Administrative Data and Economic Policy Evalution

Keynote Address, Australian Economists' Conference

Adelaide, 28 September 2009

Lorraine Dearden

Institute of Fiscal Studies and

Institute of Education, University of London

Email: ldearden@ifs.org.uk

Introduction

- Administrative data is increasingly being used to evaluate policy initiatives and to measure effectiveness of public service provision
  - E.g. school and hospital league tables in UK, evaluating government policy initiatives
  - Topical debate in Australia at the moment - should you have school league tables based on new NAPLAN tests?

Introduction (cont)

- BUT administrative data typically doesn't have extensive background information on students, patients , welfare recipients etc which also impact on outcomes of interest
  - Scandinavian Administrative data exception here
  - Means typically have to be very careful using this data to inform public policy and evaluating public sector performance

Outline of talk

- Focus on 3 examples of how administrative data has been used in the UK with varying degrees of success
  - Examples not randomly chosen – involve areas where I have current research interest BUT
- The examples I hope show why:
  1. Holders of administrative data should allow researchers to access this data
  2. Administrative data cannot and should not be used to answer all questions
     - and indeed one generally needs either very rich administrative data and/or some natural experiment coupled with a sound methodological approach to get the right answer

My examples

1. Using school and university administrative data to look at the impact of month of birth/length of schooling on educational outcomes

2. Using School Administrative Data to measure school performance

3. Using DWP administrative data to look at whether ethnic minorities receive parity in treatment from Job Centre Plus staff (equivalent of Centrelink staff)

My examples

1. Using school and university administrative data to look at the impact of day of birth/length of schooling on educational outcomes

2. Using School Administrative Data to measure school performance

3. Using DWP administrative data to look at whether ethnic minorities receive parity in treatment from Job Centre Plus staff (equivalent of Centrelink staff)

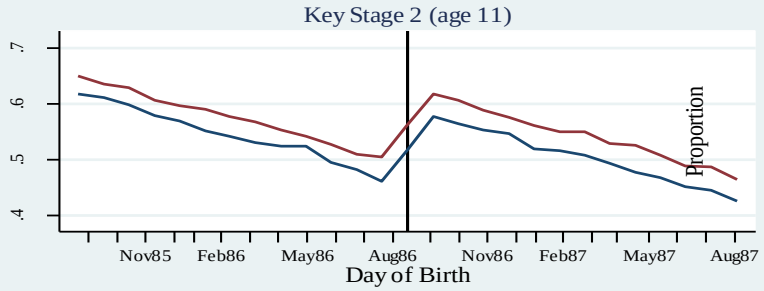
Background

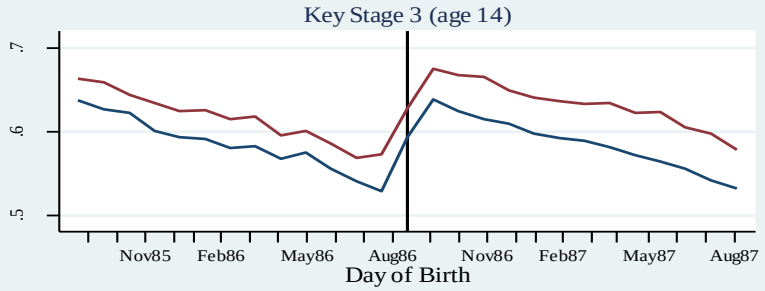
- This is joint work with Claire Crawford and Costas Meghir and funded by the DCSF in UK
- Children in England must have started school by the beginning of the term after they turn five
- Local Authorities (LAs) are free to set admissions policies within this framework
  - Single entry point, 2 entry points or 3 entry points
- Academic year: 1st September to 31st August
  - Expect children born at the end of the academic year to perform more poorly than children born at the start of the academic year – but for how long?

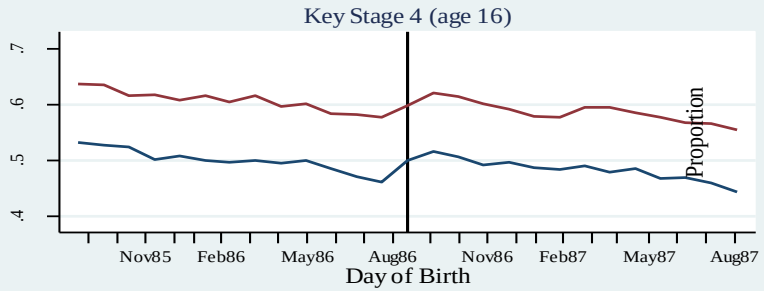
School, FE and HE linked data

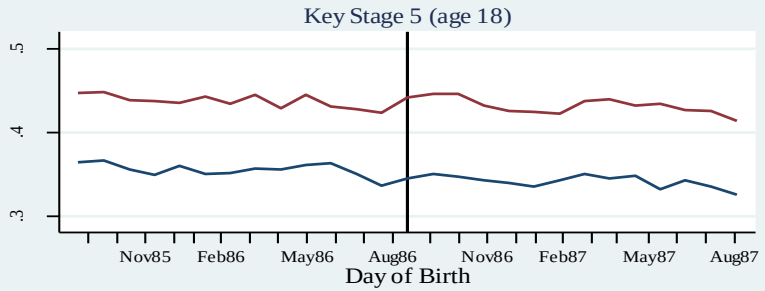
- For two cohorts of children, we have linked school, FE and HE data which has academic outcomes at age 11, 14, 16, 18 and 19/20
- Our sample includes children in state schools for whom we have academic results up to the end of compulsory schooling (age 16)
- Post 16, we see whether they gained a level 3 qualification at 18 (pre-requisite for university entrance)
- Whether they started university by ages 18/19 or 19/20

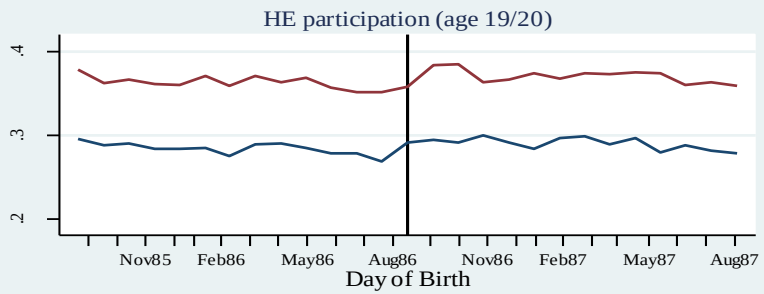
Raw differences (proportion getting expected level)











Background

- Why might this be?
  - Age of sitting the test (absolute age) effect
    - They are younger when they sit the tests
  - Age of starting school effect
    - They start school at a younger age
  - Length of schooling effect
    - They receive less schooling prior to the test
  - Age position (relative age) effect
    - They are the youngest relative to others in their class

Previous research

- Children born at the end of the academic year do perform worse:
  - e.g. Puhani & Weber (2005), Bedard & Dhuey (2006), etc
- Some studies attempt to disentangle the effects of these four factors:
  - But only for post-compulsory schooling outcomes
    - Fredriksson & Ockert (2005)
    - Black, Devereux & Salvanes (2008)
  - More difficult for compulsory schooling outcomes
    - Age at test = age at starting school + length of schooling

Our contribution

- Regional variation in admissions policies and regression discontinuity allows us to break this linear relationship *for compulsory schooling outcomes*
  - Children born on the same day (who sit tests at the same age) start school at different ages
  - Children born on virtually the same day (1 September or 31 August) who sit tests at different ages
- We can separately identify:
  - Absolute age effect
  - Age of starting school/length of schooling effect
  - Age position effect

How large is the August birth penalty?

		Reached expected level at KS2 (age 11)	Reached expected level at KS3 (age 14)	Reached expected level at KS4 (age 16)	Reached expected level at KS5 (age 18)	College participation (age 19/20)
Boys	August	-0.158** [0.004]	-0.104** [0.003]	-0.068** [0.003]	-0.016** [0.003]	-0.014** [0.003]
	September (base)	60.2	64.2	53.1	36.0	29.7
Girls	August	-0.151** [0.004]	-0.089** [0.003]	-0.057** [0.003]	-0.013** [0.003]	-0.015** [0.003]
	September (base)	63.8	67.3	63.1	44.7	38.1

Notes: ** indicates significance at 1% level; * at 5% level. Standard errors are corrected for clustering at school level.

Model also includes controls for individual characteristics (including ethnicity, language, low income indicator) and neighbourhood characteristics (particularly area-based deprivation measures), plus cohort dummies and school fixed effects.

Summary

- August-born children experience significantly poorer education outcomes than September-born children
- Almost entirely due to differences in the age at which they sit the tests
- Starting school earlier/having more terms of school is marginally better for August born children at younger ages (not shown)

The policy dilemma

- Results presented emphasise August birth penalty, but findings also apply more generally
  - On average, the younger you are the worse you do
- Ideally need to create a level playing field for all children *regardless of date of birth*
  - But also need to have school years, so someone will always be the youngest
- Points to need to age normalise tests as day of birth has long lasting impacts
- Long-term consequences suggest should not do nothing

Australia could provide some more answers

- 8 different education systems with different starting ages and lengths of schooling before sitting NAPLAN tests in years 3, 5, 7 and 9
  - In Queensland have switch from 12 to 13 years of education and switch in school starting age in 2007
- Analysing Australian data could separate not only age of sitting test effects but length of schooling versus age of starting school effects (cf UK)
- Age adjustment problematic as for some ages will have all States represented and for other ages only one state (except if non-standard entry)
- May be able to use NAPLAN data to inform on best system and move to common system Australia wide?

But

- Researchers need to have access to Australia wide data
- UK has got that absolutely right – National Pupil Database widely available to researchers
  - this has been beneficial to policy makers and researchers and has helped inform government policy

My examples

1. Using school and university administrative data to look at the impact of month of birth/length of schooling on educational outcomes

2. Using School Administrative Data to measure school performance

3. Using DWP administrative data to look at whether ethnic minorities receive parity in treatment from Job Centre Plus staff (equivalent of Centrelink staff)

Hot topic in Australia at moment..

- Should NAPLAN data be used to measure school effectiveness?
- My view is that NAPLAN national data should be open to all researchers and if you do this then they will be able to construct their own league tables
- But it is very difficult to do well as the UK experience shows, and rankings are extremely sensitive to the methodological approach taken
- My view is that data is much better to look at other issues, and the current UK current approach is highly misleading and published tables of limited use
  - But it could be done better…….

The UK experience

- Started by publishing raw results by school at Key Stage 2 (end of primary school), Key Stage 4 (end of year compulsory schooling – GSCE results) and Key Stage 5 (end of school – A level results)
- Clearly raw results by school affected by background of students attending that school and therefore not necessarily a measure of the value added by school
- Moved to also publishing value added (VA) and then contextualised value added (CVA)
- VA used baseline measures of performance and CVA controls for other factors such as gender, ethnicity, special educational needs status, whether receiving free school meals and local area measures of deprivation

UK League tables

- At moment UK school league tables consist of:
  - Proportion getting expected level at each age
  - Average points score (capped and uncapped)
  - Contextualized value added (1000 = expected level) – meant to measure the value added by the school (after stripping out the effects of background and prior ability)
    - Primary school no proper baseline for CVA (take Y2 as baseline even though do testing in Reception)

Does this help inform parental choice?

- One of the main drivers for school league tables is to identify failing schools, to set the right incentives for schools, and to aid parents in school choice
- League tables as they stand don't do any of these things particularly well
- Why?
  - Differences in raw scores are largely driven by differences in background and prior ability rather than the school
  - Even CVA, which takes account of children's background and prior ability, doesn't tell a parent whether the school adds value for their child
    - unless they have the average characteristics of a child attending that particular school and/or the school adds the same value to every child, regardless of background

Example

- Work I am doing with Marcello Sartarelli and Anna Vignoles at the IOE
- Very preliminary at this stage – but for illustrative purposes focus on 4 secondary schools in Essex to illustrate our approach and the problems with current league tables in UK
  - Newport Free Grammar School (NF)
  - Saffron Walden County High School (SW)
  - William de Ferrers School (WF)
  - Chelmer Valley High School (CV)

League Tables for these schools 2006

School	APS (uncapped)	5 A*-C	5 A*-C incl English & Maths	CVA	CVA LCI	CVA UCI
NF	400.6	65%	58%	985.2	975.3	995.1
SW	419.2	75%	65%	1010.9	1004.0	1017.8
WF	370.8	70%	63%	1000.6	994.4	1007.8
CV	376.8	66%	51%	1000.2	991.0	1009.4

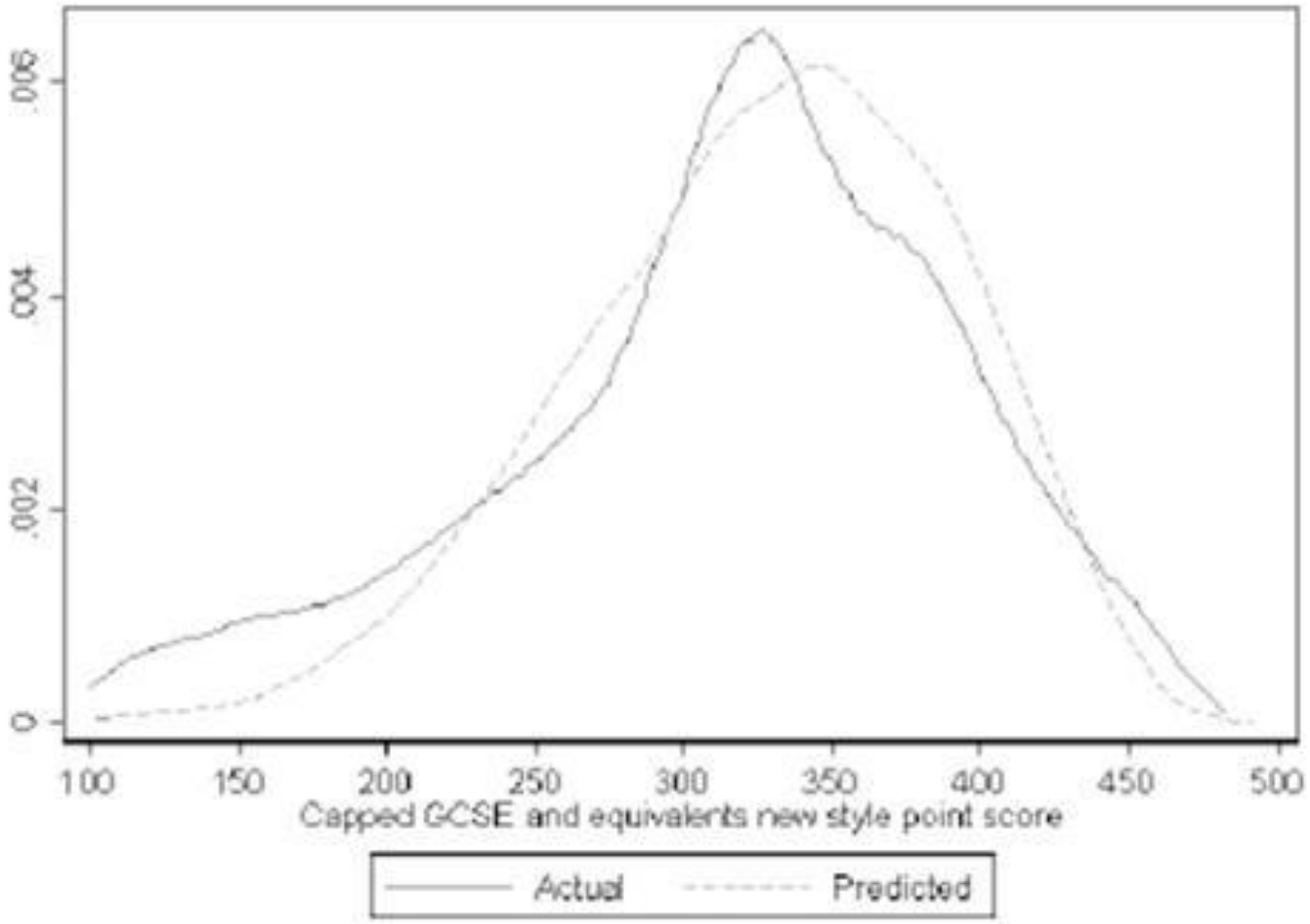
Our Approach

- Assume child has these 4 schools as choices
- Given their prior achievement and characteristics which school should they choose?
- Use some heavy econometrics (which I will skip) but does involve using GMM and relying on Conditional Independence Assumption to see which is best school given prior achievement and background characteristics (following ideas in Dehejia (2005), Fröhlich (2008), Lechner and Smith (2007)).
- Don't do pairwise comparisons between schools but multiple comparison with best (MCB) see Hsu (1996)

What we find

- If in 25th centile of Maths distribution at KS2 and average at English and Science then should go to any school but CV (with SW the best)
- If in 75th centile of Maths distribution at KS2 and average at English and Science then should go to any school but NF (with WF the best)
- Show that by reallocating 51% of students between these 4 schools could improve GCSE scores by 13 points or 14% of a standard deviation.
  - Constrained so enrolments do not change – means second best school sometimes
  - Gains are at the bottom of the distribution and 60th/70th centile of the distribution. This is where the real difference in the 4 schools lie.

Mean vs Predicted GCSE scores 4 Essex Schools



Still lots to do

- But our results suggest even with years of school league tables informing parental choice that people get decisions wrong/allocation is not optimal
- But coming up with understandable, transparent ways of doing these league tables that are informative and a true reflection of school effectiveness is very difficult
- As a starting point need rich baseline data from when child first enters primary and secondary school
  - Don't have that in UK (and Australia?)

My examples

1. Using school and university administrative data to look at the impact of day of birth/length of schooling on educational outcomes

2. Using School Administrative Data to measure school performance

3. Using DWP administrative data to look at whether ethnic minorities receive parity in treatment from Job Centre Plus staff (equivalent of Centrelink staff)

Previous work

- Considerable literature about effect of ethnicity on employment probability for JCP customers
  - e.g. Moody (2000) for NDYP, McArdle (2001) for ND25plus
- Typically aim to find degree of parity in outcomes between different ethnic groups
- But:
  - Simple regression techniques
  - Small sample sizes in survey-based studies

This paper

- Joint work with Crawford, Mesnard, Shaw and Sianesi at IFS
- Ethnic parity:
  - No difference on average between Ethnic Minority and "otherwise identical" White entering the same JCP office and accessing same program/benefit
- Our aim:
  - Get as close as possible to "otherwise identical" White and see what difference remains
  - Calculate results for a range of JCP benefits and programs

Programs and Benefits

- Incapacity benefit (IB): paid to individuals who are assessed as being incapable of work and who meet certain National Insurance contributions conditions.
- Income support (IS): a benefit for individuals on low income; usually claimants are lone parents, sick or disabled, or carers.
- Jobseeker's allowance (JSA): a benefit paid to individuals of working age who are unemployed, or who work fewer than 16 hours per week and are looking for full-time work.
- New Deal for Lone Parents (NDLP): a voluntary programme whose aim is to encourage lone parents to improve their work prospects and help them into work.
- New Deal for individuals aged 25 plus (ND25plus): a programme to help unemployed individuals aged 25 and over to find and keep a job. Participation is compulsory for individuals who have been claiming JSA for at least 18 of the previous 21 months.
- New Deal for Young People (NDYP): similar to ND25plus except that it is targeted on individuals aged 18-24. Participation is compulsory for those who have been claiming JSA for at least six months.

Controlling for selection

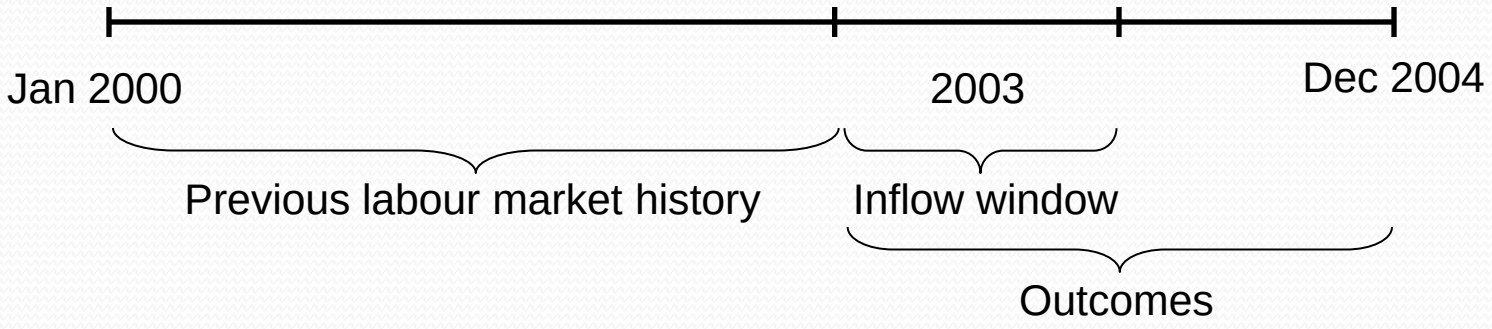
- Control for differences in *observed* characteristics between ethnic groups that may affect outcomes
- Data:
  - Detailed labour market histories
  - Individual background characteristics
- Methods:
  - Primarily propensity score matching (PSM)
  - Also regression-based methods and conditional difference in differences (DID)
- Previous LM history may have been affected by discrimination but nothing we can do about this

Sampling frame

- Sample selected on *inflow* into programme
  - Addresses differential selection off programme

Sampling frame

- Sample selected on *inflow* into programme
  - Addresses differential selection off programme
- Inflow window is 2003, allowing:
  - 3-year pre-inflow labour market history
  - 1-year follow-up



Outcomes of interest

- Two dimensions of labour market status
  - In employment (15+ days in the month)
  - On benefit (15+ days in the month)
- Benefit definition includes:
  - IS, IB, JSA, New Deal options, Basic Skills and Work-Based Learning for Adults
- Measured monthly

Data

- Primarily Work and Pensions Longitudinal Study (WPLS)
  - Benefit and employment spells for anyone on a DWP benefit since mid-1999
  - Also contains limited demographics including sex, DOB, ethnicity and postcode
- Also used National Benefit Database (NBD) and census information

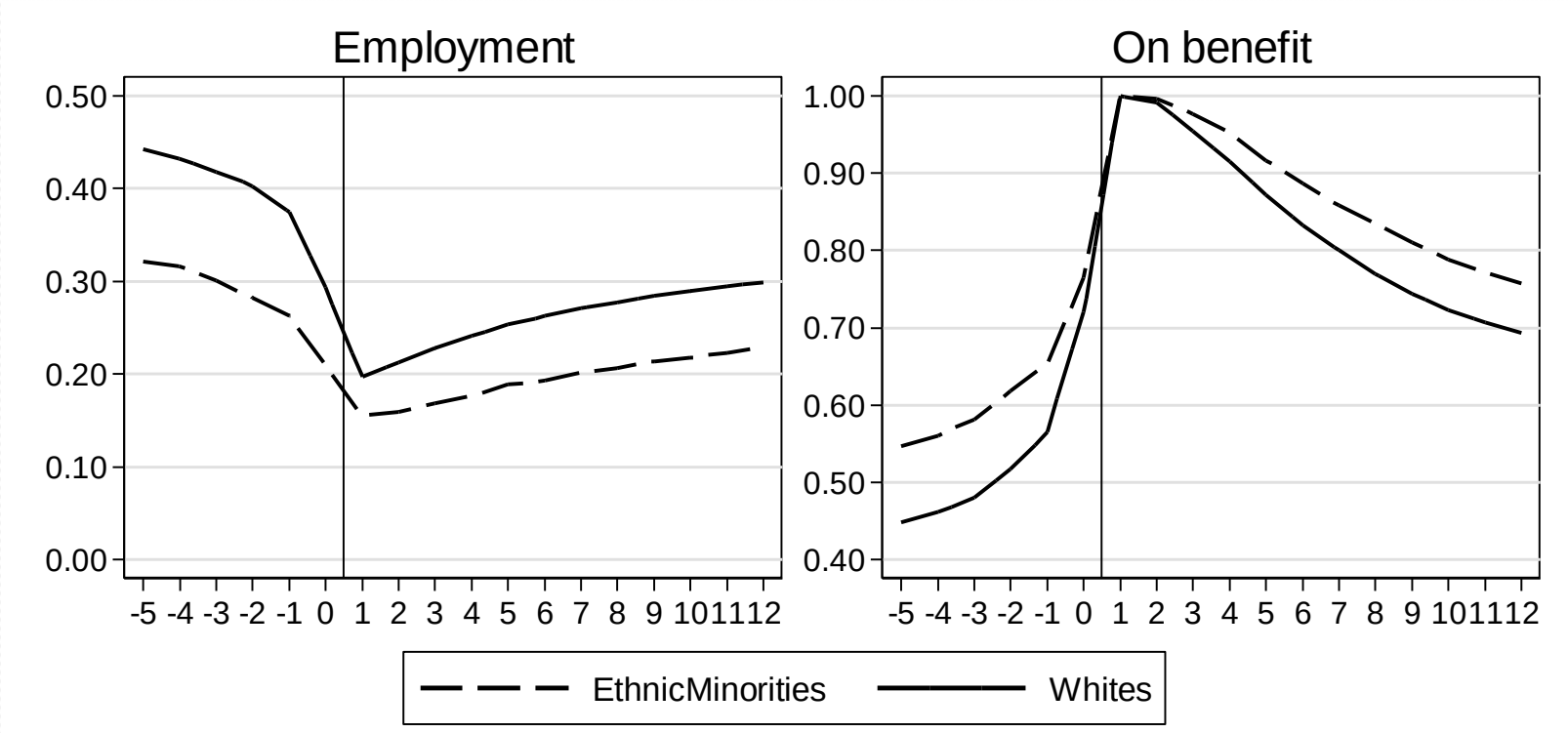
X variables

- Employment and benefit history
- Past participation in voluntary programmes
- Past participation in Basic Skills
- Individual characteristics
  - Gender, age, month of inflow
  - Proxies for education and wealth (from census)
- Local area characteristics (region, travel-to-work-area unemployment)
- Other programme-related information

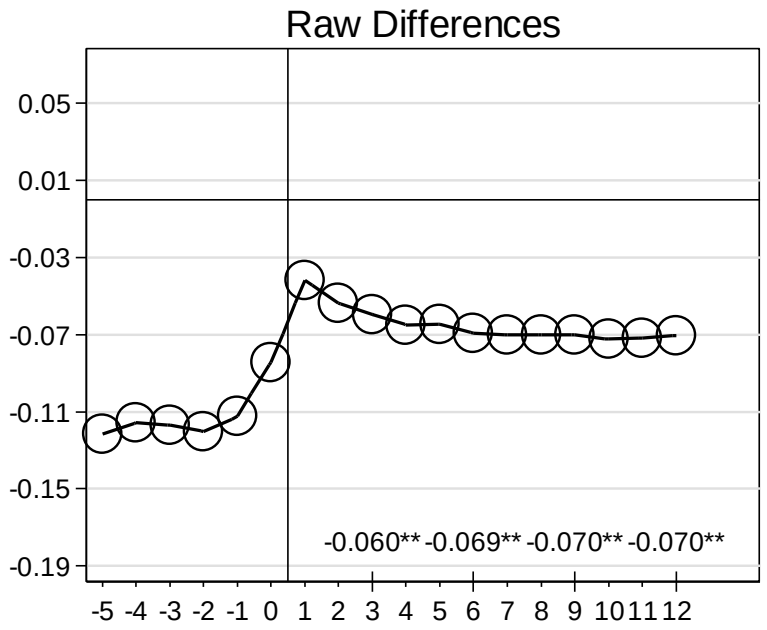
What did we find?

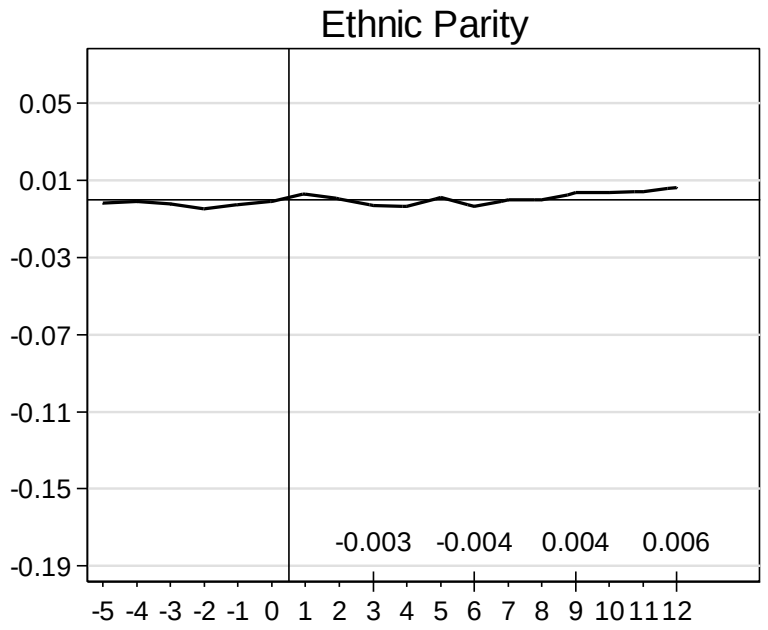
- For most programs and benefits (with exception of IS and IB), Minorities and Whites are simply too different for satisfactory estimates to be calculated and results are sensitive to the methodology used.
  - MASSIVE COMMON SUPPORT PROBLEMS
- This calls into question previous results based on simple regression techniques, which may hide the fact that observationally different ethnic groups are being compared by parametric extrapolation.
  - In some cases, depending on method used, eg NDLP we could find significant ethnic penalites in employment (raw and DID), no ethnic penalty (regression methods) and significant ethnic premium (PSM)

IB: raw labour market status



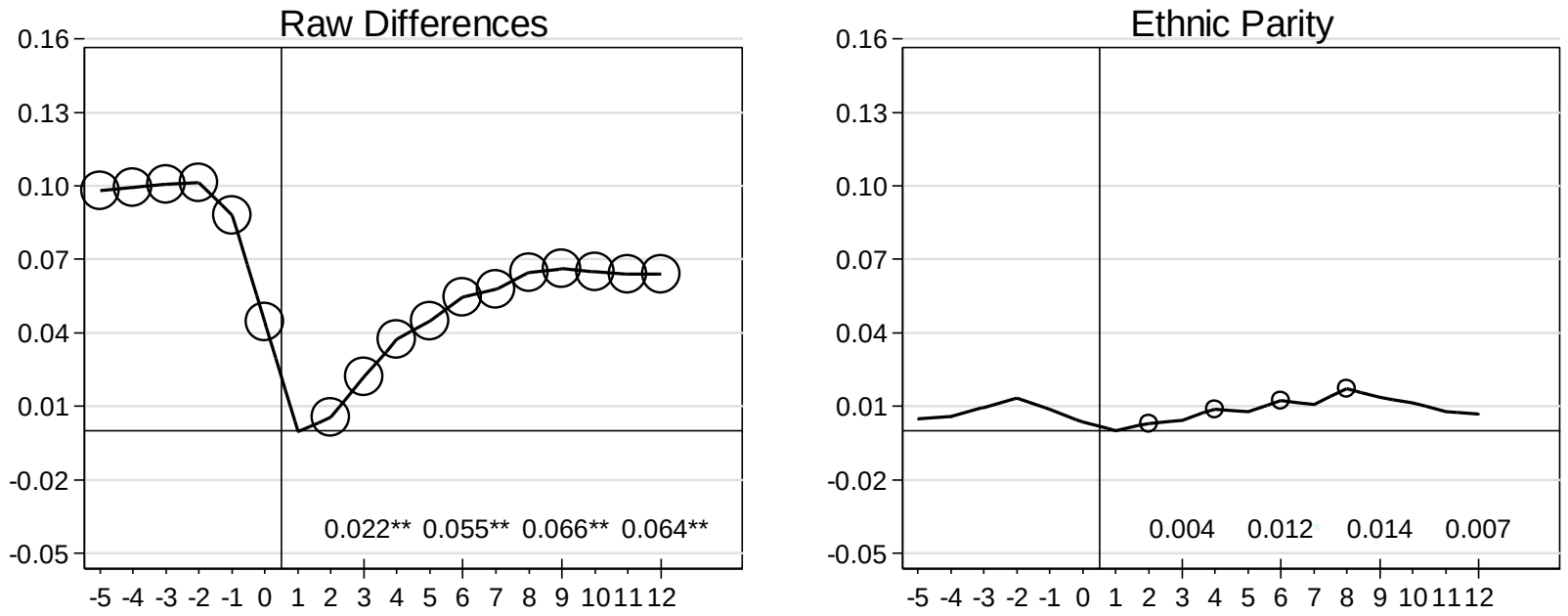
IB: overall employment result





Reliability of matching: CS(0), UC(28) (i.e. reliable according to our criteria)

IB: overall benefit result



Reliability of matching: CS(0), UC(28) (i.e. reliable according to our criteria)

Need other methods to do this properly

- Using administrative data to analyse this question very problematic
- Problem due to the fact that the Ethnic Minority and White clients accessing the same JCP office are very different in the UK with exception of IS and IB recipients
- Might not be problem in other countries but could be…….

Conclusions

- Administrative data is going to be used more and more by economic researchers
- But not answer to every question and depends crucially on question being asked, whether there is some nice natural experiment (e.g. Pilot program, historical accident, regression discontinuity, variation in policy by area) and the richness and reliability of the data
- The increasing linkage of administrative data to survey data will also allow us to test more fully the best ways of using administrative data but that is another talk…

Selected References

- Crawford, C., Dearden, L. and Meghir, C (2009), "When you are born matters: The Impact of Date of Birth on Educational Outcomes in England", mimeo, IFS.
- Crawford, C, Dearden, L. Mesnard, A., Shaw, J and Sianesi, B. (2009), "Ethnic parity in labour market outcomes for benefit claimants", IFS working paper no.
- Dearden, L., Sartarelli, M. and Vigoles, A. (2009), "Value added for students choosing a secondary school in England", mimeo, ADMIN, Institute of Education, University of London.
- Dehejia, R. (2005) "Program evaluation as a decision problem", Journal of Econometrics 125, 141-173
- Hsu, J. (1996) *Multiple Comparisons Theory and Methods*, Chapman and Hall
- Frolich, M. (2008) "Statistical treatment choice: an application to active labor market programs", Journal of the American Statistical Association 103 (482), 547-558.
- Lechner, M. and J. Smith (2006) "What is the value added by caseworkers?", Labour Economics 14, 135-151.

End