

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository:<https://orca.cardiff.ac.uk/id/eprint/131037/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Jones, Melanie K. and McVicar, Duncan 2020. Estimating the impact of disability onset on employment. *Social Science and Medicine* 255 , 113001. 10.1016/j.socscimed.2020.113001

Publishers page: <http://dx.doi.org/10.1016/j.socscimed.2020.113001>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies. See <http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



Estimating the Impact of Disability Onset on Employment

Melanie K. Jones^{a,c} and Duncan McVicar^{b,c,*}

^a Cardiff Business School
Cardiff University, Cardiff, Wales
Tel: +44 (0)2920875079
Email: jonesm116@cardiff.ac.uk

^b Queen's Management School
Queen's University Belfast, Belfast, Northern Ireland
Tel: +44 (0)2890974809
Email: d.mcvicar@qub.ac.uk

^c Institute of Labor Economics, Bonn

Revised Version: February 2020

JEL: J14, J21

Keywords: disability, employment, local labour force survey, propensity score matching

Acknowledgements

This work is based on data from the Annual Population Survey which is produced by the ONS and was accessed via the UK Data Service Secure Lab. We are grateful to the UK Data Service Team for their support. The use of these data does not imply the endorsement of the data owner or the UK Data Service at the UK Data Archive in relation to the interpretation or analysis of the data. This work uses research datasets which may not exactly reproduce National Statistics aggregates. This research did not receive any specific grant funding from agencies in the public, commercial, or not-for-profit sectors.

Declarations of interest: none

*corresponding author

Abstract

This paper examines the impact of disability onset on the probability of employment using an underexplored longitudinal dataset for Britain. It contrasts estimates based on a control group drawn from those not experiencing disability onset – a common approach in the literature – with estimates based on a control group drawn from those who do experience disability onset, but one year after the treatment group. Compared to the non-disabled control group, the control group of later-onsetters is likely to be more similar to the treatment group in terms of unobservables, with the resulting estimates therefore more plausibly interpreted as causal. Using this control group we estimate that the probability of employment falls by 11 percentage points in the year of disability onset. The equivalent estimate using the control group drawn from those not experiencing onset is about fifty percent larger. The employment effects of disability onset are also shown to be larger for those with lower qualification levels, consistent with weaker attachment to the labour market.

1. Introduction

There is an extensive international literature examining the impact of disability on labour market outcomes using cross-sectional data (e.g. Kidd, Sloane and Ferko, 2000; DeLeire, 2001). There is also a growing literature which exploits longitudinal data to examine the labour market impact of disability onset (e.g. Charles, 2003; Jenkins and Rigg, 2004; Garcia-Gomez, 2011; Garcia-Gomez, van Kippersluis, O'Donnell and van Doorslaer, 2013; Polidano and Vu, 2015; Jones, Davies and Drinkwater, 2018; Meyer and Mok, 2019). An advantage of this dynamic approach is that it can help to estimate associations which are closer to having a causal interpretation.

This paper adopts a dynamic approach to estimate the employment effects of disability onset in Britain. The focus on employment is motivated by the substantial impact of disability onset found in the existing international literature (see, for example, Meyer and Mok, 2019). The application to Britain reflects the persistence of the disability employment gap at a time of record overall employment rates and despite previous government commitments to reduce it (see Baumberg, Jones and Wass, 2015), recently renewed as part of the National Disability Strategy announced in December 2019. In contrast to the existing literature, however, we complement estimates based on comparing outcomes for those experiencing onset with those not experiencing onset (the standard approach in the literature, e.g. see Jenkins and Rigg (2004) and Polidano and Vu (2015)), with estimates comparing outcomes for those experiencing onset with those who experience disability onset on average *one year later*.

Interpreting differences in employment outcomes between those who do and do not experience disability onset as causal, even when conditioning on observable differences between individuals, is complicated by potential biases due to simultaneity (employment may impact on disability or willingness to report disability in surveys), unobserved time-invariant and

unobserved time-varying confounders. Augmenting regression or propensity score matching methods with individual fixed effects (e.g. Jones et al., 2018), differencing (e.g. Polidano and Vu, 2015) or exact matching on pre-onset outcomes (e.g. Garcia-Gomez, 2011) can reduce or eliminate some of these biases (in particular eliminating biases due to time-invariant unobserved confounders), but other biases, including those driven by remaining unobserved time-varying confounders – anything unobserved that might simultaneously influence both employment and disability – may remain. Such biases are likely to be smaller, although not necessarily entirely absent, when comparing outcomes for those experiencing disability onset at time t with those experiencing disability onset at time $t+1$, since they are likely to be more similar in terms of both time-invariant and time-varying unobservables than is the case when compared to those who do not experience disability onset at all.

In making this argument we build on a recent paper by Fadlon and Nielsen (2015). This earlier paper uses Danish administrative data to show substantial pre-treatment divergence in labour supply ahead of a spousal health shock – suggestive of time-varying confounders – when compared to those not experiencing such a shock, but no such pre-treatment divergence when comparing to those experiencing a similar shock one year later. Fadlon and Nielsen (2019) apply a similar identification strategy, again using Danish administrative data, when examining health spillovers within families. In contrast to these earlier papers, however, our focus is on the employment effects of own disability onset rather than the health and employment impacts of health shocks of family members. In a further contrast to these earlier papers, we use survey data as opposed to administrative data here. We argue that the Fadlon and Nielsen approach may have an additional benefit in this context because it can help to reduce biases that may be more likely in survey data. In particular, given the subjective nature of our treatment (self-reported disability onset), this approach may have implications for justification bias, a potential source of simultaneity.

Like Fadlon and Nielsen (2015), in our application we show diverging prior trends when comparing those experiencing disability onset with the non-disabled control group, suggesting the presence of time-varying confounders and/or justification bias. There is no such divergence, however, when comparing those experiencing disability onset to those experiencing later onset. Using a propensity score matching (PSM) approach supplemented with exact matching on base-year employment, we then show that the magnitude of employment effects of disability onset are smaller using the alternative control group of later-onsetters compared to using the standard non-disabled control group. This suggests large potential upward biases in the estimates based on the non-disabled control group. Nevertheless, our estimate based on the later-onset control group still shows a sizeable drop (of 11 percentage points on a base of a 75% employment rate) in employment in the year of disability onset, consistent with a causal impact of disability on employment in Britain.

2. Data

We exploit the 25% rotational panel structure of the Local Labour Force Survey (LLFS), included within the Annual Population Survey (APS) (Office for National Statistics, 2020), where individuals are retained in the sample for up to four years. Analysis is restricted to working-age respondents (16-64 for men, 16-59 for women) who provide information at four consecutive waves between 2004 and 2012, creating a balanced panel with maximum sample of 49,059 individuals from several overlapping cohorts (individuals enter the survey between 2004 and 2009). The questions used to identify disability changed in 2013 and preclude the inclusion of more recent data. Since the LLFS was designed to boost the APS in parts of Britain it is not fully geographically representative. Together with the balanced panel restriction, the

result is a slightly older sample that has higher rates of disability than the full APS sample (see Jones et al., 2018).

Consistent with the dynamics of disability and labour market outcomes literature (e.g. Charles, 2003; Jones et al., 2018; Meyer and Mok, 2019), a work-limiting definition of disability is adopted here. Disabled individuals respond positively to an initial long-term health (LTH) question: “Do you have any health problems or disabilities that you expect will last for more than a year?” and to either of the follow-up questions: “Does this health problem affect the kind of paid work you might do? Does this health problem affect the amount of paid work you might do?”. Individuals answering ‘no’ to the initial LTH question, or those answering ‘yes’ but ‘no’ to both the follow-up questions, are classed as non-disabled. The prevalence of disability in the (pooled) sample is 17.5%.

The subjective nature of reported disability raises established concerns about potential measurement error (which may bias estimated onset effects in an uncertain direction (see Van Soest, Andreyeva, Kapteyn and Smith (2012)) and justification or rationalization bias (which may inflate estimated onset effects, and might be reflected in diverging prior trends), although Benitez-Silva et al. (2004) presents evidence that questions the economic significance of the latter. Exact matching on base-year employment outcomes likely reduces the impact of any such justification bias (see Garcia-Gomez, 2011), as does comparing a treatment group who report disability onset to a control group who also report disability onset just one year later (because the justification mechanism potentially exists for both the treatment and control groups, acting in the same direction, rather than solely for the treatment group). The latter argument suggests an additional potential benefit from the Fadlon and Neilsen identification approach when using self-reported disability or subjective health measures from survey data in this context.

In an effort to reduce other types of measurement error and to net out the most transitory periods of disability we adopt a two-period measure of disability onset among the treatment group, similar to Jenkins and Rigg (2004), i.e. the treatment group consists of those individuals who have two periods of not reporting disability followed by two periods of reporting disability (0011), about 1.2% of the total sample. However, since we only observe these individuals for four waves we can say nothing about how disability onset fits into broader patterns of disability or the permanency of onset. We then specify two alternative control groups. Following the standard approach in the literature (e.g. Garcia-Gomez, 2011; Polidano and Vu, 2015), the first control group is drawn from those continuously non-disabled (0000), i.e. those at risk of, but who do not experience, onset, which form the majority (72.6%) of the total sample. In a break with this tradition, but following the recent paper by Fadlon and Nielsen (2015), we draw an alternative control group from those who experience disability onset one year later (i.e. 0001), about 2.7% of the sample. Note that because we are limited to four observations per individual, and because we need to retain two periods prior to reporting disability onset for the treatment group to explore pre-onset trends, we cannot restrict the control group of later-onsetters to those with two consecutive years of disability. If this control group experiences shorter or less severe disability spells on average as a result then the magnitude of the estimated treatment effect may retain some upwards bias. However, the extent of this is likely to be considerably less than when using the standard control group.

Table 1 provides the sample sizes and proportions in (ILO) employment by wave and treatment status. Note the large declines in employment at onset for both the treatment (onset between wave 2 and wave 3) and alternative control (onset between wave 3 and wave 4) group, compared to stable outcomes for the standard control group. But also note pre-onset employment declines for both the treatment and alternative control groups. This pattern –

which could reflect justification bias, prior declining health, or other time-varying confounders – is not unique to this particular application (e.g. Meyer and Mok, 2019).

<Table 1 around here>

3. Methods

Like Garcia-Gomez (2011) we use a PSM approach combined with exact matching to identify disability onset effects separately from compositional differences between the treatment and control groups, under a standard conditional independence assumption (CIA), i.e. that there are no relevant *unobserved* differences between the treatment and control groups (see Rosenbaum and Rubin, 1983). This is implemented in Stata using PSMATCH2 (Leuven and Sianesi, 2003) where, the treatment and control groups are first matched exactly on year and wave 1 employment status and then matched by PSM on the probability of treatment (i.e. disability onset). The PSM entails estimating a probit model for treatment regressed on an extensive set of observable characteristics, measured in wave 1, which are set out in Table 2. Note that the exact matching on wave 1 employment nets out relevant time-invariant unobservable differences between the treatment and control groups, so that the relative merits of the standard and alternative control groups here relate to the extent to which they ameliorate biases related to unobserved time-varying heterogeneity and, in this particular context, justification bias.

<Table 2 around here>

For each individual experiencing onset the individual in the relevant control group with the most similar probability of experiencing onset between waves 2 and 3 given their characteristics, but who did not do so, is then identified (their ‘nearest neighbour’ (NN(1))). Calculating how the treated individuals’ employment outcomes differ from their matched partners’ outcomes, and averaging these differences over all treated individuals (within the

region of common support), yields our estimates of the impact of disability onset on those who experience it, in the year of onset. The latter is a necessary restriction because the alternative control group of later-onsetters are themselves treated – they experience disability onset – between waves 3 and 4. On the other hand we are able to report estimated disability onset impacts on employment in the following year when using the standard, non-disabled, control group. Note, however, that existing evidence suggests that the majority (but not all) of the impact of disability onset is evident in the period immediately post-onset (Mok, Meyer, Charles and Achen, 2008; Meyer and Mok, 2019; Polidano and Vu, 2015).

The matching works well here, with very few off-support individuals in either case (more than 99% of the treatment group are in the region of common support in each case). If the CIA holds, given the near-100% common support, our estimates are therefore interpretable as the average treatment effect on the treated, or ATT. The critical question, of course, is whether the CIA can plausibly be expected to hold here. Table 2 reports balancing tests for the treatment group relative to each control group and confirms there are no significant differences in wave 1 observable characteristics between the matched treatment and either control group at the 95% level, which is encouraging. (Note also that in support of our argument that the control group of later-onsetters is likely to be more similar in terms of unobservables to the treatment group than is the case for the standard non-disabled control group, there are few differences in observables between the two former groups even prior to matching.) But what of time-varying confounders and justification bias?

Disability onsets necessarily encompass both shock events (e.g. as a result of a traffic or industrial accident) and gradual deteriorations in capacity that result in individuals crossing a subjective threshold at which they begin to report themselves as disabled. In survey data individuals may also start to report themselves as disabled to rationalize (or justify) non-employment, perhaps but not necessarily coupled with benefit claiming. Consistent with the

literature, we focus on disability onset as the act of crossing the threshold, but are able to explicitly exclude at least some of the impact of pre-onset deterioration in health. (There is, of course, also merit in trying to establish the broader impact of disability onset, including any prior deterioration in health, but this is subject to an arbitrary choice of a year from which the decline in health forms part of disability onset.)

For shock disability onsets, where there is no pre-onset deterioration in health, the alternative control group better accounts for other time-varying unobservables than the standard control, as per the arguments of Fadlon and Nielsen (2015). Although not always explicit in the literature, for disability onset arising from a gradual decline in health, the difference in post-onset outcomes between the treatment group and the standard non-disabled control group may include the effects of pre-onset deterioration in health. These are likely to be much more modest when compared to the later-onset control group, some of whom will also experience a pre-onset deterioration in health. As such, the alternative control group, in better adjusting for time-varying unobservables, also nets out some of the influence of any prior deterioration in health. Further, because the justification bias mechanism potentially drives some reported disability onsets for both the treatment and alternative control groups, but not the standard control group, any justification bias is likely to be reduced in the former case relative to the latter case. In these respects the alternative control provides a more conservative measure of the impact of disability onset and, therefore, estimates of the impact of disability onset from the standard and alternative control are likely to provide a range of magnitudes reflecting their differing ability to ameliorate justification bias and the influence of time varying unobservables, including prior health.

In sensitivity analysis we repeat the estimation: using alternative matching algorithms (five nearest neighbours, local linear regression, and kernel density); including wave 1 LTH in the PSM; excluding selected cohorts of those experiencing disability onset to explore whether

estimates are sensitive to differential effects of the financial crisis; and replacing the work-limiting definition of disability with an activity-limiting definition, which Oguzoglu (2012) argues may be less susceptible to justification bias. Finally, in an extension to examine whether disability onset effects are heterogeneous across different socio-demographic groups (likely differing in their pre-existing attachment to the labour market, among other things) and for different types of onset – in both cases with potential relevance for policy targeting – we repeat the matching process (with a suitably modified set of observable characteristics) for subsamples split dichotomously by gender (male/female), age (older/younger than 45), area unemployment rate (above/below the median unemployment in the NUTS2 area), marital status (married/non-married), qualification level (higher/lower than GCSE or equivalent A-C), whether disability onset reflects a main health condition that is mental or physical, a single or multiple health conditions, and whether the individual reported a LTH in wave 1. As in the main analysis, common support exceeds 99% of the treatment group in each case.

4. Estimated Employment Effects of Disability Onset

The upper panel of Table 3 shows the difference in employment rates averaged over the matched treatment and standard non-disabled control groups, for wave 2 (pre-onset), wave 3 (the year of onset) and wave 4 (the following year). The wave 3 estimate for the year of onset is our first PSM estimate of the (short-run) effect of disability onset on employment for our sample. It is very large, at -16 percentage points (on a base employment rate of 75%), and highly statistically significant. There is also a large and statistically significant employment gap of 6 percentage points in the year prior to onset, however, suggesting the presence of time-varying confounders, potentially including but unlikely to be limited to unmeasured deterioration in health, and/or justification bias. The result is that we cannot be sure how much (if any) of the estimated 16 percentage point gap in employment in the year following onset is a causal effect of onset, and how much reflects other differences between the treatment and

control groups. Further, because we do not know how much of the 6 percentage point decline in employment in the year prior to onset is due to pre-onset deterioration in health, this problem remains even if we are willing to interpret disability onset as including lead effects due to deteriorations in health between waves 1 and 2, i.e. prior to reported onset.

<Table 3 around here>

In contrast, the lower panel of Table 3 shows a smaller, although still highly statistically significant estimated disability effect on employment in the year of onset, at -11 percentage points, when the treatment group is compared to the alternative control group drawn from later-onsetters. In this case there is no evidence of a pre-onset employment gap in the year prior to onset (the gap is only 2 percentage points and not statistically significant), consistent with our conjecture that unobserved confounders and/or justification bias are less likely in this case than in the standard control group case. Although we cannot rule out such factors after wave 2, the resulting PSM estimate is more plausibly interpreted as approaching a causal estimate of disability onset than is the case for the estimate using the non-disabled control group.

The final row of the upper panel of Table 3 shows an additional divergence in employment outcomes between the treatment group and the non-disabled control group in the subsequent year following onset. For the reasons discussed above we cannot be sure whether or to what extent this further divergence captures a growing causal impact of disability onset as opposed to confounders. Nevertheless the drop in employment rates is far greater in the year of disability onset than in this second year, supporting our focus on the period immediately post-onset.

The key conclusions above are robust to the sensitivity analyses listed in Section 3 (full results are available from the authors on request). Specifically, the three alternative matching algorithms (five nearest neighbours, local linear regression, and kernel density) return estimates of the wave 3 employment effect using the standard control of -14, -15, and -17 percentage

points respectively (all with statistically significant diverging prior trends), compared to -11, -11, and -10 percentage points respectively (none with statistically significant diverging prior trends) using the alternative control. Matching using the nearest neighbour but including wave 1 LTH in the PSM returns an estimated wave 3 employment effect of -16 percentage points (with marginally significant diverging prior trend) using the standard control group and 11 percentage points (with no significant diverging prior trend) using the alternative control group. Excluding cohorts from the treatment group who experience disability onset during 2008 and 2009 (corresponding to the recession induced by the financial crisis) or those experiencing disability onset after 2007 returns estimated employment effects of -10 percentage points and -16 percentage points respectively using the standard control group (although in both cases diverging prior trends are smaller in magnitude and no longer statistically significant) compared to -8 percentage points and -10 percentage points respectively (with no diverging prior trends) using the alternative control group. Estimated employment effects using the alternative (activity-limiting) definition of disability are smaller in magnitude (-8 percentage points with significant diverging prior trend using the standard control group, -3 percentage points with no diverging prior trend using the alternative control group), as we might expect given a broader measure of disability.

Finally, consider Table 4 which presents key PSM estimates (for conciseness just the difference between the treatment and relevant control group in wave 3, the year of onset) from splitting the sample by broad socio-demographic group and by the nature of the disability onset. Since the sample sizes are necessarily smaller, we are pushing at the limits of the data here in terms of statistical power, so in most cases our conclusions are tentative. But there is evidence of a difference in the employment effects of disability onset by qualification level; those with higher levels of qualifications experience a much smaller drop in employment compared to those without (-8 percentage points compared to -22 percentage points using the standard control

group, -6 percentage points (not statistically significant) compared to -16 percentage points using the alternative control group). This is suggestive of larger disability onset effects for a group who may already have a weaker attachment to the labour market and who potentially face higher disability benefit replacement rates, and is also consistent with existing estimates in the literature (e.g. Polidano and Vu, 2015).

There are other comparisons where there appear to be differences in onset effects but where estimates are too imprecise to be confident, notably the difference in employment effects between disability onsets corresponding to single conditions and those corresponding to multiple conditions (a proxy for severity), and also by gender (larger for men), marital status (larger for non-married), and by LTH in wave 1 (larger for those not reporting a LTH condition in wave 1, which may be proxying for more sudden onsets). There are only small or unclear differences by age, area unemployment rate, and whether the disability onset reflects a main health condition that is mental or physical (albeit the sample for mental health is particularly small). Confirming the findings above, in all cases the estimated employment effect of disability onset is smaller when using the alternative control group than when using the standard control group.

<Table 4 around here>

5. Conclusions

This paper proposes an alternative control group (later disability onsetters) to the non-disabled control group commonly used in the dynamics of disability literature. In doing so we build on earlier papers by Fadlon and Nielsen (Fadlon and Nielsen, 2015; Fadlon and Nielsen, 2019) who use a similar approach to model the impacts of health shocks of family members, although we diverge from these earlier papers in using survey rather than administrative data, and in the nature and subjectivity of our treatment (self-reported disability onset). The alternative control

group is likely to be more similar to the treatment group in terms of unobservables than is the case for the standard non-disabled control group, and is also likely to be less susceptible to justification bias in this context, potentially mitigating biases that would otherwise inflate the estimated causal effect of disability onset on employment. An application to British longitudinal data suggests this is indeed the case, with an estimated employment effect of disability onset – an 11 percentage point decline in employment in the onset year – that is smaller than the equivalent estimate (a 16 percentage point decline) using the standard control group.

The clear methodological implication of this paper is that there is a case for adding this kind of control group to the toolkit of the applied researcher in the context of estimated disability onset effects, and perhaps also more generally in estimating treatment effects with longitudinal data in a health economics context. The advantages of this approach may be particularly pertinent when using measures of health that are potentially susceptible to reporting biases. We make this argument in the context of justification bias in reported disability here, but further research could explore this more generally, for example by exploiting the availability of a wider set of self-reported health outcomes in other surveys.

The economic implication of the paper is that, in line with the existing literature, disability onset has a substantial impact on employment in the short-run, even when considering the conservative estimates presented here. Nevertheless, in contrast to popular perception, most people experiencing disability onset in Britain are still in employment one year later.

If Britain is ever to seriously and sustainably narrow the disability employment gap – as is the government’s continued aim according to the recently announced National Disability Strategy – then, among other things, it is likely to require further intervention to reduce this short-run employment effect of disability onset, particularly given how difficult it appears to be to get

people with disability back into work once they have been out of the labour market, and perhaps in receipt of disability benefits, for an extended period of time (e.g. see Burkhauser, Daly, McVicar and Wilkins, 2014). The magnitude of the short-run employment effect of disability onset – which we show here to be sensitive to control group adopted in the modelling – is therefore likely to be an important parameter in terms of evaluating the effectiveness of interventions going forward, making the quality of such estimates particularly important. Further, there are (tentative) implications here in terms of how some such interventions might be targeted, including for example by qualification level. One possible way of implementing this could be additional support offered to help low wage workers retain employment in the event of disability onset. Further research, for example exploiting the larger sample size and longer time frame potentially available in Understanding Society, could subject these conclusions, and their policy implications, to further scrutiny.

References

- Abadie, A., and Imbens, G.W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74, 235–267.
- Baumberg, B., Jones, M. and Wass, V. (2015). Disability prevalence and disability-related employment gaps in the UK 1998–2012: Different trends in different surveys? *Social Science and Medicine*, 141: 72-81.
- Benitez-Silva, H., Buchinsky, M., Chan, HM, Rust, J. and Sheidvasser, S. (2004). How large is the bias in self-reported disability status? *Journal of Applied Econometrics* 19, 6, 649-670.
- Burkhauser, R.V., Daly, M.C., McVicar, D. and Wilkins, R. (2014). Disability benefit growth and disability reform in the US: lessons from other OECD nations. *IZA Journal of Labor Policy* 3:4.
- Charles, K.K. (2003). The longitudinal structure of earnings losses among work-limited disabled workers. *Journal of Human Resources*, 38, 618–646.
- DeLeire, T. (2001). Changes in wage discrimination against people with disabilities: 1984–93, *Journal of Human Resources*, 36, 144–58.
- Fadlon, I. and Nielsen. T.H. (2015). Household responses to severe health shocks and the design of social insurance. Working Paper No. 21352, National Bureau of Economic Research, Cambridge MA.
- Fadlon, I. and Nielsen. T.H. (2019). Family health behaviors. *American Economic Review* 109(9): 3162–3191.

- Garcia-Gomez, P. (2011). Institutions, health shocks and labour market outcomes across Europe, *Journal of Health Economics*, 30, 200-213.
- García-Gómez, P., van Kippersluis, H., O'Donnell, O. and van Doorslaer, E. (2013). Long-term and spillover effects of health shocks on employment and income. *Journal of Human Resources*, 48, 873-909.
- Jenkins, S. and Rigg, J. (2004). Disability and disadvantage: Selection, onset and duration effects. *Journal of Social Policy*, 33, 479-501.
- Jones, M.K., Davies, R. and Drinkwater, S., (2018). The dynamics of disability and work in Britain, *The Manchester School*, 86, 279-307.
- Kidd, M.P., Sloane, P.J. and Ferko, I. (2000). Disability and the labour market; an analysis of British males. *Journal of Health Economics* 19, 961–81.
- Leuven, E. and Sianesi, B. (2003). PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing, Statistical Software Components S432001, Boston College Department of Economics, revised 01 Feb 2018.
- Meyer, B.D. and Mok, W.K.C. (2019). Disability, earnings, income and consumption. *Journal of Public Economics* 171: 51-69.
- Mok, W.K.C., Meyer, B.D., Charles, K.K. and Achen, A.C. (2008). A note on ‘The longitudinal structure of earnings losses among work-limited disabled workers’. *Journal of Human Resources* 43, 3, 721–728.

Office for National Statistics, Social Survey Division. (2020). Annual Population Survey, 2004-2019: Secure Access. [data collection]. 15th Edition. UK Data Service. SN: 6721, <http://doi.org/10.5255/UKDA-SN-6721-14>.

Oguzoglu, U. (2012). Is there a better measure of self-assessed disability? *Applied Economics Letters* 19, 14, 1335-1338.

Polidano, C. and Vu, H. (2015). Differential labour market impacts from disability onset. *Health Economics*, 24, 302-317.

Rosenbaum, P.R. and Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects, *Biometrika*, 70, 41-55.

Van Soest, A., Andreyeva, T., Kapteyn, A. and Smith, J.P. (2012). Self-reported disability and reference groups. In Wise, DA. (ed.) *Investigations in the Economics of Aging*, pp. 237-264, University of Chicago Press.

Table 1: Proportions in employment by wave and treatment status

	Wave				<i>N</i> <i>(no. individuals)</i>
	1	2	3	4	
Treatment (0011)	0.747	0.725	0.628	0.583	578
Standard Control (0000)	0.848	0.856	0.858	0.853	34,889
Alternative Control (0001)	0.786	0.782	0.758	0.708	1,295

Note: Balanced panel LLFS. The sample is constrained to be the same as that in Tables 2 and 3.

Table 2: Descriptive statistics for matching variables by treatment and control groups

	Treatment (0011)	Standard Control (0000) (pre- matching)	Standard Control (0000) (post- matching)	Alternative Control (0001) (pre- matching)	Alternative Control (0001) (post- matching)
Age	44.965	41.564***	44.609	45.249	44.912
Male	0.505	0.496	0.486	0.501	0.505
<i>Marital status</i>					
Single	0.254	0.258	0.268	0.225	0.261
Married	0.593	0.634**	0.592	0.622	0.583
Divorced/separate/widowed	0.152	0.109***	0.140	0.153	0.156
<i>Highest qualification</i>					
Degree	0.151	0.229***	0.142	0.156	0.145
Other higher education	0.090	0.116*	0.116	0.110	0.071
A level	0.232	0.234	0.225	0.217	0.249
O level	0.235	0.225	0.225	0.221	0.220
Other	0.113	0.088**	0.119	0.121	0.102
None	0.180	0.108***	0.168	0.175	0.213
Dependent child	0.332	0.460***	0.343	0.351	0.363
Full-time student	0.054	0.056	0.055	0.042	0.057
<i>Region</i>					
Tyne and Wear	0.024	0.031	0.040	0.031	0.033
Rest of the North East	0.055	0.046	0.061	0.045	0.048
Greater Manchester	0.074	0.067	0.083	0.083	0.064
Merseyside	0.043	0.033	0.054	0.026*	0.038
Rest of the North West	0.036	0.033	0.035	0.044	0.038
South Yorkshire	‡	0.018	0.020	0.015	0.017
West Yorkshire	0.021	0.013	0.010	0.012	0.022
Rest of Yorkshire and Humberside	0.023	0.033	0.012	0.026	0.028
East Midlands	‡	0.019	0.010	0.018	0.014
West Midlands Metropolitan Country	0.043	0.038	0.048	0.041	0.047
Rest of the West Midlands	0.028	0.030	0.017	0.027	0.024
East of England	0.033	0.031	0.043	0.028	0.026
Inner London	0.019	0.020	0.029	0.022	0.016
Outer London	0.040	0.027	0.028	0.032	0.050
South East	0.085	0.085	0.059	0.080	0.081
South West	0.057	0.067	0.066	0.064	0.050
Wales	0.197	0.171*	0.209	0.186	0.201
Strathclyde	0.090	0.102	0.085	0.090	0.097
Rest of Scotland	0.099	0.137***	0.090	0.131**	0.105
<i>Housing tenure</i>					
Owned outright	0.251	0.207***	0.225	0.246	0.253
Mortgaged	0.505	0.659***	0.505	0.545*	0.516
Rented	0.244	0.135***	0.270	0.209*	0.249
Area unemployment rate	0.065	0.063***	0.066	0.064	0.065
<i>Year</i>					
2004	0.183	0.205	0.183	0.168	0.183
2005	0.190	0.197	0.190	0.167	0.190
2006	0.154	0.169	0.154	0.166	0.154
2007	0.149	0.160	0.148	0.197**	0.149
2008	0.173	0.137**	0.173	0.156	0.172
2009	0.151	0.132	0.150	0.146	0.150
Employment	0.747	0.848***	0.747	0.786*	0.747
<i>N (no. individuals)</i>	578	34,889	-	1,295	-

Notes: Balanced panel LLFS. All characteristics are measured at wave 1. Figures for age and area unemployment reflect averages. The remaining figures reflect proportions within the relevant group. Estimates are based on NN(1) PSM with replacement. ***, **, * denote significance at the 99%, 95% and 90% level respectively. ‘‡’ indicates cells based on less than 10 positive observations, suppressed for reasons of disclosure control.

Table 3: PSM estimates of disability onset employment effects

<i>Standard Control</i>	Treatment	Control (0000)	Difference	T-stat
Wave 1 (onset-2)	0.747	0.747	0.000	-
Wave 2 (onset-1)	0.725	0.784	-0.059	-3.26***
Wave 3 (onset year)	0.628	0.786	-0.157	-7.54***
Wave 4 (onset+1)	0.583	0.786	-0.202	-8.72***
<i>Alternative Control</i>	Treatment	Control (0001)	Difference	T-stat
Wave 1 (onset-2)	0.747	0.747	0.000	-
Wave 2 (onset-1)	0.725	0.747	-0.022	-1.15
Wave 3 (onset year)	0.628	0.739	-0.111	-4.48***

Notes: Balanced panel LLFS. Estimates are based on NN(1) PSM with replacement. ***, **, * denote significance at the 99%, 95% and 90% level respectively. Standard errors are calculated following Abadie and Imbens (2006). Timings (relative to the onset year) relate to the treatment group. The treatment group sample used in the analysis (N=578 individuals) is those on common support in both cases and includes more than 99% of treated individuals. Estimates for wave 4 are not presented for the alternative treatment group given they have experienced disability onset by this point.

Table 4: Heterogeneity in the PSM estimates of disability onset on employment, standard and alternative control groups

	Onset (wave 3) employment effect	
	<i>Standard Control</i>	<i>Alternative Control</i>
Male	-0.157	-0.109
Female	-0.119	-0.077
Older (aged 45 or above)	-0.160	-0.120
Younger (below age 45)	-0.153	-0.107
Area above median unemployment	-0.174	-0.124
Area below median unemployment	-0.137	-0.115
Married	-0.122	-0.065
Not married	-0.179	-0.110
Low qualification (GSCE equivalent or below)	-0.220**	-0.161**
High qualification (above GCSE or equivalent)	-0.077	-0.059
Mental health condition	-0.266	-0.148
Physical health condition	-0.135	-0.120
Single health condition	-0.104	-0.048
Multiple health conditions	-0.163	-0.148
Long-term health problem	-0.130	-0.073
No long-term health problem	-0.168	-0.123

Notes: Balanced panel LLFS. Personal characteristics (and long-term health) are measured in wave 1 and disability related characteristics are measured at onset. Estimates are based on NN(1) PSM with replacement and refer to the difference in employment between the treatment and control in wave 3. Bold denotes significance at the 95% level and ** denotes a significant difference from the relevant comparator group at the 95% level. Standard errors are calculated following Abadie and Imbens (2006). Estimates are based on the region of common support, which covers more than 99% of the treatment group in each case. The samples for the treatment group are necessarily smaller than in Table 3. The minimum treatment sample is for mental health (N=64).