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 A1. 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 2.1 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 A1: SF-12 Questionnaire

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

In our analysis, we discretize sf12mcs indicator for mental health and consider four states: Good, Average, Poor and Severe. We show in Figure A1 the distribution of the mental health score (in the initial cross section) and the cut-offs for the four health states.
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 A2 several variables from the SF-12 questionnaire against age. For 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 A2, is the sf12mcs indicator that is the variable underlying the four health categories in our analysis. Figures A2b-A2f illustrate answers to questions 6, 7, 9-11 in the SF-12 questionnaire (see Table A1). 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.
Our first observation is that the 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 2.1) 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 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 A2b, A2c, A2e and A2f). These patterns are captured by the summary indicator sf12mcs used in our analysis, as Figure A2a shows a dip for men in their forties. This small decrease of the mental health indicator is however not obvious 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’ mental state, where this state is defined using cut offs for sf12mcs from the medical literature (see Section 2.1).

### 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 $\{k_0^i\}$ 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 $g^{(s)}$ function parameters for this class to the
resulting estimates. In each class with only one individual, set the $g^{(s)}$ 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” $\left\| m_i - g^{(s)} \left( k^{(s)}_i, T^0_i \right) \right\|^2$ for each observation.
   (b) For each $k \in \{1, \ldots, K\}$, 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 $g^{(s)}$ function parameters with the new values.

4. Update the partition as follows: $k^{(s+1)}_i = \arg\min_{k=1,\ldots,K} \left\| m_i - g^{(s)} \left( k, T^0_i \right) \right\|^2$.

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**

**C.1 Parameters estimated directly from the data**

The coefficients $\beta^{(x)}$ and $\beta^{(s)}$ in [3] can be estimated directly from the data. Define $\bar{w}_{\text{net}}(x, s)$
as the maximum observed wage amongst all matches involving a type-$x$ worker in a full-time
type-$s$ job. In a large data set, this equals $\beta^{(0)} + \mathbf{1}_x \cdot \beta^{(x)} + (\mathbf{1}_x * s) \cdot \beta^{(s)} + \sigma(\omega) \cdot \bar{w}$, where $\bar{w}$ is
the upper support of $\omega$. An OLS regression of $\bar{w}_{\text{net}}(x, s)$ on worker type dummies $\mathbf{1}_x$ 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, \( \Pr\{\text{FT}|\text{FT}\} \) and \( \Pr\{\text{FT}|\text{PT}\} \), can be estimated from observed within-job changes in hours. For any \( (\ell, \ell') \in \{\text{PT}, \text{FT}\}^2 \), the probability \( \Pr\{\ell' \mid x, t, h', y = (w, s, \ell)\} \) of going from hours \( \ell \) to \( \ell' \) while staying on the same job, conditional on worker and job characteristics \( (x, t, h', y) \) can be estimated directly from the data.\(^1\) This is based on the following equality:

\[
\frac{\Pr\{\ell' = \text{PT} \mid x, t, h', y = (w, s, \text{FT})\}}{\Pr\{\ell' = \text{PT} \mid x, t, h', y = (w, s, \text{PT})\}} = 1 - \Pr(\text{FT}|\text{FT}) \over 1 - \Pr(\text{FT}|\text{PT}) \tag{A1}
\]

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 \( \Pr\{\text{FT}|\text{FT}\}/\Pr\{\text{FT}|\text{PT}\} \). Solving the resulting linear system of two equations, we obtain \( \Pr\{\text{FT}|\text{FT}\} \) and \( \Pr\{\text{FT}|\text{PT}\} \).

We now prove equation \( \text{A1} \). Consider the probability to go from full time to part time conditionally on \( (x, t, h', w, s) \). We can write this probability as:

\[
\Pr\{\ell' = \text{PT} \mid x, t, h', y = (w, s, \text{FT})\} = (1 - \delta) \left(1 - \Pr(\text{FT}|\text{FT})\right) \left(1 - \lambda_1\right) H \left[ V(x, t+1, h', y' = (w, s, \text{PT})) - U(x, t+1, h') \right] + \lambda_1 \Pr(\varepsilon', y' | x, s) \left\{ V(x, t+1, h', y' = (w, s, \text{PT})) - \varepsilon' \geq \max \langle V(x, t+1, h', y''') - \varepsilon', U(x, t+1, h') \rangle \right\}.
\]

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

\[
\Pr\{\ell' = \text{PT} \mid x, t, h', y = (w, s, \text{PT})\} = (1 - \delta) \left(1 - \Pr(\text{FT}|\text{PT})\right) \left(1 - \lambda_1\right) H \left[ V(x, t+1, h', y' = (w, s, \text{PT})) - U(x, t+1, h') \right] + \lambda_1 \Pr(\varepsilon', y' | x, s) \left\{ V(x, t+1, h', y' = (w, s, \text{PT})) - \varepsilon' \geq \max \langle V(x, t+1, h', y''') - \varepsilon', U(x, t+1, h') \rangle \right\}.
\]

The ratio of these two (observed) probabilities equals \( (1 - \Pr(\text{FT}|\text{FT})) / (1 - \Pr(\text{FT}|\text{PT})) \). This proves equation \( \text{A1} \).

The first reallocation shock \( \tilde{\delta} \) 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-
unemployment transition (call it $\Pr(u2j2u)$). The former probability, $\Pr(u2j)$ is equal to $\lambda_0$ (draw a job offer) times $1 - \tilde{\delta}$ (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 ($\lambda_0$), of accepting a job offer and of the job offer being destroyed before the start of the next period ($\tilde{\delta}$). Hence, $\tilde{\delta}$ can be identified from the ratio $\Pr(u2j2u)/[\Pr(u2j2u) + \Pr(u2j)]$, which is readily available from the data.

The other reallocation shock, $\tilde{\lambda}$ cannot be estimated directly however the ratio $\tilde{\lambda}/\lambda_0$ 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 $\tilde{\lambda}/\lambda_0$ equals $\frac{1 - \tilde{\delta}}{\lambda} \cdot \frac{\Pr(j2u2j)}{\Pr(u2j2u)}$, which is known\(^2\) Hence, we just need to identify $\lambda_0$ and $\tilde{\lambda}$ follows.

C.2 Indirect inference

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 4.4). 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 $y_o$ is given by:\(^3\)

$$
\Pr(U2J, y_o|x, t, h') = \left(1 - \frac{\tilde{\delta}}{\lambda}\right) \lambda_0 \Pr(y_o|x) \cdot H \left[ V(x, t + 1, h', y_o) - U(x, t + 1, h') \right].
$$  \(A2\)

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

$$
\Pr(J2J, y_o|x, t, h', y = (w, s, \ell)) = (1 - \delta) \lambda_1 \Pr(y_o|x, s) \cdot H \left[ V(x, t + 1, h', y_o) - U(x, t + 1, h') \right] \\
\cdot \Pr_{\ell'|\ell} \left\{ V(x, t + 1, h', y') > V(x, t + 1, h', y' = (w, s, \ell')) \right\}, \quad (A3)
$$

where the last term is the probability that job $y_o$ is preferred to job $y$ but the new full-/part-

\(^2\)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.

\(^3\)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 $t$: $y = \emptyset$. 

6
time status $\ell'$ 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. Lastly, remember that the job offer distribution for employed workers depends on their current job’s health content $s$ in the following way:

$$\text{Pr}(y^o|x, s) = \rho \cdot 1 \{ s^o = s \} \text{Pr}(w^o, \ell^o|x, s^o = s) + (1 - \rho) \cdot \text{Pr}(y^o|x)$$

Let us first consider a case where $s^o \neq s$. Taking the ratio between [A2] and [A3], we get:

$$\frac{\text{Pr}(J2J, y^o|x, t, h', y)}{\text{Pr}(U2J, y^o|x, t, h')} = \frac{\lambda_1}{\lambda_0} \cdot \frac{1 - \delta}{1 - \delta} \cdot (1 - \rho) \cdot \text{Pr}_{\ell'|\ell} \left\{ V(x, t + 1, h', y^o) > V(x, t + 1, h', y' = (w, s, \ell')) \right\}$$

(A4)

An implicit assumption here is that the denominator is not zero. What follows holds if for any job $y^o$ 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 not zero. 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 (A4) (observed in the data) over $(y^o, y, x, t, h')$ to obtain $\rho \cdot \lambda_1/\lambda_0$ (recall that the $\delta$ and $\tilde{\delta}$ parameters were already identified).

Now we consider the case $s^o = s$. The ratio between [A2] and [A3] becomes:

$$\frac{\text{Pr}(J2J, y^o|x, t, h', y)}{\text{Pr}(U2J, y^o|x, t, h')} = \frac{\lambda_1}{\lambda_0} \cdot \frac{1 - \delta}{1 - \delta} \cdot \left[ (1 - \rho) + \rho \text{Pr}(s^o = s|x) \right] \cdot \text{Pr}_{\ell'|\ell} \left\{ V(x, t + 1, h', y^o) > V(x, t + 1, h', y' = (w, s, \ell')) \right\}$$

Using a similar weak assumption as above (this time with the constraint that $s = s^o$) we can identify the term on the right-hand side of the first line and, given that we know $\lambda_1/\lambda_0$ from above, we get $\frac{\lambda_1\rho}{\lambda_0 \text{Pr}(s^o = s|x)}$. Summing over the support of $s$ and combining with $\lambda_1\rho/\lambda_0$ we identify $\lambda_1/\lambda_0$ and $\rho$ separately. This also gives us the offer distribution of job health content $\text{Pr}(s^o|x)$. The ratio $\text{Pr}(y^o|x, s) / \text{Pr}(y^o|x)$ is then identified for any $x$, $y^o$ and any $s$.

Once we know these parameters, we can compute $1 \{ V(x, t + 1, h', y^o) > V(x, t + 1, h', y) \}$
already identified). We can write equation (A6) for each value of \( y \) in the discrete support of the job offer distribution (remember that the ratio \( \Pr (y^o|x, s) / \Pr (y^o|x) \) is already identified). All the terms in equation (A6) are known except for \( 1/\lambda_0 \) and \( \Pr (y^o|x) \) for any point on the discrete support of the job offer distribution. Using (A2) to substitute for the \( H[\cdot] \) term on the right-hand side of (A5), we can re-write (A5) as:

\[
\frac{\Pr (\text{stay}|x, t, h', y) \cdot \Pr (y|x) \cdot (1 - \tilde{\delta})}{\Pr (\text{U2J}, y^o|x, t, h') \cdot (1 - \delta)} = \frac{1}{\lambda_0 \Pr (y^o|x)} - \frac{\lambda_1}{\lambda_0} \sum_{y^o} \frac{\Pr (y^o|x, s)}{\Pr (y^o|x)} 1\{V(x, t + 1, h', y^o) > V(x, t + 1, h', y)\}. \tag{A6}
\]

Lastly consider the probability to stay in a job \( y = (w, s, \ell) \), where the new full-/part-time status is realized (and observed for job stayers):

\[
\frac{\Pr (\text{stay}|x, t, h', y) \cdot \Pr (y|x)}{\Pr (\text{U2J}, y^o|x, t, h') \cdot (1 - \delta)} = \frac{1}{\lambda_0 \Pr (y^o|x)} - \frac{\lambda_1}{\lambda_0} \sum_{y^o} \frac{\Pr (y^o|x, s)}{\Pr (y^o|x)} 1\{V(x, t + 1, h', y^o) > V(x, t + 1, h', y)\}. \tag{A5}
\]

Using (A2) to substitute for the \( H[\cdot] \) term on the right-hand side of (A5), we can re-write (A5) as:

\[
\frac{\Pr (\text{stay}|x, t, h', y) \cdot \Pr (y|x) \cdot (1 - \tilde{\delta})}{\Pr (\text{U2J}, y^o|x, t, h') \cdot (1 - \delta)} = \frac{1}{\lambda_0 \Pr (y^o|x)} - \frac{\lambda_1}{\lambda_0} \sum_{y^o} \frac{\Pr (y^o|x, s)}{\Pr (y^o|x)} 1\{V(x, t + 1, h', y^o) > V(x, t + 1, h', y)\}. \tag{A6}
\]

All the terms in equation (A6) are known except for \( 1/\lambda_0 \) and \( \Pr (y^o|x) \) for any point on the discrete support of the job offer distribution (remember that the ratio \( \Pr (y^o|x, s) / \Pr (y^o|x) \) is already identified). We can write equation (A6) for each value of \( y, x, t \) or \( h' \). Doing so yields
a linear system with many more equations than unknowns so, should transition probabilities show enough variations, we can identify $\lambda_0$ and $\Pr(y|x)$ for any value of $x$ and $y$\footnote{Note that we do not need to exclude $(t,h)$ 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.}. We then get $\lambda_1$ as we already know the ratio $\lambda_1/\lambda_0$. Importantly for what follows, once we know $\lambda_0$ and the offer distribution, we identify the selection probability $H[V(x,t+1,h',y') - U(x,t+1,h')]$ from (A2) for any job $y$ and worker $(x,t,h)$.

We now turn to the preference parameters in the instantaneous utility function. First, we re-write the Bellman equations (1):

$$U(x,t,h) = b + (1+r)^{-1}\left\{ \sum_{h'} \Pr(h',y' = \emptyset|x,t,h,y = \emptyset) \cdot U(x,t+1,h') + \sum_{h',y' \neq \emptyset} \int_{\varepsilon'} \Pr(h',y' = y^o,\varepsilon'|x,t,h,y = \emptyset) \cdot [V(x,t+1,h',y^o) - \varepsilon'] \, dH(\varepsilon') \right\}$$ (A7)

and (2):

$$V(x,t,h,y) = u(x,t,h,y) + (1+r)^{-1}\left\{ \sum_{h'} \Pr(h',y' = \emptyset|x,t,h,y) \cdot U(x,t+1,h') + \sum_{h',y' \neq \emptyset} \int_{\varepsilon'} \Pr(h',y' = y^o,\varepsilon'|x,t,h,y) \cdot [V(x,t+1,h',y^o) - \varepsilon'] \, dH(\varepsilon') \right\}.$$ (A8)

The last line of (A7) can be rewritten as follows:

$$\sum_{h',y'} \Pr(h',y' = y^o|t,h,y = \emptyset) \cdot [V(t+1,h',y^o) - \mathbf{E}(\varepsilon'|h',y' = y^o,t,y = \emptyset)]$$ (A9)

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

Recall that we can identify $H[V(x,t,h',y') - U(x,t,h')]$ from (A2). We know that $\varepsilon$ follows a log-normal distribution so its cdf is invertible. Assuming that we know its standard error $\sigma^4$ we can then identify the function $(y,x,t,h') \mapsto V(x,t,h',y') - U(x,t,h')$ over the set of $(y,x,t,h')$ such that $V(x,t,h',y) > U(x,t,h')$. That is enough to identify the conditional expectation in (A9) from the data (as whenever $V(x,t,h',y) \leq U(x,t,h')$, the probability of

\footnote{Proving the identification of the standard deviation $\sigma$ of the labor supply shock is not straightforward however (note that its mean has been normalised in section 4.3). 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.}
observing that job being taken up, \( \Pr(h', y' = y^0 | t, h, y = \emptyset) \), is zero).

Once we know the conditional expectation in (A9), we have identified all the terms in equation (A7) 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 (A8). 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 \( y' \) at \( t + 1 \) and was in \( y \) at \( t \) only yields information on the shock through the fact that \( y' \) is preferred to unemployment. Knowing that \( y' \) is potentially better than \( y \) (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 (A8) is also conditional on \( \varepsilon' \leq V(x, t + 1, h', y^0) - U(x, t + 1, h') \) so we can identify it the same way we did the one in (A9).

To summarize, let us stack all values \( U(x, t, h) \) and \( V(x, t, h, y) \) over the discrete support of \( (x, t, h, y) \) in a column vector denoted as \( C_{\text{val}} \). Likewise, we stack all instantaneous utilities \( b(x, t, h) \) and \( u(x, t, h, y) \) in a column vector \( C_{\text{util}} \). We can then rewrite the system of equations (A7)-(A8) as:

\[
C_{\text{util}} = P \cdot C_{\text{val}} + B,
\]

where \( P \) is a square matrix with coefficients equal to 0, -1 or combinations of transition probabilities \( \Pr(h', y'|x, t, h, y) \) (and the discount factor which we assume to be known) and \( B \) is a known vector of (combinations of) the conditional expectations of the labour supply shock. Note that \( P \) 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, \( P \) and \( B \) can be directly estimated using data on health and labour market transitions.

Then, if \( P \) is invertible equation (A10) means we can write the value function at each \( (x, t, h, y) \) 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 \( u(x, t, h, y) = w - h(\alpha s + \beta t) \), where \( b, \alpha \) and \( \beta \) are the parameters to identify. Recall that a
job $y$ is a triplet $(w, s, \ell)$. Multiplying both sides of $\text{(A10)}$ by $P^{-1}$ leads to:

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

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

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:

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

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 $\text{(A2)}$ we can then estimate the parameters of interest $\alpha, \beta$ and $b$. 
D Additional estimation results

Figure A3: Model fit - Auxiliary regressions

a: Log wage
b: Log wage change
c: Employment
d: Job stress change
e: Wage variance
f: Job stress pdf

Notes: Red dots: Model-produced moments
Blue line segments: 95% bootstrap confidence interval of data moments.
E  Details of the partially directed search model

E.1 Derivation

For convenience, we reiterate the value function for employed workers presented in Section 7:

\[ V(x,t,h,y) = w - c(\ell,s,t,h) + (1 + r)^{-1} \sum_{h' \in H} \sum_{\ell' \in \mathcal{L}} \Pr\{h'|y,x,t,h\} \Pr\{\ell'|\ell\} \]
\[ \times \left[ \delta \left(1 - \tilde{\lambda}\right) U(x,t+1,h') + \delta \lambda \mathbb{E}_{\zeta} \max_k \langle S(x,t+1,h',\emptyset,s_k) + \zeta_k \rangle \right. \]
\[ \left. + (1 - \delta) \mathbb{E}_{\zeta} \max_k \langle S(x,t+1,h',y',s_k) + \zeta_k \rangle \right] \]

and explicate the value function of unemployed workers:

\[ U(x,t) = b(x,t) + (1 + r)^{-1} \sum_{h' \in H} \Pr\{h'|x,t,h\} \mathbb{E}_{\zeta} \max_k \langle S(x,t+1,h',y',s_k) + \zeta_k \rangle \]

with:

\[ S(x,t,h,y,s_k) = (1 - \lambda_1(s_k|s)) \int \max \{V(x,t,h,y) - \varepsilon; U(x,t,h)\} dH(\varepsilon) \]
\[ + \lambda_1(s_k|s) \int \max \{V(x,t,h,y^o) - \varepsilon; V(x,t,h,y) - \varepsilon; U(x,t,h)\} dF(y^o|x, s = s_k) dH(\varepsilon) \]

and for unemployed workers (\(y = \emptyset\)):

\[ S(x,t,h,\emptyset,s_k) = \left(1 - \lambda_0(s_k|s)\right) \left(1 - \tilde{\delta}\right) U(x,t,h) \]
\[ + \lambda_0(s_k) \left(1 - \tilde{\delta}\right) \int \max \{V(x,t,h,y^o) - \varepsilon; U(x,t,h)\} dF(y^o|x, s = s_k) dH(\varepsilon) \]

The value of search for a type-\(s_k\) job, \(S(\cdot)\), can be rewritten as:

\[ S(x,t,h,y,s_k) = U(x,t,h) + (1 - \lambda_1(s_k)) \mathcal{H}[V(x,t,h,y) - U(x,t,h)] \]
\[ + \lambda_1(s_k|s) \int \mathcal{H}[\max \{V(x,t,h,y^o); V(x,t,h,y)\} - U(x,t,h)] dF(y^o|x, s = s_k) \quad (A11) \]

and

\[ S(x,t,h,\emptyset,s_k) = U(x,t,h) + \lambda_0(s_k) \left(1 - \tilde{\delta}\right) \int \mathcal{H}[V(x,t,h,y^o) - U(x,t,h)] dF(y^o|x, s = s_k) \]
\[ \quad (A12) \]
where again $\mathcal{H}(X) = \int_{-\infty}^X H(x)dx$. Next, substituting the solution $\mathbb{E}_\zeta \max_k \langle S(x, t + 1, h', y', s_k) + \zeta_k \rangle = \ln \left( \sum_{k=1}^K e^{S(x, t, h, y, s_k)} \right) + \gamma$ into the definitions of $V(\cdot)$ and $U(\cdot)$ (and ignoring the constants):

$$V(x, t, h, y) = w - c(\ell, s, t, h) + \frac{1}{1 + r} \sum_{h' \in H} \sum_{h' \in L} \Pr \{ h'|y, x, t, h \} \Pr \{ \ell'|\ell \} \times \left[ \delta \left( 1 - \tilde{\lambda} \right) U(x, t + 1, h') + \tilde{\delta} \tilde{\lambda} \ln \left( \sum_{k=1}^K e^{S(x, t+1, h', y, s_k)} \right) + (1 - \delta) \ln \left( \sum_{k=1}^K e^{S(x, t+1, h', y, s_k)} \right) \right]$$

$$U(x, t, h) = b(x, t, h) + \frac{1}{1 + r} \sum_{h' \in H} \Pr \{ h'|\emptyset, x, t, h \} \ln \left( \sum_{k=1}^K e^{S(x, t+1, h', y, s_k)} \right)$$

Further substituting (A11) and (A12), the value function equations can be rewritten as:

$$V(x, t, h, y) = w - c(\ell, s, t, h) + \frac{1}{1 + r} \sum_{h' \in H} \sum_{h' \in L} \Pr \{ h'|y, x, t, h \} \Pr \{ \ell'|\ell \} \times \left[ U(x, t + 1, h') + (1 - \delta) \mathcal{H} \left[ V(x, t + 1, h', y) - U(x, t + 1, h') \right] \right]$$

$$+ \tilde{\delta} \lambda \ln \left( \sum_{k=1}^K \exp \left\{ \lambda_0(s_k) \left( 1 - \delta \right) \int \mathcal{H} \left[ V(x, t + 1, h', y^o) - U(x, t + 1, h') \right] dF(y^o|x, s = s_k) \right\} \right)$$

$$+ (1 - \delta) \ln \left( \sum_{k=1}^K \exp \left\{ \lambda_1(s_k|s) \int \max \mathcal{H} \left[ V(x, t + 1, h', y') - U(x, t + 1, h') \right] - \mathcal{H} \left[ V(x, t + 1, h', y) - U(x, t + 1, h') \right] ; 0 \right\} dF(y^o|x, s = s_k) \right\}$$

$$U(x, t, h) = b(x, t, h) + \frac{1}{1 + r} \sum_{h' \in H} \Pr \{ h'|\emptyset, x, t, h \} \left[ U(x, t + 1, h') \right]$$

$$+ \ln \left( \sum_{k=1}^K \exp \left\{ \lambda_0(s_k) \left( 1 - \delta \right) \int \mathcal{H} \left[ V(x, t + 1, h', y^o) - U(x, t + 1, h') \right] dF(y^o|x, s = s_k) \right\} \right)$$

Those value functions can be tabulated, solving the model backwards starting from the retirement age $T$ (as workers do not search at age $T + 1$).
E.2 Parameterization, estimation, and model fit

The partially directed search model is formally very similar to the random search model presented in the main body of the paper. For the directed search model, we specify

\[
\lambda_0(s_k) = \frac{\lambda_0 s_k^{\alpha_1 - 1} \cdot (1 - s_k)^{\alpha_2 - 1}}{\sum_{j=1}^{5} s_j^{\alpha_1 - 1} \cdot (1 - s_j)^{\alpha_2 - 1}}
\]

and

\[
\lambda_1(s_k|s) = \lambda_1 \left( \frac{s_k^{\alpha_1 - 1} \cdot (1 - s_k)^{\alpha_2 - 1}}{\sum_{j=1}^{5} s_j^{\alpha_1 - 1} \cdot (1 - s_j)^{\alpha_2 - 1}} + \rho \cdot 1\{s_k = s\} \right)
\]

where \(\lambda_0\), \(\lambda_1\) and \(\rho\) are constants. The rest of the directed search model specification is identical to that of the random search model.

Estimation is carried out by indirect inference, following the exact same steps as in the random search case. Parameter estimates are collated in Tables A2 and A3 (which mirror Tables 5 and 6 from the random search model). Value indicated in italics in Tables A2 and A3 are estimated independently of the random vs directed specification of the model, and as such they are exactly identical between the two specifications.

Table A2: Job offer characteristics and arrival rates - Directed Search

<table>
<thead>
<tr>
<th>Job offers</th>
<th>Arrival rates and sampling distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\beta^{(0)})</td>
<td>6.484</td>
</tr>
<tr>
<td>(\beta_2^{(x)})</td>
<td>0.531</td>
</tr>
<tr>
<td>(\beta_3^{(x)})</td>
<td>0.672</td>
</tr>
<tr>
<td>(\beta_4^{(x)})</td>
<td>0.907</td>
</tr>
<tr>
<td>(\alpha_1)</td>
<td>1.622</td>
</tr>
<tr>
<td>(\delta)</td>
<td>0.056</td>
</tr>
<tr>
<td>(\epsilon)</td>
<td>4.753</td>
</tr>
<tr>
<td>(\kappa_1^{(t)})</td>
<td>0.333</td>
</tr>
<tr>
<td>(\kappa_1^{(s)})</td>
<td>1.174</td>
</tr>
<tr>
<td>(\kappa_1^{(t)})</td>
<td>1.539</td>
</tr>
<tr>
<td>(\kappa_1^{(t)})</td>
<td>0.551</td>
</tr>
</tbody>
</table>

Table A3: Utility estimates - Directed Search

<table>
<thead>
<tr>
<th>Disutility of labor</th>
<th>Supply shock</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\kappa_1^{(t)})</td>
<td>0.333</td>
</tr>
<tr>
<td>(\kappa_1^{(s)})</td>
<td>1.174</td>
</tr>
<tr>
<td>(\kappa_1^{(t)})</td>
<td>1.539</td>
</tr>
<tr>
<td>(\kappa_1^{(t)})</td>
<td>0.551</td>
</tr>
</tbody>
</table>

Figures A4 and A5 summarize the directed search model fit, echoing Figures 3 and A3 in the random search case. Those figures suggest that quality of fit is similar in both models.
Table A4: Labor supply cost estimates (GBP/month) - Directed Search

<table>
<thead>
<tr>
<th>Health state</th>
<th>age 30</th>
<th></th>
<th>age 50</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$s_1$</td>
<td>$s_2$</td>
<td>$s_3$</td>
<td>$s_4$</td>
</tr>
<tr>
<td>Severe</td>
<td>238</td>
<td>274</td>
<td>309</td>
<td>344</td>
</tr>
<tr>
<td>Poor</td>
<td>94</td>
<td>108</td>
<td>122</td>
<td>136</td>
</tr>
<tr>
<td>Average</td>
<td>82</td>
<td>94</td>
<td>107</td>
<td>119</td>
</tr>
<tr>
<td>Good</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

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

Figure A4: Model fit - Work and health outcomes by age - Directed Search

a: Log wage

b: Part-time rate

c: Unemployment rate

d: Pr(job-to-job)

e: Pr(unemp-to-job)

f: Pr(quit)

g: Health - FT Employed

h: Health - Unemployed

i: Job content

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.
Figure A5: Model fit - Auxiliary regressions - Directed Search

Notes: Red dots: Model-produced moments
Blue line segments: 95% bootstrap confidence interval of data moments.
E.3 Counterfactuals

In this final subsection, we collate results from all of the counterfactual experiments discussed in Section 6, this time performed using our estimated directed search model. We comment on these results in section 7.2.

E.3.1 Job loss

Table A5: Value of a full-time job (vs unemployment) at age 30 and 50

<table>
<thead>
<tr>
<th>Initial health</th>
<th>Wage decile</th>
<th>Age 30</th>
<th></th>
<th>Wage decile</th>
<th>Age 50</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2nd</td>
<td>5th</td>
<td>8th</td>
<td>2nd</td>
<td>5th</td>
</tr>
<tr>
<td>Severe</td>
<td>14,427 (11.0%)</td>
<td>34,816 (22.3%)</td>
<td>69,116 (35.5%)</td>
<td>11,996 (9.4%)</td>
<td>31,112 (20.5%)</td>
</tr>
<tr>
<td>Poor</td>
<td>17,103 (12.6%)</td>
<td>37,509 (23.4%)</td>
<td>71,815 (36.2%)</td>
<td>15,377 (11.5%)</td>
<td>34,548 (22.0%)</td>
</tr>
<tr>
<td>Average</td>
<td>17,584 (12.8%)</td>
<td>38,002 (23.6%)</td>
<td>72,314 (36.3%)</td>
<td>16,040 (11.8%)</td>
<td>35,255 (22.3%)</td>
</tr>
<tr>
<td>Good</td>
<td>19,001 (13.7%)</td>
<td>39,427 (24.2%)</td>
<td>73,746 (36.6%)</td>
<td>17,815 (12.9%)</td>
<td>37,057 (23.1%)</td>
</tr>
</tbody>
</table>

Note: Value differences $\Delta \ell(t,y,h)$ as defined in (8), measured in GBP.
Relative differences in parentheses: $\Delta \ell(t,y,h)/E_{x,\epsilon}[V(x,t,h,y) - \epsilon]$.

Figure A6: Health effects of job loss at age 30 and 50

Note: The vertical axis shows the difference in the probability of being in severe or poor health, given initial health at age 30 (or 50), between an unemployed worker and a full-time worker at the median wage/medium job stress content. The horizontal axis shows age (in years).
Figure A7: Labor market effects of job loss at age 30 and 50

Age 30

Conditionnal on being in ‘Average’ health

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

Relative to being in ‘Average’ health

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

Age 50

Conditionnal on being in ‘Average’ health

g: Unemployment rate  
h: Income (log)  
i: Job health content

Relative to being in ‘Average’ health

j: Unemployment rate  
k: Income (log)  
l: 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.
### E.3.2 Health shocks

#### Table A6: Value of being in the average vs severe health state at age 30 and 50

<table>
<thead>
<tr>
<th></th>
<th>Age 30</th>
<th>Age 50</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Employed full-time at the median wage in a...</td>
<td>Employed full-time at the median wage in a...</td>
</tr>
<tr>
<td></td>
<td>medium-stress job</td>
<td>high-stress job</td>
</tr>
<tr>
<td>( \Delta^{(h)c} )</td>
<td>2,425</td>
<td>2,979</td>
</tr>
<tr>
<td>( \Delta^{(h)V} )</td>
<td>4,297</td>
<td>5,161</td>
</tr>
<tr>
<td>( MW P^{(h)} )</td>
<td>0.048</td>
<td>0.059</td>
</tr>
</tbody>
</table>

**Notes:** Differences in costs \( \Delta^{(h)c} \) and values \( \Delta^{(h)V} \) are in GBP.  
\( \Delta^{(h)c} \) is the labor cost in average health minus the labor cost in severe health.  
\( \Delta^{(h)V} \) is the job value in average health minus the job value in severe health.  
Medium (high) stress refers to the 3rd out of 5 (highest) job health content.

#### Figure A8: Effects of being in severe vs average health at age 30 and 50

**Age 30**

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

**Age 50**

- **g:** Prob. of Severe or Poor health
- **h:** Income (log)
- **i:** Unemployment rate

**Notes:** Difference between the outcomes of individuals in severe vs average health at age 30 or 50

*Age on horizontal axis.*
### E.3.3 Stressful jobs

Table A7: Value of a medium- vs high-stress job at age 30 and 50

<table>
<thead>
<tr>
<th>Age 30</th>
<th>Health</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Severe</td>
<td>Poor</td>
<td>Average</td>
<td>Good</td>
</tr>
<tr>
<td>$\Delta(s)_C$</td>
<td>846</td>
<td>333</td>
<td>292</td>
<td>13</td>
</tr>
<tr>
<td>$\Delta(s)V$</td>
<td>3,592</td>
<td>2,848</td>
<td>2,635</td>
<td>2,258</td>
</tr>
<tr>
<td>$MW P(s)$</td>
<td>0.038</td>
<td>0.029</td>
<td>0.027</td>
<td>0.023</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Age 50</th>
<th>Health</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Severe</td>
<td>Poor</td>
<td>Average</td>
<td>Good</td>
</tr>
<tr>
<td>$\Delta(s)_C$</td>
<td>3,263</td>
<td>1,286</td>
<td>1,127</td>
<td>49</td>
</tr>
<tr>
<td>$\Delta(s)V$</td>
<td>10,414</td>
<td>7,734</td>
<td>7,062</td>
<td>5,677</td>
</tr>
<tr>
<td>$MW P(s)$</td>
<td>0.090</td>
<td>0.067</td>
<td>0.061</td>
<td>0.049</td>
</tr>
</tbody>
</table>

Note: Differences $\Delta(s)_C$ and $\Delta(s)V$ are in GBP.  
$\Delta(s)_C$ is the labor cost in a stressful job minus the cost in a medium-stress job.  
$\Delta(s)V$ is the value in a medium-stress job minus the value in a stressful job.

Figure A9: Effects of going from a medium to most stressful job type

Age 30

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

Age 50

- g: Prob. of Severe or Poor health
- h: Job health content
- i: Unemployment rate

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


