

Local Level Incapacity Benefits Rolls in Britain: Correlates and Convergence

Duncan McVicar
Queen's University School of Management
Queen's University Belfast
University Road
Belfast BT7 1NN
Tel: +44(0)2890973297
d.mcvicar@qub.ac.uk

This Version: December 2009

JEL: I38, J23

Keywords: incapacity benefits, Local Authority Districts, disability, hidden unemployment, convergence

Abstract

There is considerable variation across British Local Authority Districts in the proportion of the working age population claiming incapacity benefits, ranging from 2% in Hart to 16% in Merthyr Tydfil and Easington in 2008. These claimant rates are strongly correlated with local unemployment rates and self-reported disability rates. But spatial differences in claimant rates for incapacity benefits were even greater a decade ago. This paper describes the extent of this claimant rate convergence and discusses possible explanations for it.

Acknowledgements: Thanks to Michael Anyadike-Danes, Steve Fothergill, Anne Green and David Webster for helpful comments on an earlier draft and to David Webster (and originally the Department for Work and Pensions) for the data on Pathways to Work roll out dates. The usual disclaimer applies.

1. Introduction

In November 2008 there were over 2.5 million working age claimants of earnings replacement disability-related benefits in Britain, or just over seven percent of the working age population. Thirty years ago there were less than one million claimants. This growth has recently led to major national-level policy reforms in the UK aimed at encouraging many of these benefit claimants back into work, including Pathways to Work (PtW) (for early evaluation evidence see ADAM et al. 2006), the replacement of existing incapacity benefits with the new Employment and Support Allowance (ESA) (for a concise review see FRANCO, 2008; for the White Paper see DEPARTMENT FOR WORK AND PENSIONS, 2008), with both intended to help achieve the overarching target of getting one million claimants off the incapacity benefits register by 2016 (see DEPARTMENT FOR WORK AND PENSIONS, 2006 p4). Similar growth has also been experienced by the US and many other OECD countries. BOUND and BURKHAUSER (1999) review the earlier US empirical literature on this growth and conclude that a combination of falling demand for lower skilled workers and the characteristics of the benefits themselves – relative generosity and screening procedures – is to blame. MCVICAR (2008) reviews the UK literature and draws a broadly similar conclusion: falling labour market demand for some groups of workers coupled with the attractiveness of incapacity benefits relative to unemployment benefits have been key drivers.¹

But this growth in disability benefit rolls has not been uniform across space within countries and wide cross sectional dispersions between regions, states and smaller spatial units have emerged

¹ Most studies in the UK and beyond reject increased prevalence of disability as a key explanation of this growth in claimant numbers.

(MCVICAR, 2006). In the US, for example, disability insurance rolls are generally higher in the South than in the rest of the country. In Britain, they are generally higher in Wales and the North, with particular 'hotspots' in old industrial and mining areas in these regions and in some inner cities. These spatial differences are very large, particularly at more local levels. Local Authority District (LAD) claimant rates in November 2008, for example, ranged from 2% in Hart (a rural area around a market town in the greater southeast) to 16% in Merthyr Tydfil and in Easington (old mining areas in Wales and the Northeast respectively).

Looking beyond national level aggregate figures to focus on more local level data on claimant rates is crucially important for at least two reasons. First and foremost, it highlights those areas that may be most in need of policy interventions to try to redress the balance between work and welfare dependency. If the recent policy reforms in the UK are to meet the government's target of getting one million claimants off incapacity benefits then it may be in these areas that much of this has to happen (FOTHERGILL and WILSON, 2007). But second, a more disaggregate focus also allows us as researchers to exploit cross sectional variation in claimant rates and other factors, and changes in such cross sectional variation over time, to try to reach a better understanding of the drivers of claimant growth over recent decades. BLACK et al. (2002) and AUTOR and DUGGAN (2003) are examples of such studies using US data at county and state levels, and they too emphasise the role of labour market conditions in driving disability insurance rolls in support of Bound and Burkhauser's earlier conclusions. The key contributions in this respect in the UK literature have come from the work of Christina Beatty, Steve Fothergill and various collaborators, which has similarly emphasised local labour market conditions as the main determinant of spatial variations in incapacity benefits claimant rates. In particular they

were the first to show that the old industrial and mining areas of Wales and the North – areas that have been most exposed to adverse structural change in the labour market over the 1970s, 1980s and 1990s – were the areas with the highest male claimant rates for incapacity benefits (e.g. BEATTY and FOTHERGILL, 1996; BEATTY et al., 2000; BEATTY and FOTHERGILL, 2005), and more recently for female claimant rates (BEATTY et al., 2009).

The first contribution of this paper is to provide an up to date (November 2008) description of LAD level claimant rates in Britain for both men and women, very much in the spirit of and showing a similar spatial picture to these earlier works. The correlations between claimant rates, disability prevalence and unemployment rates are then explored via a simple OLS cross-section regression. The paper then goes on to explore the possible joint determination of incapacity benefits rolls, unemployment rates and (self-reported) disability prevalence in this simple cross sectional framework using an instrumental variables (IV) approach. The evidence from these regressions is consistent with the prevailing view that labour market conditions are a key determinant – for both men and women – of spatial variations in incapacity benefits claimant rates. But there is evidence that disability prevalence plays a role too, at least for women. This analysis is presented in Section 3 of the paper.

The paper makes a further contribution in highlighting the *convergence* in LAD level incapacity benefits claimant rates that has taken place over the last decade or so in Britain. Although we tend to see the same claiming hotspots in 2008 that others have previously identified using earlier data, the cross sectional dispersion of claimant rates has fallen quite substantially over the period. For example, whereas 2008 claimant rates ranged from 2% to 16% as previously noted,

in 1999 claimant rates ranged from 2% (Hart again) to 23% (Merthyr Tydfil again). This convergence in British incapacity benefits rolls has only recently been noted by the UK literature – at the regional level by MCVICAR (2006), MCVICAR and ANYADIKE-DANES (2009) and ANYADIKE-DANES (2009), and at the LAD level by WEBSTER et al. (2009) – so Section 4 of this paper attempts to add a little more ‘flesh’ to the convergence ‘bone’ in terms of detail. Having established that incapacity benefit rolls have to some extent converged over the last decade in Britain, the paper goes on to consider possible reasons for why we’ve seen convergence. We explore three (not necessarily mutually exclusive) conjectures: first, that convergence has been driven by policy intervention that may to some extent have been targeted on those areas with highest incapacity benefits claimant rates in 1999; second, that convergence has been driven by convergence in disability prevalence; third, as suggested by WEBSTER et al. (2009), that convergence has been driven by the improving labour market, very much as it has driven convergence in unemployment rates over the same period. The third conjecture is the most consistent with the Fothergill and Beatty interpretation of many incapacity benefits claimants, particularly in the claiming hotspots, as being hidden unemployed.

2. Definitions and Data

There are essentially three types of disability related benefits in Britain: earnings replacement benefits, tax credits for those in work with low wages and additional costs benefits. This paper is concerned with earnings replacement or out-of-work disability benefits – what we have been calling incapacity benefits – which are intended to (at least partly) replace income from employment for those not working on grounds of disability or ill health. For the period between

1999 and 2008 the main earnings replacement incapacity benefit was called Incapacity Benefit (IB). This was a contributory benefit, i.e. eligibility required a work history. Incapacity for work was determined by government doctors by means of a Personal Capability Assessment (medical screening). IB was paid at one of three flat rates depending on the length of time the individual had been unable to work: a short term lower rate for the first 28 weeks, a short term higher rate for the next 24 weeks, and a higher long term rate subsequently. Those that become sick or disabled while in work, however, were generally ineligible for IB during the first 28 weeks of a spell out of work and instead claimed Statutory Sick Pay, for which employers were responsible. Limited means testing for new IB claimants with significant pension income was introduced in 2001. Eligibility criteria and benefit rates for IB were set nationally.

There were two forms of incapacity benefit for those unable to work because of ill health or disability that did not meet the contributions based eligibility criteria, e.g. because their disabilities prevented them building up sufficient work histories. The first – for which new claimants were no longer eligible after 2001 – was Severe Disablement Allowance (SDA). Those not eligible for IB or SDA may have been eligible to have their National Insurance credits paid (contributions towards the state pension). ‘Credits Only’ claimants usually also received means tested minimum income benefits known as Income Support, often with a ‘disability premium’. Following convention, recipients of IB, SDA and ‘Credits Only’ are collectively referred to as *incapacity benefits claimants* in this paper.

Since the 27th October 2008 IB (and Credits Only) has been replaced by ESA for *new* claimants, although *existing* claimants of IB have continued to receive it. For more details on ESA see

Franco (2008). The 2008 data used in this paper refer to November 2008, so for all but a very few claimants IB/SDA/Credits Only would still have been the relevant benefit regime.

There are 408 LADs in Britain (pre-2009 LADs). They have an average working age population of 90,000 (in 2008) but this ranges from 1,200 (Isles of Scilly) to over 600,000 (Birmingham). These are administrative areas and as such do not necessarily correspond to local labour markets. But we use them here for the two (not unrelated) reasons that data are easily available at the LAD level and that the discourse to date on spatial aspects of incapacity benefits claiming has been largely conducted at the LAD or Government Office Region level. Our working age incapacity benefits (IB, SDA, Credits Only) claimant data for these areas, separately by sex, are taken from the Department for Work and Pensions website (www.dwp.gov.uk). Data on the proportions of LAD working age populations reporting either a work-limiting disability, a disability according to the 1995 Disability Discrimination Act definition² or both are taken from the Labour Force Survey (LFS), now called the Annual Population Survey (APS), sourced from NOMIS (www.nomis.ac.uk). Administrative data on residence-based claimant count unemployment rates are also taken from NOMIS.³ Although these are the main variables considered here, we do at times use others as instruments or additional controls. These include the roll out schedule of PtW⁴, working age population proportions with higher level qualifications (NVQ4 and above) or with no qualifications, both taken from the LFS/APS, and

² The Disability Discrimination Act defines a disabled person as someone who has a physical or mental impairment that has a substantial and long-term adverse effect on his or her ability to carry out normal day-to-day activities.

³ The LFS/APS contains data on unemployment using the alternative ILO definition – out of work, wanting to work, actively looking for work in the last 4 weeks – but there are some missing values at the LAD level because the LFS/APS suppresses data where numbers are so small as to be potentially unreliable. Residence-based claimant count unemployment rates refer to the proportion of the working age resident population in receipt of unemployment benefits. Note this differs from the more standard ‘proportion of the labour force’ definition of the claimant count.

⁴ Thanks to David Webster for sending us these data, originally sourced from DWP.

LAD level life expectancy data available from the website of the Office for National Statistics (www.ons.gov.uk).

3. Incapacity Benefits Claims at the LAD Level in 2008

The middle three columns of Table 1 provide descriptive statistics on the key variables for 2008. Nationally almost six percent of women and just over seven percent of men of working age claimed incapacity benefits in 2008. LAD level claimant rates for 2008, for men and women respectively, are given by Figures 1 and 2. They show a spatial pattern of IB claiming that – following the work of Christina Beatty, Steve Fothergill and various others – is by now familiar to many, with higher claimant rates in much of the north and west, particularly old industrial and mining areas, and lower claimant rates in much of the greater southeast. This is true for both men and women, with a correlation coefficient between male and female LAD claimant rates in 2008 of .98. Although we don't present similar maps for unemployment rates and disability rates, they are highly and positively correlated with these LAD claimant rates (see Table 2): unemployment and disability prevalence also tends to be higher in the north and west, particularly old industrial and mining areas, than in the greater southeast.⁵ At the national level note that the incapacity benefits claimant rate is considerably higher than the unemployment claimant count rate, as has been the case since the mid 1990s (ANYADIKE-DANES and MCVICAR, 2008). Also note that only around a third of those that report having a disability are actually claiming incapacity benefits in 2008.

⁵ Note that female claimant rates are more highly correlated with male unemployment rates than with female unemployment rates. For further discussion of the inter-related nature of male and female local labour markets in this context see BEATTY et al. (2009).

We can further explore the relationships between incapacity benefits claimant rates, unemployment rates and disability prevalence by estimating a simple OLS regression on the 2008 data (with subscript i denoting LAD) of the following form:

$$(1) IB_i = \alpha + \beta CC_i + \gamma DIS_i + u_i$$

There are few existing examples of regressions of this type in the UK incapacity benefits literature, but NOLAN and FITZROY (2003) is one exception.⁶ They estimate an equation for claimant rates regressed on mortality rates, hospital visits, regional dummies and a selection of socio-economic indicators on three separate LAD cross sections for 1998/9, 1999/0 and 2000/1 (not separately by gender). They generally find statistically significant correlations with the expected signs, but note the omission of any indicator of labour market state in addition to the use of mortality rates to proxy for disability prevalence.⁷

The relevant regression results are presented in the first 3 columns of Table 4. As we would expect claimant rates are positively and significantly related to disability prevalence and to the unemployment rate. For both males and females a one percentage point increase in disability prevalence corresponds to a .2 percentage point increase in the claimant rate. A one percentage point increase in the claimant count unemployment rate corresponds with a one (two) percentage

⁶ An earlier but rather different draft of this current paper – MCVICAR (2007) – regressed claimant rates on self-reported disability prevalence, the ILO unemployment rate among the non-disabled, median weekly earnings and the proportion of the working age population with no educational qualifications on 2003 data. Coefficients were statistically significant taking the expected signs. Note that in contrast to McVicar (2007), equation (1) does not attempt to separate unemployment and earnings effects – both of which can be thought of as manifestations of underlying labour demand – and instead treats the unemployment rate as a single proxy for geographical variation in labour demand.

⁷ For a discussion of the pros and cons of proxying for disability with mortality rates or other ‘objective’ health measures see BOUND (1991).

point increase in incapacity benefits claims for males (females).⁸ Taken at face value – placing a possibly naïve causal interpretation on the estimated relationships – these results suggest that geographical variations in disability prevalence and in unemployment rates appear to influence incapacity benefits rolls to a similar extent, together explaining between one half and three quarters of the total variation in claimant rates.⁹

Rather than the simple causal structure implied by (1) it is perhaps more likely that incapacity benefits claimant rates, unemployment rates and self-reported disability prevalence are jointly determined. First consider self-reported disability prevalence. If there are omitted factors that partly determine both self-reported disability and incapacity benefits claiming, e.g. cultural or social identity factors, then estimating (1) by OLS is likely to impart an upward bias on γ . One example of such a bias is the so-called justification bias where, for a given type and severity of condition, those on incapacity benefits are more likely to report themselves as disabled in surveys such as the LFS/APS than others. On the other hand, if self-reported disability prevalence measures ‘true’ disability prevalence with error, then the OLS estimate of γ may be downward biased. These data problems may also impart bias on the estimate of β . BOUND (1991) provides a detailed discussion of these issues.

⁸ Given the difference in sample means the elasticities (magnitudes expressed in percentage change terms) for men and women are both around one half. These estimates are not particularly robust to replacing the claimant count unemployment rate with the ILO unemployment rate. The least missing values for this alternative unemployment rate measure are for the data on males and females together, for which the corresponding elasticities for disability prevalence and unemployment are closer to one and one quarter respectively. The estimates are, however, reasonably robust to the inclusion of regional dummies intended to capture common omitted factors across proximate LADs. When these are included the coefficients on disability prevalence fall by around a third but the coefficients on the unemployment rate are very similar. Perhaps unsurprisingly the largest regional dummy coefficients are for Wales, Scotland, the Northeast and Northwest.

⁹ Elasticities and R^2 s are broadly similar when estimated on 1999 cross section data. Further, if self-reported disability prevalence is replaced by 1997-1999 male life expectancy we get similarly plausible results, with negative and significant coefficients on life-expectancy in each case and positive and significant coefficients on unemployment rates. The unemployment rate coefficients are smaller, however, by an order of magnitude for females and for all, but not for males.

Next consider unemployment rates. Although included as a proxy for labour demand, we can only really interpret the above estimate of β as the impact of unemployment rates on incapacity benefits claiming rates rather than of the unobserved, underlying labour demand factors on incapacity benefits claiming rates, even though the latter may be more interesting from a theory point of view. But this more limited interpretation may still be beset by biases, even if unemployment rates are a ‘good proxy’ (see WOOLDRIDGE, 2002 p299). For example, if we accept, at least to some extent, the argument of BEATTY et al. (2000) that unemployment and incapacity benefits claiming are substitutes for one another – that some incapacity benefits claimants are hidden unemployed – then we might expect OLS estimates of β that don’t account for this to be downward biased. Measurement error in the unemployment rate variables may also lead to downward bias in the OLS coefficients.¹⁰ On the other hand common unobserved factors other than labour demand that may influence both unemployment rates and incapacity benefits claiming rates in a similar direction, e.g. socio-demographic factors, may impart an upward bias on β . Again, such problems may also impart bias on estimates of γ .

To try to obtain unbiased estimates of the impacts of unemployment rates and disability prevalence on incapacity benefits claimant rates equation (1) is re-estimated by Two Stage Least Squares (2SLS)¹¹, where self-reported disability prevalence and unemployment rates are instrumented by the proportions of the male and female working age population holding higher qualifications in 1999 and male and female life expectancies averaged over the 2002-2004

¹⁰ This is one possible explanation for the smaller elasticity on ILO unemployment rates compared to claimant count unemployment rates, given that the LFS/APS is based on small samples at the LAD level.

¹¹ Three Stage Least Squares (3SLS) estimates of the coefficients in (1) are identical to the 2SLS estimates which are preferred because they allow robust standard errors to be estimated.

period.¹² Existing regressions for claimant rates using British data have tended to either ignore the potential endogeneity of these variables or to omit one or both altogether (MCVICAR, 2006). Discussion of the joint determination of disability insurance rolls and labour demand measures has been more of a feature of the US literature (e.g. BLACK et al., 2002; AUTOR and DUGGAN, 2003).

The relevant 2SLS results are presented in columns 4-6 of Table 4. For males, disability prevalence has a slightly larger coefficient but is no longer statistically significant given the increase in the standard error. For females and all, the coefficients on disability prevalence double and both remain statistically significant. The implication, assuming we have confidence in the instruments used, is that the OLS coefficients were downwards biased, particularly for women, suggesting bias due to measurement error dominates justification and similar biases. Contrast these cross sectional findings with the widespread view in the literature (e.g. BOUND and BURKHAUSER, 1999) that *temporal* variation in disability prevalence has not been a major driving force of variation in claimant rates for incapacity benefits/disability insurance. The 2SLS coefficients on unemployment rates change little in magnitude and although standard errors are larger they remain statistically significant. Evidently biases on the OLS estimates are small or they cancel each other out.¹³ So the 2SLS results, which may be closer to capturing causal relationships than the OLS estimates, serve to reinforce the earlier conclusion, and the widespread belief in the literature, that geographical variations in labour markets are a key driver

¹² The system is therefore over-identified allowing over-identification testing of instrument validity (see the notes to Table 4 for more details). Hausman tests for exogeneity of disability prevalence returned the following chi-square statistics (p-values): 6.54 (.038), 29.8 (.000) and 13.6 (.001) for male, females and all respectively. Equivalent tests for exogeneity of unemployment rates return the following statistics (p-values): 35.4 (.000), 19.5 (.000) and 32.3 (.000). These tests suggest both variables should be treated as endogenous in all three cases.

¹³ These results are robust to the inclusion or otherwise of regional dummies in terms of sign and statistical significance, although the coefficients on unemployment rates are lower for women and for all by close to an order of magnitude when these dummies are included.

of geographical variations in incapacity benefits rolls. But they also suggest geographical variations in disability prevalence play a role, albeit perhaps more so for women than for men.

4. Changes Over Time in LAD Claimant Rates and Convergence

Table 1 also presents descriptive statistics on the key variables for 1999 and for the change between 1999 and 2008. The national claimant rate for incapacity benefits was higher in 1999 than in 2008, with the fall driven entirely by reduced claiming rates among men (female claimant rates have barely changed). Over the same period prevalence rates of self-reported disability have barely changed for either men or women, but unemployment rates have continued to fall from their cyclical peak in 1993. If we take these figures at face value they suggest a working age population that is stable in terms of disability prevalence but that is facing a slowly improving labour market. Indeed ANYADIKE-DANES and MCVICAR (2008) provide regression evidence at the Government Office Region level for Britain over a similar period that suggests the fall in male incapacity benefits rolls can be at least partly attributed to the improving labour market and cannot be attributed to falling disability prevalence. Simple pairwise correlations at the LAD level suggest the change in claimant rates is positively correlated with both the change in unemployment rates and the change in disability prevalence, albeit both rather weakly (see Table 3).

What is most striking about the LAD level changes in claimant rates between 1999 and 2008, however, is that the biggest falls in male claimant rates have occurred in those LADs with the highest initial claimant rates in 1999 and vice versa. The picture is similar for women, although

more areas experience growth and those that experience falls tend to experience smaller falls. The correlations between the 1999-2008 change and the 1999 level of incapacity benefits claimant rates are -.87 and -.74 respectively. So, although we mostly still see the same incapacity benefits ‘hotspots’ in 2008 as in 1999, many of these hotspots are moving closer to the national mean from above whilst much of the greater southeast, for example, is moving closer to the national mean from below.

One possible explanation for this apparent convergence is that it is no more than a statistical reversion to the mean, or an example of ‘Galton’s fallacy’.¹⁴ In other words, claimant rate hotspots in 1999 were hotspots simply because they got a ‘bad draw’ from the claimant rate distribution, and they are closer to the mean in 2008 because they got less bad draws, rather than because of any causal relationship between initial level and change over time. But we can rule this out as an explanation (or at least rule it out as the only explanation) because there has also been a large fall in the cross sectional dispersion of claimant rates over the period – the 1999 bad draws are not being replaced by different 2008 bad draws – as evidenced by the falling standard deviations shown in Table 1.¹⁵ Also most of the incapacity benefits hotspots have been hotspots for a long time.

So what are the other possible explanations for such convergence? This brings us back to our earlier conjectures. Convergence in disability prevalence may have played a role – there is a weak positive correlation with changes in claimant rates – but because the changes over the

¹⁴ See QUAH (1993) for a detailed discussion of this point in the context of the economic growth literature.

¹⁵ The growth literature parallel here is the difference between beta convergence (the negative correlation of growth rates with country initial income levels) and sigma convergence (falling cross sectional dispersion in country incomes over time).

period in disability prevalence have been so small this is unlikely to have been a major driving factor.

It is less easy to rule out the possibility that national policy interventions or area specific interventions in the existing hotspots have led to greater reductions in claimant rates in these areas compared to other areas. WEBSTER et al. (2009) provide an in depth case study of such programmes in Glasgow – ranked 7th highest LAD in 2008 in terms of incapacity benefit claimant rates – and conclude that although they have led to some reduction in claimant numbers their overall impact has been small. They reach a similar conclusion regarding differential impacts of PtW – a major national programme among other things tightening the conditionality of incapacity benefits – due to its introduction according to an area-by-area roll out between October 2003 and April 2008. But note that PtW does appear to have been introduced earlier in LADs with higher incapacity benefits claimant rates – the correlation coefficient between 1999 claimant rates and the time since introduction of PtW is 0.38 for males and 0.37 for females – so PtW *was* targeted to some extent. The change in incapacity benefits claimant rates and length of time since PtW introduction are also correlated ($\rho=-0.36$ for males and $\rho=-0.31$ for females).

Our third conjecture is that it is the improving labour market that lies behind the convergence in incapacity benefits claiming rates over the last 10 years. There are two (observationally equivalent) ways of interpreting this argument, both of which view incapacity benefits claimants, or at least *some* incapacity benefits claimants, as being hidden unemployed. First, it could be that there have been bigger increases in labour demand in the 1999 incapacity benefits hotspot areas than in other areas, e.g. because of favourable structural change or because of capital inflow to

take advantage of spare capacity and/or lower wages. But second, it could simply be that the nationally improving labour market has had a bigger impact on claimant rates in those areas with initially high claimant rates than in other areas. Here we can draw a parallel with pressures for unemployment convergence during booms in a world of downwards sticky wages and low migration: in high unemployment areas increased labour demand sucks the previously unemployed into employment; in low unemployment areas increased labour demand is more likely to drive up wages. If we accept that some incapacity benefits claimants are hidden unemployed then the same arguments predict incapacity benefits rolls convergence during booms: in high claimant areas increased labour demand sucks the more marginal claimants into the labour market; in low claimant areas few of those claiming incapacity benefits are able to work whatever the state of the labour market.

We know from the previous section that claimant rates are highly correlated with unemployment rates. We also know from Table 1 that there has been some convergence in unemployment rates over this period, although they were already low in 1999. Plus we know from Table 3 that the changes in unemployment rates and in incapacity benefits claimant rates are positively, albeit weakly, correlated. But if this hidden unemployment explanation is correct we would also expect to see some convergence in employment rates and perhaps some divergence in median earnings. Employment rates have indeed converged over the period (their cross sectional dispersion at LAD level has fallen), although these changes appear to be only weakly correlated with changes in incapacity benefits claimant rates. Median earnings have indeed diverged over the period (their cross sectional dispersion at the LAD level has increased), with earnings changes strongly correlated with changes in claimant rates in the expected direction (the smallest rises in earnings

correspond to the biggest falls in claimant rates), at least for males.¹⁶ Taken together these indicators do suggest, as argued by WEBSTER et al. (2009), that the improving labour market has been a driver of convergence in incapacity benefits claimant rates. Such convergence would therefore seem to strengthen the incapacity benefits as hidden unemployment argument of BEATTY and FOTHERGILL (1996), BEATTY et al. (2000), BEATTY and FOTHERGILL (2005) and BEATTY et al. (2009).

We can further explore these issues by estimating a simple OLS regression for the change in claimant rates over the 1999-2008 period. The 1999 claimant rate, the change in the unemployment rate, the change in disability prevalence and, following WEBSTER et al. (2009), the length of time since introduction of PtW, are the right hand side variables, giving us equation (2) below. A negative estimate of λ shows evidence of claimant rate convergence not otherwise ‘explained’ by changes in unemployment rates, disability prevalence or PtW.¹⁷

$$(2) \Delta IB_i = \varphi + \lambda IB_{i,1999} + \mu \Delta CC_i + \eta \Delta DIS_i + \psi PtW_i + v_i$$

Regressions of this form have been presented in the UK incapacity benefits literature by Beatty et al. (2009) for women and by WEBSTER et al. (2009) for men and women together. The Beatty et al. regression, for the longer period from 1981 to 2008, has the following right hand side variables: 1981 female economic activity rate, the change in male employment from 1981-2001 (to capture job competition from men) and the change in sub-region employment from

¹⁶ An alternative (equilibrium labour markets type) explanation of claimant rate convergence is that incapacity benefits claimants have been migrating from the ‘hotspots’ to the greater southeast. We would expect to see convergence in median earnings across LADs, however, if this were the dominant factor.

¹⁷ The growth literature parallel is beta convergence, and QUAH’s (1993) Galton fallacy critique applies.

1981-2006 (to capture labour market changes in surrounding districts). They find all three right hand side variables to be statistically significant, with the correct sign (negative) on male employment growth and surrounding area employment growth, with an R^2 of just over one half. The Webster et al. regression, for the same period as the current paper, includes the following right hand side variables: the 1999 incapacity benefits claimant rate, the percentage change in employee jobs from 1998-2007 (the labour demand proxy) and the length of time since the roll out of PtW in each district. All three variables are statistically significant with the correct signs, although the Pathways variable is only marginally so. The 1999 level variable takes the anticipated negative sign consistent with the convergence hypothesis.

The relevant results for estimation of equation (2) are presented in columns 1-3 of Table 5. In each case the coefficient on the 1999 claimant rate is negative and highly statistically significant, consistent with the discussion of convergence above and with the findings of WEBSTER et al. (2009). The largest coefficient in this respect is for males suggesting convergence has been stronger for males than for females over this period. The interpretation of the male coefficient is that a LAD with a one percentage point higher initial claiming rate has seen a .28 percentage point greater fall in claiming rates between 1999 and 2008 than an otherwise similar LAD with a one percentage point lower initial claiming rate. At this rate of convergence the current gap between Merthyr and Hart, for example, would be halved in just over 20 years, other things being equal. The change in the unemployment rate, the change in disability prevalence and the PtW roll out schedule are all statistically insignificant in each of the three regressions.

But just as in the 2008 levels regression, there are reasons to suspect that not all the right hand side variables in (2) can necessarily be treated as exogenous. For unemployment rates and disability prevalence the same arguments might equally apply to changes as well as to levels (see Section 3). If the PtW roll-out was targeted according to some (here) unobserved criteria, themselves correlated with trends in incapacity benefits claiming rates, then the PtW variable may also be endogenous. There may also be omitted socio-demographic or other factors which affect both the 1999 level and the change over time of claimant rates.¹⁸ Further, although we don't specify a formal dynamic model here, we can think of equation (2) as being driven by some such underlying model, in which it is possible that the lagged level claimant rate term is correlated with the error term in the change regression giving us another potential source of bias in the OLS estimates of λ . Hausman tests suggest only the 1999 claimant rate and the change in unemployment rates cannot be treated as exogenous. Equation (2) was therefore re-estimated using 2SLS¹⁹, with the 1999 male and female higher qualifications rates together with 1999 male and female working age populations used as instruments.²⁰ Neither BEATTY et al. (2009) nor WEBSTER et al. (2009) consider the possible endogeneity of the right hand side variables in their regressions.

¹⁸ The cohort hypothesis discussed below – that spatial variation in claimant rates is driven by a particular cohort of ex-miners and ex-industrial workers that has recently reached retirement age – is one possible example.

¹⁹ As for the previous section, 3SLS estimates of the coefficients in (2) are identical to the 2SLS estimates which are preferred because they allow robust standard errors to be estimated.

²⁰ The system is therefore over-identified allowing over-identification testing of instrument validity (see the notes to Table 5 for more details). Hausman tests for exogeneity of the 1999 claimant level returned the following chi-square statistics (p-values): 50.1 (.000), 56.8 (.000) and 69.2 (.000) for male, females and all respectively. Equivalent tests for exogeneity of the change in disability prevalence return the following statistics (p-values): 4.83 (.305), 3.40 (.494) and 1.09 (.897); for change in unemployment rates we have 18.0 (.001), 9.96 (.019) and 18.5 (.001); for PtW we have 1.32 (.859), 2.87 (.579) and 0.15 (.997).

The relevant results are presented in columns 4-6 of Table 5. As in the OLS case the coefficients on the 1999 claimant rates are negative and significant in all three cases, but smaller in magnitude. The change in disability prevalence is again statistically insignificant in all cases. But the change in unemployment rates now takes a positive sign and is statistically significant in all three cases. For males, a one percentage point fall in the unemployment rate over the 1999-2008 period leads to a .27 percentage point fall in the incapacity benefits claimant rate; for females a one percentage point fall in unemployment rates corresponds to a one percentage point fall in claimant rates; and for all the coefficient is in between. Evidently the OLS coefficients on the unemployment rate change were downward biased, suggesting measurement error and/or the substitutability between incapacity benefits claiming and unemployment benefits claiming are the dominant biases. As in the 2008 levels regression, the results suggest that geographical variations in labour markets, partly captured by geographical variations in unemployment rates, are a key driver of geographical variations in incapacity benefits rolls. Finally note that the PtW variable now takes a negative and statistically significant coefficient in all three cases: those areas that introduce PtW earlier have experienced a greater fall or smaller rise in claimant rates. But, as suggested by WEBSTER et al. (2009), the magnitude of this effect is small: a one month increase in PtW duration leads to just a one hundredth of a percentage point change in claimant rates.

In terms of our earlier conjectures regarding the drivers of convergence the analysis presented here suggests that trends in disability prevalence have played little or no role and the roll out schedule of PtW has played only a minor role. Of course it could be that *other* policy interventions targeted on hotspots could have added to the overall impact of policy interventions

– this would be captured by the lagged level incapacity benefits claimant rate term in (2) – but there is no straightforward way of quantifying such interventions. On the other hand there is clear evidence here that growth in labour demand – here proxied by unemployment rates that have seen larger falls in previously high-unemployment LADs than in previously low-unemployment LADs – has played a statistically and economically significant role in driving convergence in incapacity benefits rolls.

Additional labour market effects may well be captured by the lagged level incapacity benefits term along with any other possible drivers of convergence, but without better data we cannot get much further in terms of identifying them. One such potential explanation, at least for males, however, is worth brief consideration given it appears to have some proponents in policy circles. This explanation is related to the hidden unemployment hypothesis but hidden unemployment of a specific cohort – those that lost their jobs in mining and heavy industry in the massive restructuring of the 1980s and early 1990s – that has gradually been flowing off incapacity benefits into retirement (for more details of this cohort of claimants and projections of claimant rates as they retire see FOTHERGILL and WILSON, 2007). Stated simply this *cohort hypothesis* can be interpreted as follows: that convergence is entirely driven by the retirement of this cohort of ex-miners and ex-industrial workers who have remained heavily concentrated in the hotspots, and that once these claimants have all reached age 65 convergence will cease. WEBSTER et al. (2009) argue against this for the case of Glasgow – they find that off flow rates have hardly changed for over 50s but have risen for younger age groups – but they don't explore it as a possible explanation for national level convergence. Given that LAD level incapacity benefits caseloads are now reported by age on the DWP website, however, we can do so. If convergence

between 1999 and 2008 has been largely driven by the retirement of this older male cohort then we should see strong convergence in male claimant rates for over 55s and little convergence in male claimant rates for under 55s. In fact we see strong convergence for both age groups, with correlation coefficients between the change and the 1999 levels of -.92 and -.80 respectively. So, whilst it is the case that male convergence appears stronger for the older age group, it is certainly not *restricted* to the older age group. Plus, of course, we see convergence in female incapacity benefits rolls over the same period.

5. Concluding Remarks

This paper has used recent cross section data for Britain at the LAD level to highlight the big differences across space in both male and female claiming rates for incapacity benefits and the link between these spatial differences in claiming and spatial differences in unemployment and disability prevalence. Using IV analysis to try to draw causal inferences from the data suggests that differences in the strength of local labour markets are a key driving factor behind the dispersion of claimant rates for incapacity benefits. In this respect the paper is consistent with – contributing new cross section evidence to – the general view in the literature that disability related benefit rolls in Britain and elsewhere are closely related to labour market factors. But differences in disability prevalence also appear to play a role, albeit one that looks stronger for women than for men. This is less consistent with the general view in the literature that disability prevalence is not closely related to claimant rates. Time series and cross section analyses can lead us to somewhat different conclusions in this respect, although not necessarily incompatible ones.

These spatial differences in claiming rates were even greater at the end of the 1990s. This process of convergence has taken place for both men and women, although it has been stronger for men. Again using IV analysis to try to draw causal inferences from the data, we find that changes in disability prevalence between 1999 and 2008 have had no significant impact on incapacity benefits claimant rates. The introduction of PtW (and more specifically the manner of its introduction) has played a role in driving convergence, albeit a small one. The main finding, however, is that there is a strong association with changes over time in unemployment rates. These patterns in the data can be most easily reconciled if we accept the ‘incapacity benefits as hidden unemployment’ argument of Beatty, Fothergill and others. In areas of the country where the labour market was already strong prior to 1999, where both unemployment and incapacity benefits claimant rates were low, there has been little scope for ongoing labour demand growth since 1999 to lead to further falls in unemployment or incapacity benefits claiming. But in those areas where the labour market was weaker, with higher unemployment and incapacity benefits claiming prior to 1999, labour demand growth would first have acted to reduce unemployment rates over the mid-late 1990s and then acted to encourage the more marginal incapacity benefits claimants – those with less severe disabilities which we might think of as being hidden unemployed – back into the labour market. In fact it is difficult to come up with a convincing alternative explanation for the convergence that is consistent with the data presented here. So from this perspective the observed convergence in incapacity benefits claimant rates over the period actually lends further support to the hidden unemployment hypothesis

The implications of this for future claiming rates and for policy are inextricably linked. The current recession has brought the long period of labour demand growth to an end, so the labour market is unlikely to drive any further convergence in claimant rates in the near future, unless perhaps the downturn falls hardest on the greater southeast. Even if the current hotspots are hit no harder than elsewhere in the country, however, there may be renewed divergence both because the newly non-employed join a longer existing job queue and because more of those losing jobs in these areas will have been previous claimants, or will know existing claimants and therefore perhaps be less stigmatised by claiming, than in other areas. So convergence in incapacity benefits claimant rates between 1999 and 2008 is certainly not an argument for non-intervention in the hotspots. Even if the labour market continued to improve the rates of convergence suggested here would still leave LADs such as Merthyr Tydfil and Easington with four times as many claimants per head as Hart in 2030, although we should of course be wary of taking such extrapolations too seriously. It seems more likely, however, that the process of convergence will now cease and possibly even reverse. All this makes it harder to envisage how the government will meet its one million off incapacity benefits target by 2016, short of a wholesale, and not necessarily sustainable, shift from incapacity benefits to JSA.

References

ADAM, S., EMMERSON, C., FRAYNE, C. and GOODMAN, A. (2006) *Early Quantitative Evidence on the Impact of Pathways to Work Pilots*, Research Report no. 354, London: Department for Work and Pensions.

ANYADIKE-DANES, M. (2009) What is the problem exactly? The distribution of Incapacity Benefit claimants' conditions across the British regions. Mimeo, Economic Research Institute of Northern Ireland, Belfast.

ANYADIKE-DANES, M. and MCVICAR, D. (2008) Has the boom in Incapacity Benefit claimant numbers passed its peak? *Fiscal Studies*, **29**, 4, 415-434.

AUTOR, D.H. and DUGGAN, M.G. (2003) The rise in the disability rolls and the decline in unemployment. *Quarterly Journal of Economics* **118**, 157-206.

BEATTY, C. and FOTHERGILL, S. (1996) Labour market adjustment in areas of chronic industrial decline: the case of the UK coalfields. *Regional Studies* **30**, 627-40.

BEATTY, C., FOTHERGILL, S. and MACMILLAN, R. (2000) A theory of employment, unemployment and sickness. *Regional Studies*, **34**, 617-30.

BEATTY, C. and FOTHERGILL, S. (2005) The diversion from 'unemployment' to 'sickness' across British regions and Districts. *Regional Studies*, **39**, 837-54.

BEATTY, C., FOTHERGILL, S., HOUSTON, D., POWELL, R. and SISSONS, P. (2009) *Women on Incapacity Benefits: A Statistical Overview*. Centre for Regional Economic and Social Research, Sheffield Hallam University, Sheffield.

BLACK, D., KERMIT, D. and SANDERS, S. (2002) The impact of economic conditions on participation in disability programs: Evidence from the coal boom and bust. *American Economic Review* **92**, 27-50.

BOUND, J. and BURKHAUSER, R.V. (1999) Economic analysis of transfer programs targeted on people with disabilities, in O. ASHENFELTER and D. CARD (eds) *Handbook of Labor Economics* Volume 3C, 3417-3528 (Elsevier: North Holland).

BOUND, J. (1991) Self-reported versus objective measures of health in retirement models. *Journal of Human Resources*, **26**, 106-38.

DEPARTMENT FOR WORK AND PENSIONS (2006) *A New Deal for Welfare: Empowering People to Work*. The Stationary Office, Norwich.

DEPARTMENT FOR WORK AND PENSIONS (2008) *No One Written Off: Reforming Welfare to Reward Responsibility*. The Stationary Office, Norwich.

FOTHERGILL, S. and WILSON, I. (2007) A million off incapacity benefit: how achievable is labour's target? *Cambridge Journal of Economics*, 171:1007–1023.

FRANCO, A. (2008) The economic impact of the introduction of Employment and Support Allowance. *Benefits*, 16, 3, 279-85.

MCVICAR, D. (2006) Why do disability benefit rolls vary between regions? A review of evidence from the US and UK. *Regional Studies*, 40, 5, 519-533.

MCVICAR, D. (2007) What explains spatial variation in British disability benefit rolls? Mimeo, Queen's University Belfast, Belfast.

MCVICAR, D. (2008) Why have UK disability benefit rolls grown so much? *Journal of Economic Surveys*, 22, 1, 114-139.

MCVICAR, D. and ANYADIKE-DANES, M. (2009) Panel estimates of the determinants of British regional male incapacity benefits rolls 1998-2006. *Applied Economics* (forthcoming).

NOLAN, M. and FITZROY, F. (2003) Inactivity, sickness and unemployment in Great Britain: Early analysis at the level of Local Authorities. Mimeo, University of Hull.

QUAH, D. (1993) Galton's Fallacy and tests of the convergence hypothesis. CEPR Discussion Paper No. 820, London.

WEBSTER, D., ARNOTT, J., BROWN, J., TUROK, I., MITCHELL, R. and MACDONALD, E.B. (2009) Falling Incapacity Benefit claims in a former industrial city: policy impacts or labour market improvement? Forthcoming *Policy Studies*.

WOOLDRIDGE, J.M. (2002) *Introductory Econometrics; A Modern Approach (2nd Edition)*. Thomson South-Western: Ohio.

Figure 1: Male IB/SDA Claimant Rate by Local Authority District, 2008

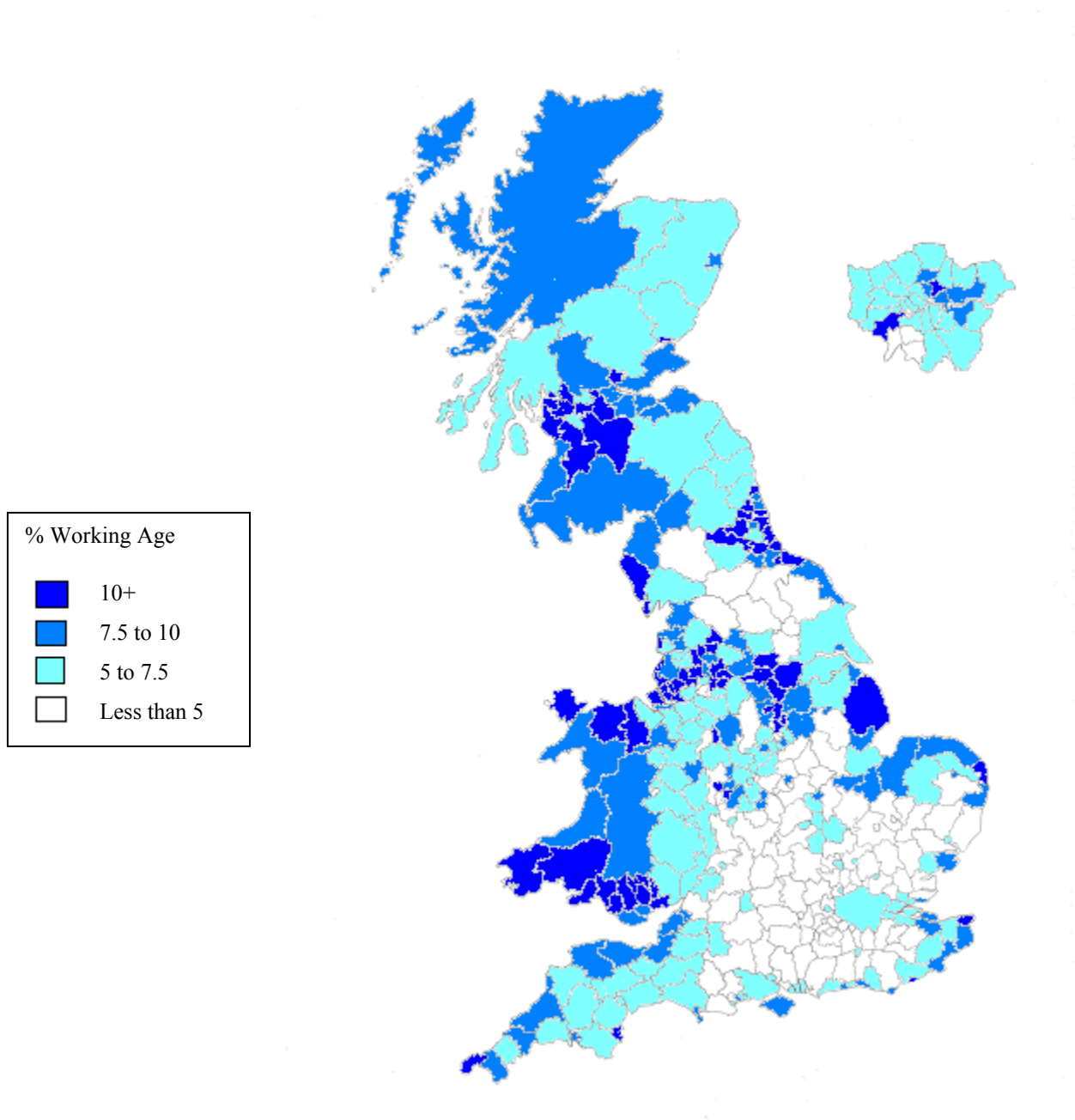


Figure 2: Female IB/SDA Claimant Rate by Local Authority District, 2008

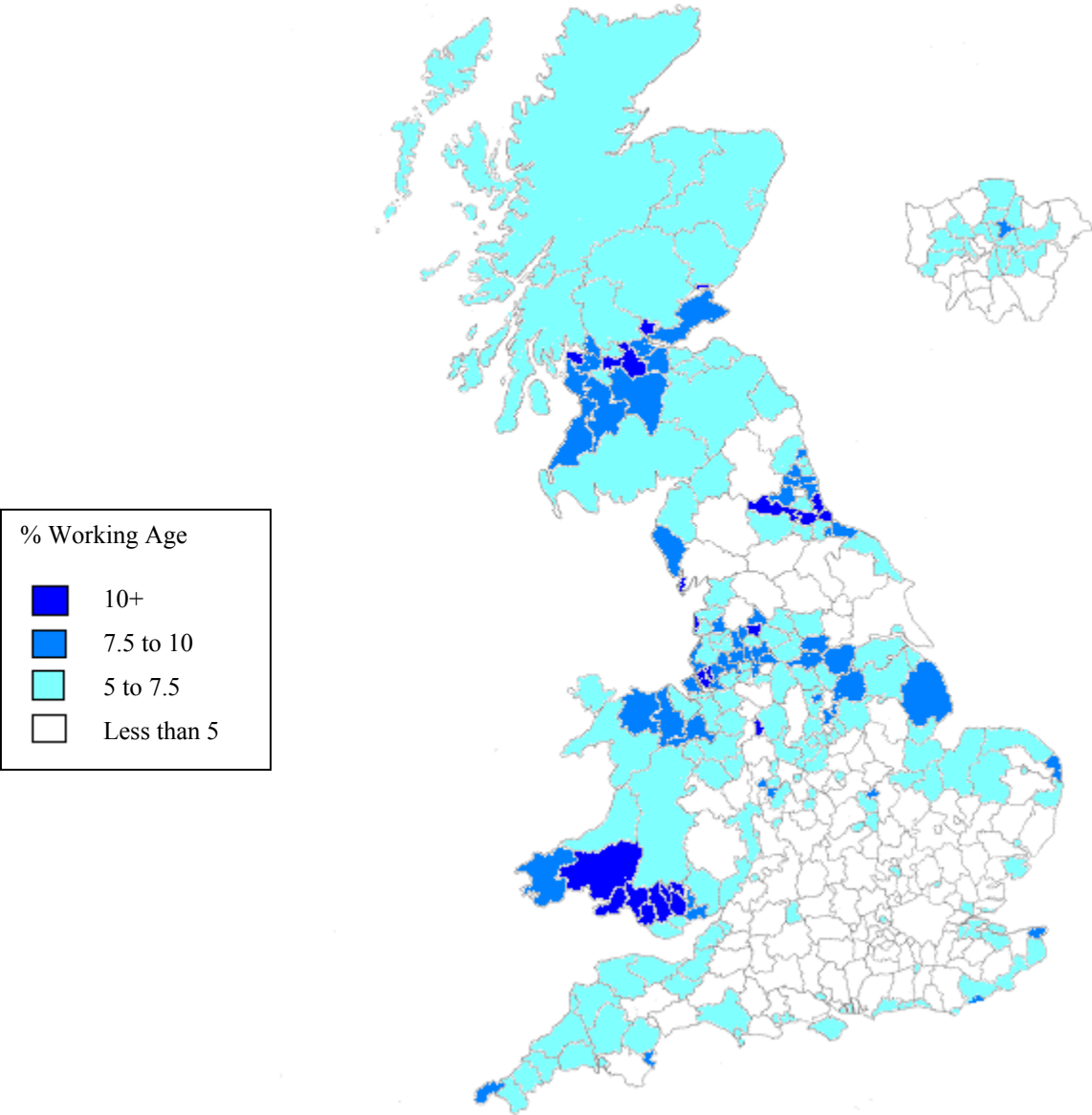


Figure 3: Change in Male IB/SDA Claimant Rate by Local Authority District, 1999-2008

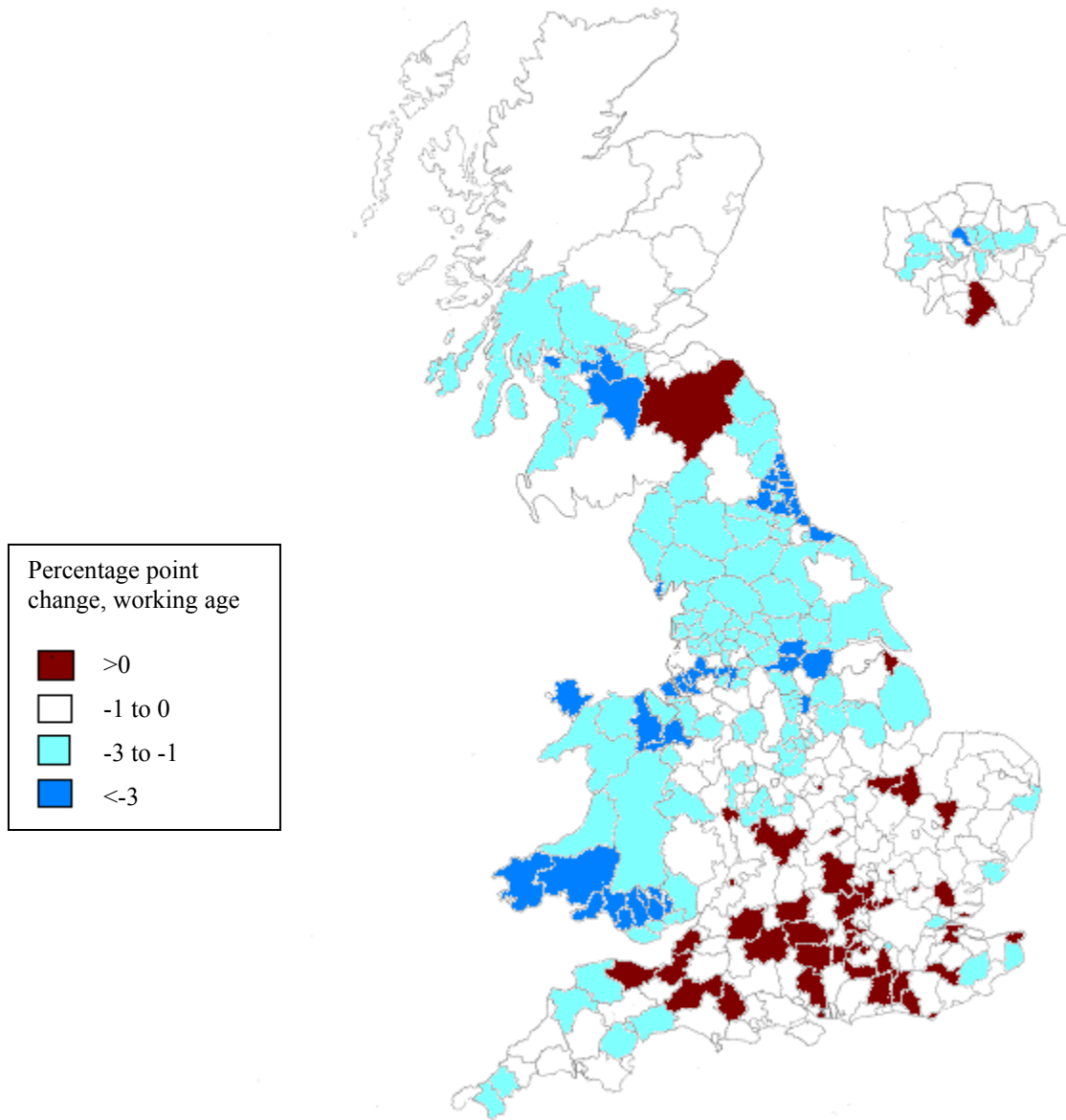


Figure 4: Change in Female IB/SDA Claimant Rate by Local Authority District, 1999-2008

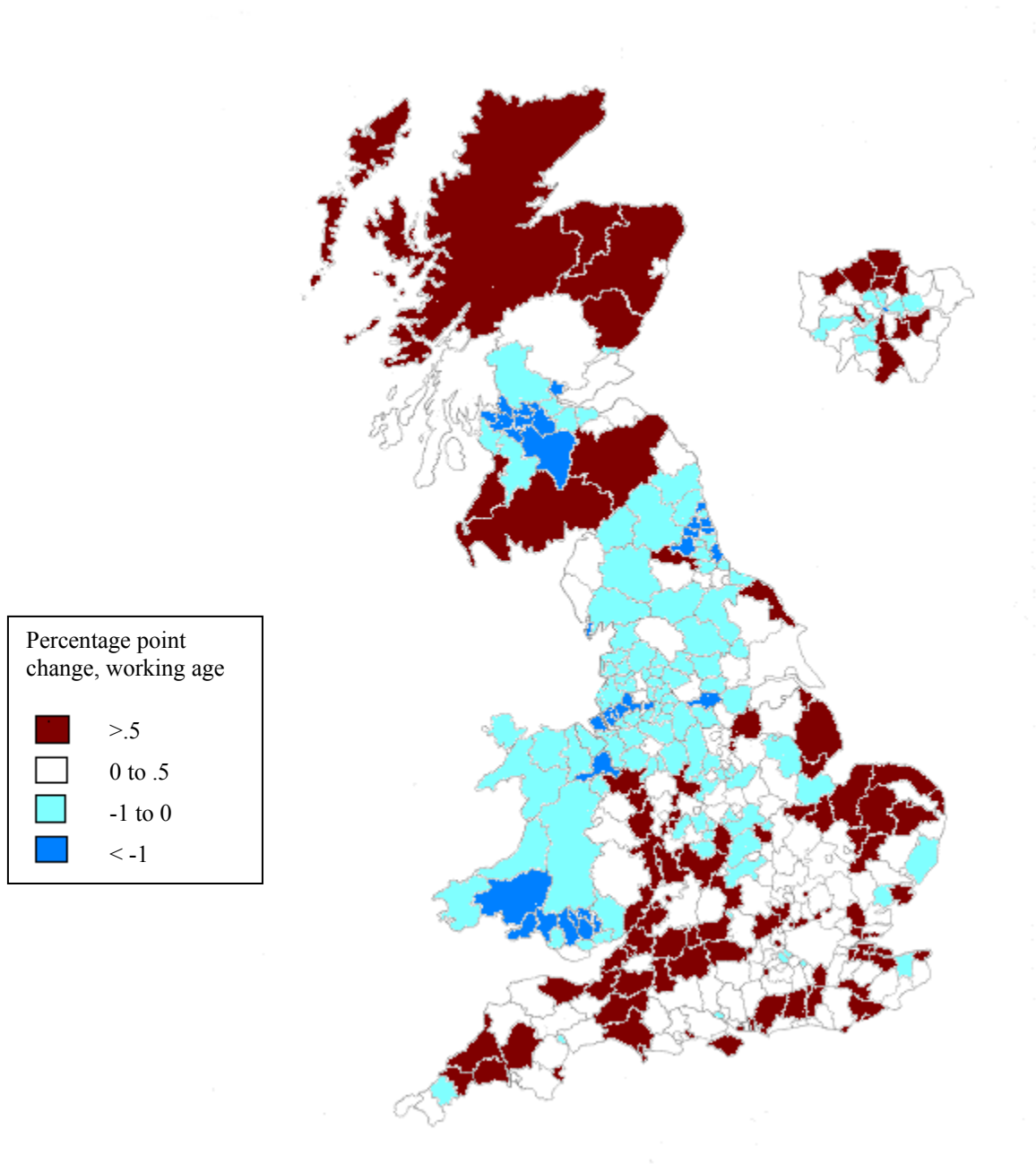


Table 1: Sample Means (Standard Deviations)

	Males 1999	Females 1999	All 1999	Males 2008	Females 2008	All 2008	Males Change	Females Change	All Change
IB claimant rate, proportion of working age population	.084 (.043)	.058 (.028)	.071 (.036)	.072 (.032)	.058 (.023)	.066 (.027)	-.012 (.014)	.0004 (.008)	-.006 (.011)
Disabled rate, proportion of working age population	.183 (.052)	.176 (.050)	.181 (.046)	.187 (.052)	.179 (.046)	.183 (.044)	.004 (.054)	.003 (.053)	.003 (.043)
Claimant Count unemployment rate	.040 (.023)	.014 (.007)	.028 (.015)	.035 (.016)	.014 (.005)	.025 (.011)	-.005 (.012)	-.0005 (.004)	-.003 (.008)

Source: Author's calculations using data from NOMIS and DWP.

Table 2: Pairwise Correlations, 2008

	IB female	Dis male	Dis female	CC male	CC female
IB male	.98	.57	.60	.80	.66
IB female		.60	.59	.79	.65
Dis male			.55	.38	.28
Dis female				.43	.35
CC male					.92

Table 3: Pairwise Correlations, Changes 1999-2008

	Δ IB female	IB male [-1]	IB female [-1]	Δ Dis male	Δ Dis female	Δ CC male	Δ CC female	Pathways
Δ IB male	.88	-.87	-.85	.14	.10	.28	.17	-.36
Δ IB female		-.69	-.74	.12	.11	.08	.04	-.31
IB male [-1]			.98	-.11	-.07	-.34	-.19	.38
IB female [-1]				-.11	-.07	-.26	-.13	.37
Δ Dis male					.29	.12	.08	-.03
Δ Dis female						-.02	-.07	.08
Δ CC male							.84	-.08
Δ CC female								-.08

Table 4: Determinants of 2008 Claimant Rates

	Males OLS	Females OLS	All OLS	Males 2SLS	Females 2SLS	All 2SLS
Disabled rate	.176* (.019)	.197* (.020)	.233* (.022)	.214 (.205)	.457* (.150)	.404* (.174)
Claimant count rate	1.28* (.070)	2.09* (.170)	1.47* (.091)	1.47* (.304)	1.96* (.741)	1.43* (.395)
Constant	-.005 (.003)	-.006 (.003)	-.014* (.003)	-.020 (.028)	-.051* (.017)	-.045* (.022)
R ²	.724	.570	.727	.700	.283	.651
Observations	407	406	407	365	365	365

Notes: Robust (Huber/White) standard errors in parentheses. * denotes statistically significant at 5%. The 2SLS estimates instrument for self-reported disability prevalence and unemployment rates with male and female life expectancy and the 1999 proportions of the male and female working age population with no qualifications. The relevant Hansen-Sargan test statistics (p-values) are 2.87 (.239), 2.98 (.225) and 1.55 (.461) for male, females and all respectively. The F statistics (p-values) for the significance of the excluded instruments in the first stage regressions for disability prevalence are 19.5 (.000), 24.4 (.000) and 27.2 (.000) for males, females and all respectively. The associated F statistics for the first stage regressions for unemployment rates are 171.2 (.000), 107.9 (.000) and 169.0 (.000). Centered R²s reported for 2SLS. Variables are expressed as proportions so coefficients can be interpreted as the percentage point change in the incapacity benefits claimant rate of a one percentage point change in the relevant explanatory variable.

Table 5: Determinants of Change in IB Claimant Rates, 1999-2008

	Males OLS	Females OLS	All OLS	Males 2SLS	Females 2SLS	All 2SLS
1999 IB claimant rate	-.278* (.014)	-.209* (.016)	-.251* (.015)	-.192* (.020)	-.095* (.032)	-.146* (.023)
ΔDisabled rate	.012 (.007)	.005 (.005)	.009 (.008)	.011 (.007)	.016 (.009)	.009 (.011)
ΔClaimant count rate	-.002 (.038)	-.113 (.075)	-.061 (.046)	.271* (.119)	1.14* (.415)	.465* (.187)
Pathways to Work	-.00003 (.00002)	-.00002 (.00002)	-.00002 (.00002)	-.0001* (.00003)	-.00006* (.00003)	-.0001* (.00003)
Constant	.012* (.001)	.013* (.001)	.012* (.001)	.008* (.001)	.007* (.001)	.007* (.001)
R ²	.754	.546	.692	.678	.078	.497
Observations	371	373	373	371	373	373

Notes: Robust (Huber/White) standard errors in parentheses. * denotes statistically significant at 5%. The 2SLS estimates instrument for the 1999 claimant level and the change in unemployment rates with the 1999 male and female working age populations and the 1999 proportions of these populations with higher qualifications. The relevant Hansen-Sargan test statistics (p-values) are .425 (.808), 4.41 (.110) and 1.98 (.371) for male, females and all respectively. The F statistics (p-values) for the significance of the excluded instruments in the first stage regressions for the 1999 claimant level are 35.1 (.000), 28.4 (.000) and 32.0 (.000) for males, females and all respectively. The associated F statistics for the first stage regressions for change in unemployment rates are 10.3 (.000), 7.49 (.000) and 10.5 (.000). Centered R²s reported for 2SLS. Coefficients can be interpreted as the percentage point impact on the change in incapacity benefits claimant rates of a one percentage point change in the relevant explanatory variable, with the exception of the PtW variable for which the coefficient can be interpreted as the percentage point impact on the change in incapacity benefits claimant rates of a one month increase in the duration of PtW. Incomplete coverage of the PtW variable cuts the sample size down from 400+ in each case.