): "The Effects of Health, Wealth, and Wages on Labour Supply and Retirement Behavior," *Review of Economic Studies*, 72, 395–427.

French, E. and J. Jones (2011): "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," *Econometrica*, 79, 693–732.

Fryers, T., D. Melzer, and R. Jenkins (2003): "Social Inequalities and the Common Mental Disorders," *Social Psychiatry and Psychiatric Epidemiology*, 38, 229–37.

Gill, S., P. Butterworth, B. Rodgers, and A. Mackinnon (2007): "Validity of the Mental Health Component Scale of the 12-Item Short-Form Health Survey (MCS-12) as Measure of Common Mental Disorders in the General Population," *Psychiatry Research*, 152, 63–71.

Gilleskie, D. (1998): "A Dynamic Stochastic Model of Medical Care Use and Work Absence," *Econometrica*, 66, 1–45.

Gruber, J. and B. Madrian (2004): *Health Insurance, Labor Force Participation, and Job Choice: A Critical Review of theLiterature*, Urban Institute Press.

Harris, M. (2019): "The Impact of Body Weight on Occupation Mobility and Career Development," *International Economic Review*, 60, 631–60.

Jacobs, L. and S. Piyapromdee (2016): "Labor Force Transitions at Older Ages: Burnout, Recovery, and Reverse Retirement," Finance and Economics Discussion Series 2016-053, Washington: Board of Governors of the Federal Reserve System.

Jarosch, G. (2015): "Searching for Job Security and the Consequences of Job Loss," mimeo, University of Chicago.

Kaplan, G. and S. Schulhofer-Wohl (2018): "The Changing (Dis-)Utility of Work," *Journal of Economic Perspectives*, 32, 239–58.

Layard, R. (2017): "The Economics of Mental Health," *IZA World of Labor*.

Lise, J. and J.-M. Robin (2017): "The Macro-dynamics of Sorting between Workers and Firms," *American Economic Review*, 107, 1104–35.

Low, H., C. Meghir, and L. Pistaferri (2010): "Wage Risk and Employment Risk over the Life Cycle," *American Economic Review*, 100, 1432–67.

Michaud, A. and D. Wiczer (2018a): "The Disability Option: Labor Market Dynamics with Economic and Health Risks," mimeo, University of Western Ontario.

——— (2018b): "Occupational hazards and social disability insurance," *Journal of Monetary Economics*, 96, 77–92.

OECD (2015): "Fit Mind, Fit Job: From Evidence to Practice in Mental Health and Work," in *Mental Health and Work*, OECD Publishing, Paris.

Papageorge, N. (2016): "Why Medical Innovation is Valuable: Health, Human Capital, and the Labor Market," *Quantitative Economics*, 7, 671–725.

Salvati, F. (2020): "Female Health and Labor Supply at Older Ages," mimeo, University College London.

Salyers, M., H. Bosworth, J. Swanson, J. Lambpagone, and F. Osher (2000): "Reliability and Validity of the SF-12 Health Survey among People with Severe Mental Illness," *Medical Care*, 38, 1141–50.

Sullivan, D. and T. Von Wachter (2009): "Job Displacement and Mortality: An Analysis Using Administrative Data," *Quarterly Journal of Economics*, 124, 1265–306.

Tran, T. (2017): "Too Stressed to Work: The Effects of Job Stressors on Health and Employment Outcomes," mimeo, University of North Carolina at Chapel Hill.

Ware, J., M. Kosinski, and S. Keller (1996): "A 12-Item Short-Form Health Survey: Construction of Scales and Preliminary Tests of Reliability and Validity," *Medical Care*, 34, 220–33.

WHO (2019): "Special Initiative for Mental Health (2019-2023)," *World Health Organization*, WHO/MSD/19.1.

APPENDIX

A Health variables from the Short-Form Survey: presentation and robustness checks

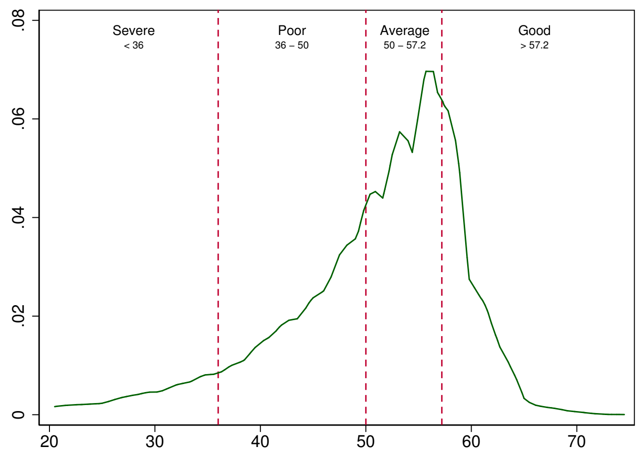
The SF-12 consists of respondents' replies to the twelve questions in Table 12. Answers to these questions are then processed to produce two continuous health indicators: sf12mcs for mental health and sf12pcs for physical health. The former (mental) health indicator is the variable we use in our analysis to assign individuals to one of the four health states (see Section 3.1 for details). As we mention in the main text, the sf12mcs indicator, has been successfully tested in several studies in the medical literature against other diagnosis measures of mental health (see Ware et al., 1996, Salyers et al., 2000 and Gill et al., 2007) and is thus considered to be a reliable variable to assess individuals' health.

Table 12: SF-12 Questionnaire

1	How is your health in general?	7	Did your mental health mean you work less carefully?
2	Does your health limit moderate activities?	8	Did the pain interfere with your work?
3	Does your health limit walking up flights of stairs?	9	Did you feel calm and peaceful?
4	Did your physical health limit the amount of work you do?	10	Did you have a lot of energy?
5	Did your physical health limit the kind of work you do?	11	Did you feel downhearted and depressed?
6	Did your mental health mean you accomplish less?	12	Did you health interfere with your social life?

In our analysis, we discretize sf12mcs indicator for mental health and consider four states states: Good, Average, Poor and Severe. We show in Figure 10 the distribution of the mental health score (in the initial cross section) and the cut-offs for the four health states.

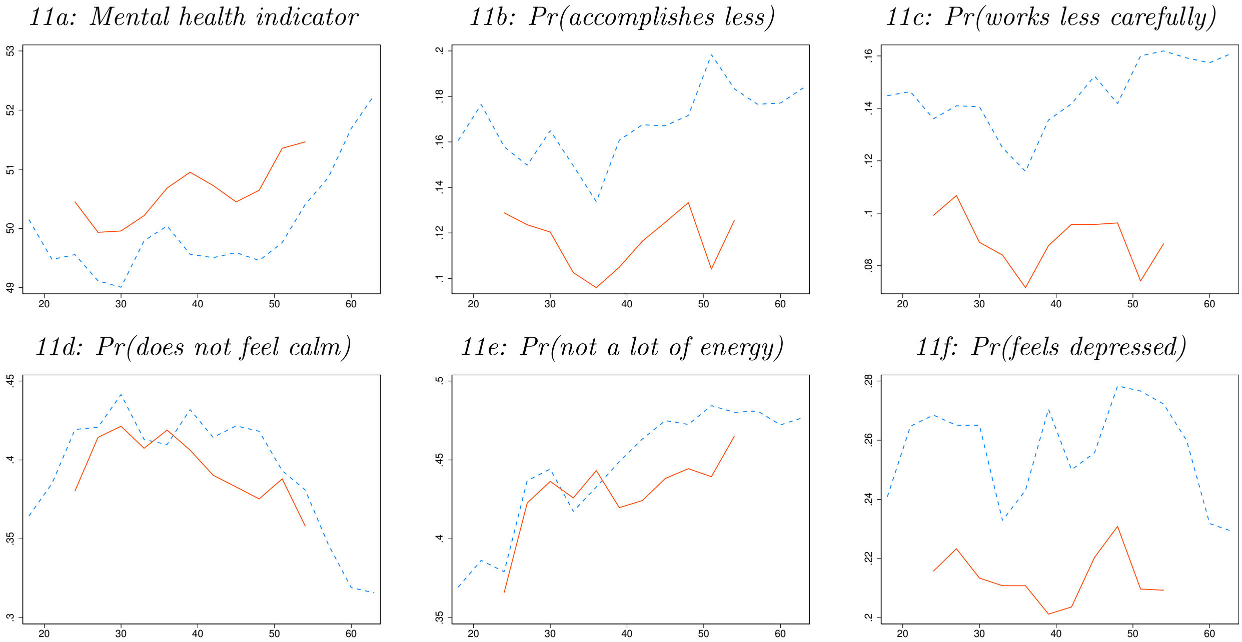
Figure 10: SF-12 Mental health score and health categories



x-axis: SF-12 mental health score, y-axis: score density in initial cross section

To give more insights into the age profiles of mental health shown in the main text in Figure 2a and to check the robustness of those patterns with respect to the selection of our work sample, we plot in Figure 11 several variables from the SF-12 questionnaire against age. For each each variable, we show the average level by age, where averages are taken in our work sample or in the raw UKLHS data (before any selection takes place, except for gender). The first variable, in Figure 11, is the sf12mcs indicator that is the variable underlying the four health categories in our analysis. Figures 11b-11f illustrate answers to questions 6, 7, 9-11 in the SF-12 questionnaire (see Table 12). For each of these questions, the individual answer is set to 1 if individuals say that they “accomplish less”, “work less carefully”, etc. some of the time, most of the time or all the time. The graphs then show the average of these dummies over male workers in our work sample (solid line) or over of the whole population of men in the original UKLHS data (dashed line). Accordingly, the work sample averages only cover ages 25 to 55 while the original data sample averages go from age 18 to 65.

Figure 11: Age profiles of several health indicators



Age (in years) on the horizontal axis. Solid line: work sample. Dashed line: original data sample

Our first observation is that the age profiles are qualitatively similar whether we consider our work sample or the original UKLHS data. We do not expect the two patterns to be superimposed as the two populations are different given the selection criteria used in the construction of our sample (see Section 3.1) and indeed the averages differ across samples (by only a few percentage points) for some of the variables. However, the selection applied to the original data sample does not seem to create age profiles that were not already in the data.

The second observation is that there seems to be a small increase in several indicators of depression between ages 40 and 50 (in Figures, 11b, 11c, 11e and 11f). These patterns are captured by the summary indicator sf12mcs used in our analysis, as Figure 11a shows a dip for men in their forties. This small decrease of the mental health indicator is however not obvious in the age profiles of the four health categories shown in Figure 2a, thus meaning that the deterioration in mental health between 40 and 50 is not large enough to make a substantial proportion of men enter the ’severe’ or ’poor’ mental states, where these states are defined using cut offs for sf12mcs from the medical literature (see Section 3.1).

B Algorithm for the classification step

We present a simple iterative algorithm based on the one that has been developed by Bonhomme and Manresa (2015) to implement their group fixed effect approach. Since our setting shares many features with theirs, we can use a similar, though slightly modified, algorithm for our classification step. The number of classes is set to $K \geq 1$.

Iterative Algorithm:

1. Set $s = 0$ and let $\left\{k_i^0\right\}$ be an initial partition.
2. In each class $k^{(s)}$ with at least two elements, regress by OLS each of the M individual moments on a constant, T and T^2. Set the $\mathbf{g}^{(s)}$ function parameters for this class to the resulting estimates. In each class with only one individual, set the $\mathbf{g}^{(s)}$ slope parameters to zero and the constant to the value of the individual's moment.
3. If at least one class is empty:
   (a) Compute the "distance" $\left\|\mathbf{m}_i - \mathbf{g}^{(s)}\left(k_i^{(s)}, T_i^0\right)\right\|^2$ for each observation.
   (b) For each $k \in \{1, ..., K\}$, if class $k^{(s)}$ is empty, assign to $k^{(s)}$ the observation with the highest distance and re-set this distance to 0.
   (c) Do step 2 and replace the $\mathbf{g}^{(s)}$ function parameters with the new values.
4. Update the partition as follows: $k_i^{(s+1)} = \underset{k=1,...,K}{\text{argmin}} \left\|\mathbf{m}_i - \mathbf{g}^{(s)}\left(k, T_i^0\right)\right\|^2$.
5. If no observations changes class at Step 4 then stop. Otherwise, set $s = s + 1$ and go to Step 2.

This algorithm is a variant of one of the first heuristic clustering procedures, referred to as H-means algorithms. More precisely, a H-means algorithm would not include Step 3 and could thus terminate with empty classes. The addition of Step 3 overcomes this issue and leads to what is sometimes labelled a H-means+ algorithm (see Brusco and Steinley, 2007, for a description of several clustering algorithms).

C Proof of identification

We start by proving the equality 10 used to identify the probabilities Pr(FT|FT) and Pr(FT|PT) directly from the data. Consider the probability to go from full time to part time conditionally on (x, t, h', w, s). We can write this probability as:

$$\begin{aligned}
&\Pr\left\{\ell' = \text{PT} \mid x, t, h', \mathbf{y} = (w, s, \text{FT})\right\} \\
&\qquad = (1-\delta)\left(1 - \Pr\{\text{FT}|\text{FT}\}\right)\Bigg[(1-\lambda_1) H\left[V\left(x, t+1, h', \mathbf{y}' = (w, s, \text{PT})\right) - U(x, t+1, h')\right] \\
&\quad + \lambda_1 \Pr_{(\varepsilon', \mathbf{y}^o)}\left\{V\left(x, t+1, h', \mathbf{y}' = (w, s, \text{PT})\right) - \varepsilon' \geq \max\left\langle V(x, t+1, h', \mathbf{y}^o) - \varepsilon', U(x, t+1, h')\right\rangle\right\}\Bigg].
\end{aligned}$$

Similarly, the conditional probability to stay in part time within the same job is:

$$\begin{aligned}
&\Pr\{\ell' = \mathrm{PT} \mid x, t, h', \mathbf{y} = (w, s, \mathrm{PT})\} \\
&\quad = (1-\delta)\left(1 - \Pr\{\mathrm{FT}|\mathrm{PT}\}\right)\Big[(1-\lambda_1) H\left[V\left(x, t+1, h', \mathbf{y}' = (w, s, \mathrm{PT})\right) - U(x, t+1, h')\right] \\
&+ \lambda_1 \Pr_{(\varepsilon', \mathbf{y}^o)}\left\{V\left(x, t+1, h', \mathbf{y}' = (w, s, \mathrm{PT})\right) - \varepsilon' \geq \max\left\langle V(x, t+1, h', \mathbf{y}^o) - \varepsilon', U(x, t+1, h')\right\rangle\right\}\Big].
\end{aligned}$$

The ratio of these two (observed) probabilities equals $\left(1 - \Pr\{\mathrm{FT}|\mathrm{FT}\}\right) / \left(1 - \Pr\{\mathrm{FT}|\mathrm{PT}\}\right)$. This proves equation 10

We now prove that the parameters obtained in the last step of our estimation procedure, through simulated moments, are identified from the data (cf. Section 5.4). First we show that the arrival rates and distribution of job offers are identified, then we use Bellman equations to uncover a known mapping between the value and instantaneous utility functions and lastly we combine these two results to identify the preference parameters.

We begin with the arrival rates and distribution of job offers. The conditional joint probability of leaving unemployment (U2J) for a job $\mathbf{y}^o$ is given by:[22]

$$\Pr\left(\mathrm{U2J}, \mathbf{y}^o | x, t, h'\right) = \left(1 - \widetilde{\delta}\right) \lambda_0 \Pr\left(\mathbf{y}^o\right) \cdot H\left[V(x, t+1, h', \mathbf{y}^o) - U(x, t+1, h')\right]. \tag{12}$$

Likewise, we can write the conditional probability of making a job-to-job transition (J2J) from a type-$\mathbf{y} = (w, s, \ell)$ job towards a type-$\mathbf{y}^o$ job as:

$$\begin{aligned}
\Pr\left(\mathrm{J2J}, \mathbf{y}^o | x, t, h', \mathbf{y} = (w, s, \ell)\right) &= (1-\delta)\lambda_1 \Pr\left(\mathbf{y}^o\right) \cdot H\left[V(x, t+1, h', \mathbf{y}^o) - U(x, t+1, h')\right] \\
&\cdot \Pr_{\ell'|\ell}\left\{V(x, t+1, h', \mathbf{y}^o) > V\left(x, t+1, h', \mathbf{y}' = (w, s, \ell')\right)\right\}, \quad (13)
\end{aligned}$$

where the last term is the probability that job $\mathbf{y}^o$ is preferred to job $\mathbf{y}$ but the new full-/part-time status ℓ' in this latter job is random and unobserved. Note that the last probability does not depend on the labor supply shock: being additive, that shock does not affect preferences across jobs.

Taking the ratio between (12) and (13), we get:

$$\frac{\Pr\left(\mathrm{J2J}, \mathbf{y}^o | x, t, h', \mathbf{y}\right)}{\Pr\left(\mathrm{U2J}, \mathbf{y}^o | x, t, h'\right)} = \frac{\lambda_1}{\lambda_0} \cdot \frac{1-\delta}{1-\widetilde{\delta}} \cdot \Pr_{\ell'|\ell}\left\{V(x, t+1, h', \mathbf{y}^o) > V\left(x, t+1, h', \mathbf{y}' = (w, s, \ell')\right)\right\} \tag{14}$$

An implicit assumption here is that the denominator is not zero. What follows holds if for any job $\mathbf{y}^o$ there is at least one type of worker (x, t, h) for whom this probability is not zero. This would hold in an equilibrium model as no firm would post an offer to which unemployment is preferred by all workers.

We now assume that there is at least one pair of jobs (say A and B) and one worker (of a certain type, age, and health state) such that the individual prefers job A over job B, regardless of working hours in job B (that is, the individual prefers job A over both the part-time and full-time versions of job B). Formally, there is at least one triple $(\mathbf{y}^o, w, s)$ and one value of (x, t, h') such that $V\left(x, t+1, h', \mathbf{y}^o\right) > \max\left\langle V\left(x, t+1, h', \mathbf{y} = (w, s, \mathrm{PT})\right), V\left(x, t+1, h', \mathbf{y} = (w, s, \mathrm{FT})\right)\right\rangle$. This is a very weak assumption as it is very likely that at least one worker will prefer a highly paid non stressful job to a badly paid stressful one (part time or full time). If this holds, we can take the maximum of the left-hand side of (14)

[22] Note that we condition on the newly realized health h' rather than on health at the start of the period h as h' is observed, realized independently of the events occurring between t and $t+1$ and h becomes irrelevant once h' is known. Also, we drop the conditioning on not having a job at date t: $\mathbf{y} = \emptyset$.

(observed in the data) over $(\mathbf{y}^o, \mathbf{y}, x, t, h')$ to obtain the ratio λ_1/λ_0 (recall that the δ and $\widetilde{\delta}$ parameters were already identified).

Once we know λ_1/λ_0 we can compute $\mathbf{1}\{V(x,t+1,h',\mathbf{y}^o) > V(x,t+1,h',\mathbf{y})\}$ for any pair of jobs and any worker by using equation (14) as follows:

$$\begin{aligned}
&\frac{\Pr(\text{J2J}, \mathbf{y}^o | x, t, h', \mathbf{y})}{\Pr(\text{U2J}, \mathbf{y}^o | x, t, h')} \cdot \frac{\lambda_0}{\lambda_1} \cdot \frac{1-\widetilde{\delta}}{1-\delta} \\
&\quad = \mathbf{1}\{V(x,t+1,h',\mathbf{y}^o) > \max \langle V(x,t+1,h',\mathbf{y}=(w,s,\text{PT})), V(x,t+1,h',\mathbf{y}=(w,s,\text{FT}))\rangle\} \\
&\quad + 0 \times \mathbf{1}\{V(x,t+1,h',\mathbf{y}^o) \leq \min \langle V(x,t+1,h',\mathbf{y}=(w,s,\text{PT})), V(x,t+1,h',\mathbf{y}=(w,s,\text{FT}))\rangle\} \\
&\quad + \Pr(\text{PT}|\ell) \times \mathbf{1}\{V(x,t+1,h',\mathbf{y}=(w,s,\text{PT})) < V(x,t+1,h',\mathbf{y}^o) < V(x,t+1,h',\mathbf{y}=(w,s,\text{FT}))\} \\
&\quad + \Pr(\text{FT}|\ell) \times \mathbf{1}\{V(x,t+1,h',\mathbf{y}=(w,s,\text{FT})) < V(x,t+1,h',\mathbf{y}^o) < V(x,t+1,h',\mathbf{y}=(w,s,\text{PT}))\}.
\end{aligned}$$

The above equation states that the observed left-hand side ratio can only take one of four values (1, 0, $\Pr(\text{PT}|\ell)$ or $\Pr(\text{FT}|\ell)$) depending on the individual's preference between job $\mathbf{y}^o$ and $\mathbf{y} = (w, s, \text{FT})$ or $\mathbf{y} = (w, s, \text{PT})$. We can then use this equation to get the sign of $V(x,t,h,\mathbf{y}^o) - V(x,t,h,\mathbf{y})$ for any individual (x, t, h) and any pair of jobs $(\mathbf{y}, \mathbf{y}^o)$. Note that this slightly convoluted way of obtaining individual rankings of jobs is due to the fact that the full-/part-time status can change within job. This prevents us from directly obtaining the rankings from job-to-job transitions as we do not know whether the previous job was full or part time when the decision to change was taken.

Lastly consider the probability to stay in a job $\mathbf{y}$, where the new full-/part-time status is realized (and observed for job stayers):

$$\begin{aligned}
\Pr(\text{stay}|x,t,h',\mathbf{y}) &= (1-\delta) \cdot H\left[V(x,t+1,h',\mathbf{y}) - U(x,t+1,h')\right] \\
&\times \left[1 - \lambda_1 \sum_{\mathbf{y}^o} \Pr(\mathbf{y}^o)\, \mathbf{1}\{V(x,t+1,h',\mathbf{y}^o) > V(x,t+1,h',\mathbf{y})\}\right]. \qquad (15)
\end{aligned}$$

Using (12) to substitute for the $H[\cdot]$ term on the right-hand side of (15), we can re-write (15) as:

$$\frac{\Pr(\text{stay}|x,t,h',\mathbf{y}) \cdot \Pr(\mathbf{y}) \cdot \left(1-\widetilde{\delta}\right)}{\Pr(\text{U2J}, \mathbf{y}|x,t,h') \cdot (1-\delta)} = \frac{1}{\lambda_0} - \frac{\lambda_1}{\lambda_0} \sum_{\mathbf{y}^o} \Pr(\mathbf{y}^o)\, \mathbf{1}\{V(x,t+1,h',\mathbf{y}^o) > V(x,t+1,h',\mathbf{y})\}. \qquad (16)$$

All the terms in equation (16) are known except for $1/\lambda_0$ and $\Pr(\mathbf{y})$ for any point on the discrete support of the job offer distribution. We can write equation (16) for each value of $\mathbf{y}$, x, t or h'. Doing so yields a linear system with many more equations than unknowns so, should transition probabilities show enough variations, we can identify λ_0 and $\Pr(\mathbf{y})$ for any value of $\mathbf{y}$.[23] We then get λ_1 as we already know the ratio λ_1/λ_0. Importantly for what follows, once we know λ_0 and the offer distribution, we identify the selection probability $H\left[V(x,t+1,h',\mathbf{y}') - U(x,t+1,h')\right]$ from (12) for any job $\mathbf{y}$ and worker (x, t, h).

We now turn to the preference parameters in the instantaneous utility function. First, we re-write

[23]Note that we do not need to exclude (x, t, h) from the offer distribution to identify the offer distribution but doing so gives us more variation in transition probabilities and thus helps improve identification in practice.

the Bellman equations (1):

$$U(x,t,h) = b + (1+r)^{-1}\Bigg\{\sum_{h'} \Pr\left(h', \mathbf{y}' = \emptyset | x, t, h, \mathbf{y} = \emptyset\right) \cdot U(x, t+1, h') \\ + \sum_{h', \mathbf{y}^o} \int_{\varepsilon'} \Pr\left(h', \mathbf{y}' = \mathbf{y}^o, \varepsilon' | x, t, h, \mathbf{y} = \emptyset\right) \cdot \left[V(x, t+1, h', \mathbf{y}^o) - \varepsilon'\right] dH(\varepsilon')\Bigg\} \quad (17)$$

and (2):

$$V(x,t,h,\mathbf{y}) = u(x,t,h,\mathbf{y}) + (1+r)^{-1}\Bigg\{\sum_{h'} \Pr\left(h', \mathbf{y}' = \emptyset | x, t, h, \mathbf{y}\right) \cdot U(x, t+1, h') \\ + \sum_{h', \mathbf{y}^o} \int_{\varepsilon'} \Pr\left(h', \mathbf{y}' = \mathbf{y}^o, \varepsilon' | x, t, h, \mathbf{y}\right) \cdot \left[V(x, t+1, h', \mathbf{y}^o) - \varepsilon'\right] dH(\varepsilon')\Bigg\}. \quad (18)$$

The last line of (17) can be rewritten as follows:

$$\sum_{h', \mathbf{y}^o} \Pr\left(h', \mathbf{y}' = \mathbf{y}^o | t, h, \mathbf{y} = \emptyset\right) \cdot \left[V(t+1, h', \mathbf{y}^o) - \mathbf{E}\left(\varepsilon' | h', \mathbf{y}' = \mathbf{y}^o, t, \mathbf{y} = \emptyset\right)\right] \quad (19)$$

The last term, $\mathbf{E}\left(\varepsilon' | h', \mathbf{y}' = \mathbf{y}^o, t, \mathbf{y} = \emptyset\right)$, pertains to selection over the unobserved labour supply shock and can be re-written as $\mathbf{E}\left[\varepsilon' | \varepsilon' \leq V(x, t+1, h', \mathbf{y}^0) - U(x, t+1, h')\right]$.

Recall that we can identify $H\left(V(x,t,h',\mathbf{y}) - U(x,t,h')\right)$ from (12). We know that ε follows a log-normal distribution so its cdf is invertible. Assuming that we know its standard error σ,[24] we can then identify the function $(\mathbf{y}, x, t, h') \mapsto V(x,t,h',\mathbf{y}) - U(x,t,h')$ over the set of $(\mathbf{y}, x, t, h')$ such that $V(x,t,h',\mathbf{y}) > U(x,t,h')$. That is enough to identify the conditional expectation in (19) from the data (as whenever $V(x,t,h',\mathbf{y}) \leq U(x,t,h')$, the probability of observing that job being taken up, $\Pr\left(h', \mathbf{y}' = \mathbf{y}^o | t, h, \mathbf{y} = \emptyset\right)$, is zero).

Once we know the conditional expectation in (19), we have identified all the terms in equation (17) except for the instantaneous utility and value functions. We can then write b as a weighted sum of the value functions $U(x,\cdot,\cdot)$ and $V(x,\cdot,\cdot,\cdot)$ taken at different dates (t and $t+1$), health status and job characteristics. From what we have just shown, these weights are identified from the data.

We can proceed similarly for the employed value function in (18). The only difference with what we have done above is that the expectation of the labour supply shock is now conditional on being employed in a given job at t (rather than being unemployed). However, since the labour supply shock is additive, knowing that the worker is in job $\mathbf{y}'$ at $t+1$ and was in $\mathbf{y}$ at t only yields information on the shock through the fact that $\mathbf{y}'$ is preferred to unemployment. Knowing that $\mathbf{y}'$ is potentially better than $\mathbf{y}$ (in the case of a job-to-job transition) does not tell us anything about the labour supply shock. Hence the expectation of the labour supply shock in (18) is also conditional on $\varepsilon' \leq V(x,t+1,h',\mathbf{y}^0) - U(x,t+1,h')$ so we can identify it the same way we did the one in (19).

To summarize, let us stack all values $U(x,t,h)$ and $V(x,t,h,\mathbf{y})$ over the discrete support of $(x,t,h,\mathbf{y})$ in a column vector denoted as $\mathbf{C}_{val}$. Likewise, we stack all instantaneous utilities $b(x,t,h)$ and $u(x,t,h,\mathbf{y})$

[24] Proving the identification of the standard deviation σ of the labor supply shock is not straightforward however (note that its mean has been normalised to $-\sigma^2/2$ in section 5.3). In a simpler model where decisions are based on instantaneous utilities, identification would directly follow from the normalized coefficient of the wage in the utility function. With expected value functions however, the point is difficult to prove formally so we carry on assuming that σ is known.

in a column vector $\mathbf{C}_{util}$. We can then rewrite the system of equations (17)-(18) as:

$$\mathbf{C}_{util} = \mathbf{P} \cdot \mathbf{C}_{val} + \mathbf{B}, \tag{20}$$

where $\mathbf{P}$ is a square matrix with coefficients equal to 0, -1 or combinations of transition probabilities $\Pr(h', \mathbf{y}'|x, t, h, \mathbf{y})$ (and the discount factor which we assume to be known) and $\mathbf{B}$ is a known vector of (combinations of) the conditional expectations of the labour supply shock. Note that $\mathbf{P}$ is block diagonal as individuals always stay in the same x group and, for a given instantaneous utility at date t, the weight put on value functions at dates other than t or $t+1$ is zero. Importantly, $\mathbf{P}$ and $\mathbf{B}$ can be directly estimated using data on health and labour market transitions.

Then, if $\mathbf{P}$ is invertible equation (20) means we can write the value function at each $(x, t, h, \mathbf{y})$ as an affine function of instantaneous utilities thus allowing identification of the structural preference parameters under our model specification. Indeed, let $b(x, t, h) = b$ and $u(x, t, h, \mathbf{y}) = w - h(\alpha s + \beta t)$, where b, α and β are the parameters to identify. Recall that a job $\mathbf{y}$ is a triplet (w, s, ℓ). Multiplying both sides of (20) by $\mathbf{P}^{-1}$ leads to:

$$U(x,h,t) = \sum_{\mathbf{y}',h',t'} \mu\left(h',t',\mathbf{y}',x,h,t,\emptyset\right) \cdot \left[w' - h'(\alpha s' + \beta t')\right] + \sum_{h',t'} \mu\left(h',t',\emptyset,x,h,t,\emptyset\right) \cdot b + \psi(x,h,t,\emptyset),$$

$$V(x,h,t,\mathbf{y}) = \sum_{\mathbf{y}',h',t'} \mu\left(h',t',\mathbf{y}',x,h,t,\mathbf{y}\right) \cdot \left[w' - h'(\alpha s' + \beta t')\right] + \sum_{h',t'} \mu\left(h',t',\emptyset,x,h,t,\mathbf{y}\right) \cdot b + \psi(x,h,t,\mathbf{y}),$$

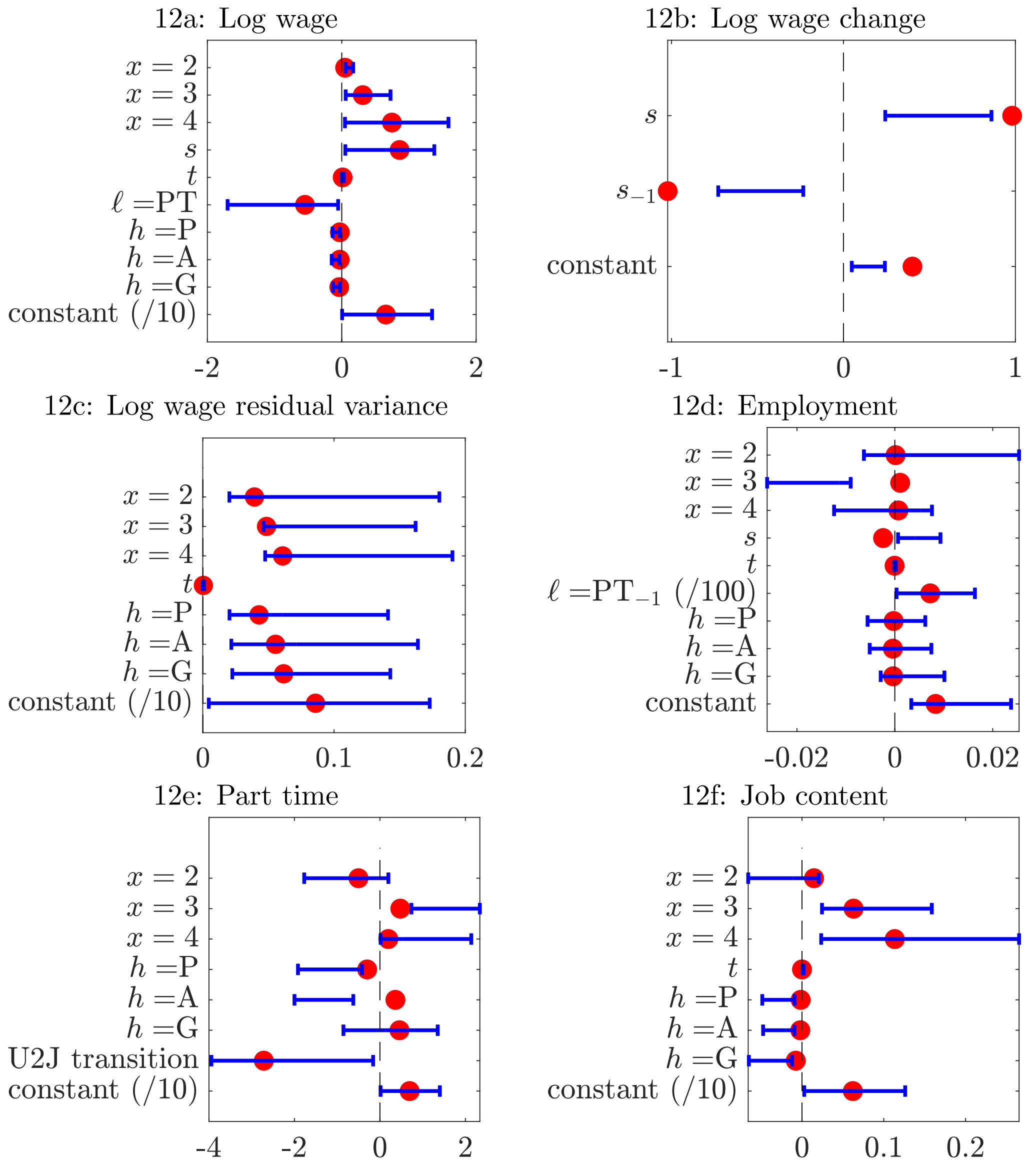
where all $\mu(\cdot)$ and $\psi(\cdot)$ are identified from the data. The difference between the value of a job and that of unemployment can then be written as:

$$V(x,h,t,\mathbf{y}) - U(x,h,t) = \alpha \cdot g_\alpha\left(x,t,h,\mathbf{y}\right) + \beta \cdot g_\beta\left(x,t,h,\mathbf{y}\right) + b \cdot g_b\left(x,t,h,\mathbf{y}\right) + g_\alpha\left(x,t,h,\mathbf{y}\right),$$

where $g_\alpha(\cdot)$, $g_\beta(\cdot)$, $g_b(\cdot)$ and $g(\cdot)$ are known from the data. Taking this into the likelihood of (for instance) leaving unemployment for a given job (12) we can then estimate the parameters of interest α, β and b.

D Additional estimation results

Figure 12: Model fit - Auxiliary regressions



Notes: Red dots: Model-produced moments
Blue line segments: 95% bootstrap confidence interval of data moments.