The Gender Wage Gap Among Those Born in 1958: A Matching Estimator Approach

(ESRC Grant No. ES/S012583/1)

Alex Bryson
UCL

UCL SRI Gender Equality Workshop

21st April 2022

Project Overview

- Part of an ESRC funded project examining the GWG over the life course using birth cohort data
- The UCL team:
  - Alex Bryson (PI)
  - Heather Joshi (co-investigator)
  - David Wilkinson (co-investigator)
  - Francesca Foliano (Research Fellow)
  - Bozena Wielgoszewska (Research Fellow)
- All information on the project can be found here: https://www.ucl.ac.uk/ioe/departments-and-centres/departments/social-science/gender-wage-gap-evidence-cohort-studies

Motivation

1. Drawbacks in parametric estimation of the gender wage gap (GWG)
    - Failure to compare ‘like’ men and women
2. Common to condition on potentially endogenous variables
    - Biases ‘true’ estimates of the GWG
3. Data from the National Child Development Survey (NCDS) provide good basis for tackling these issues:
    - Match men and women on a rich set of variables liable to impact wage formation over the life cycle which might conceivably differ by gender
    - Measured pre-labour market entry and thus less liable to be endogenous with respect to wage formation
        - Birth, 7, 11 and 16 years collected prospectively

Preview of Results

1. Large raw GWG rising until 40s then falls but remains sizeable to age 55
2. Raw GWG is larger when use propensity score matching (PSM)
   - Very different to usual regression-adjusted estimates which are often half the size of the raw GWG
3. Implication: women have pre-labour market traits which reduce their earnings later in life relative to men
   - Chief among these is occupational expectations
   - Not true for all traits

Previous Literature

1. Studies indicate inverted u-shape in the GWG over the life course
    - Small in early years, widening in 30s/40s, narrowing thereafter
2. Falls across cohorts
3. Raw gap tends to close by (roughly) one half when condition on other variables
    - Depends somewhat on data set and conditioning variables
4. Frequently treats education and fertility decisions as exogenous when, in fact, might be endogenous and partials out some of GWG
    - Same could be said of job traits
5. Some exceptions using structural estimation in an effort to tackle endogenous decision-making
    - Adda, Dustmann and Stevens 2017 “The Career Costs of Children”, Journal of Political Economy

Value of Matching Estimators

1. Linear estimation (and decompositions on which most are based) based on unnecessarily restrictive assumptions regarding functional form
2. By ignoring common support, compare wages of women to men who may not be reasonable comparators
3. Matching may make a substantive difference to the estimation of the GWG
    - Strittmatter and Wunsch (2021) explain more of GWG when estimated with PSM
        - Substantial common support issue in their data
        - Combine exact matching on key wage determinants with PSM (radius) matching

PSM v OLS

1. Both assume relevant differences between treated and non-treated are captured by their observables (conditional independence assumption)
   - violated if analysis does not incorporate all factors affecting participation and outcome of interest
   - the assumption is not testable
2. Advantages of PSM relative to OLS
   - semi-parametric so does not require assumption of linearity in outcome equation
   - individual causal effect is completely unrestricted so heterogeneous treatment effects can be captured (no assumption of constant additive effects)
   - highlights problem of common support since women must have ‘like’ counterparts in male population. Thus, avoids extrapolating beyond CS but implications if many treated individuals remain ‘unmatched’

Data and Methods

1. National Child Development Study (NCDS)
2. Log hourly wages at ages 23, 33, 42, 50, 55
    - Deflated to January 2000 prices
    - Rerun matching for each wage outcome
3. Propensity score matching (PSM) matching women to men on single index (the propensity score) derived from probit (0,1) if woman
4. Using pre- labour market covariates from mother, cohort member, teacher
    - Parental background; pregnancy/birth; ages 7, 11, 16
5. Theory driven as opposed to data driven (Machine Learning)
6. Plausibility of conditional independence assumption in this case
7. 5 nearest neighbours (Froelich) to recover ATT
    - enforces common support with 0.005 caliper
    - Bootstrapping
8. Accommodate selection into employment using zero wages

Covariates used in matching

Wave	Variables
Pre-birth/birth	Gestation (days); mother smoked during pregnancy; white; birthweight (ounces); mother smoking 4 months after birth
Age 7	Southgate reading test score; arithmetic problems; N Rutter symptoms; Score on Bristol Social Adjustment Guide; N child illnesses; N hospital admissions; laterality in hands
Age 11	Occupational expectations when aged 25; standardized reading score; standardized maths score; type of school attended;
Age 16	In trouble with police; teacher rating on capability relating to maths, English, Languages, practical issues; mother's assessment of over/underweight; disability; alcohol consumption; smoking behaviour

Others tried: mum's and dad's social class; breast fed; region; housing tenure; siblings and household size; verbale and non-verbale test scores aged 11; female teacher; teacher rating of child aged 11; child's expectations on schooling; mum's and dad's interest in education of child; financial hardship and FSMs aged 7 and 11

Occupational Expectations

	Male	Female
Professional	9	4
Other non-manual, scientific	6	4
Typist, clerical	2	11
Shop assistant	1	7
Junior non-managerial	3	1
Personal services	1	9
Foreman, manual	<1	<1
Skilled manual	18	1
Semi-skilled manual	3	<2
Unskilled manual	<1	<1
Self-employed	1	1
Farm worker	2	2
HM Forces	7	<1
Sports man/woman	9	<1
Student	<1	<1
Teacher/nurse	2	20
Unclassifiable	34	38

Match Bias

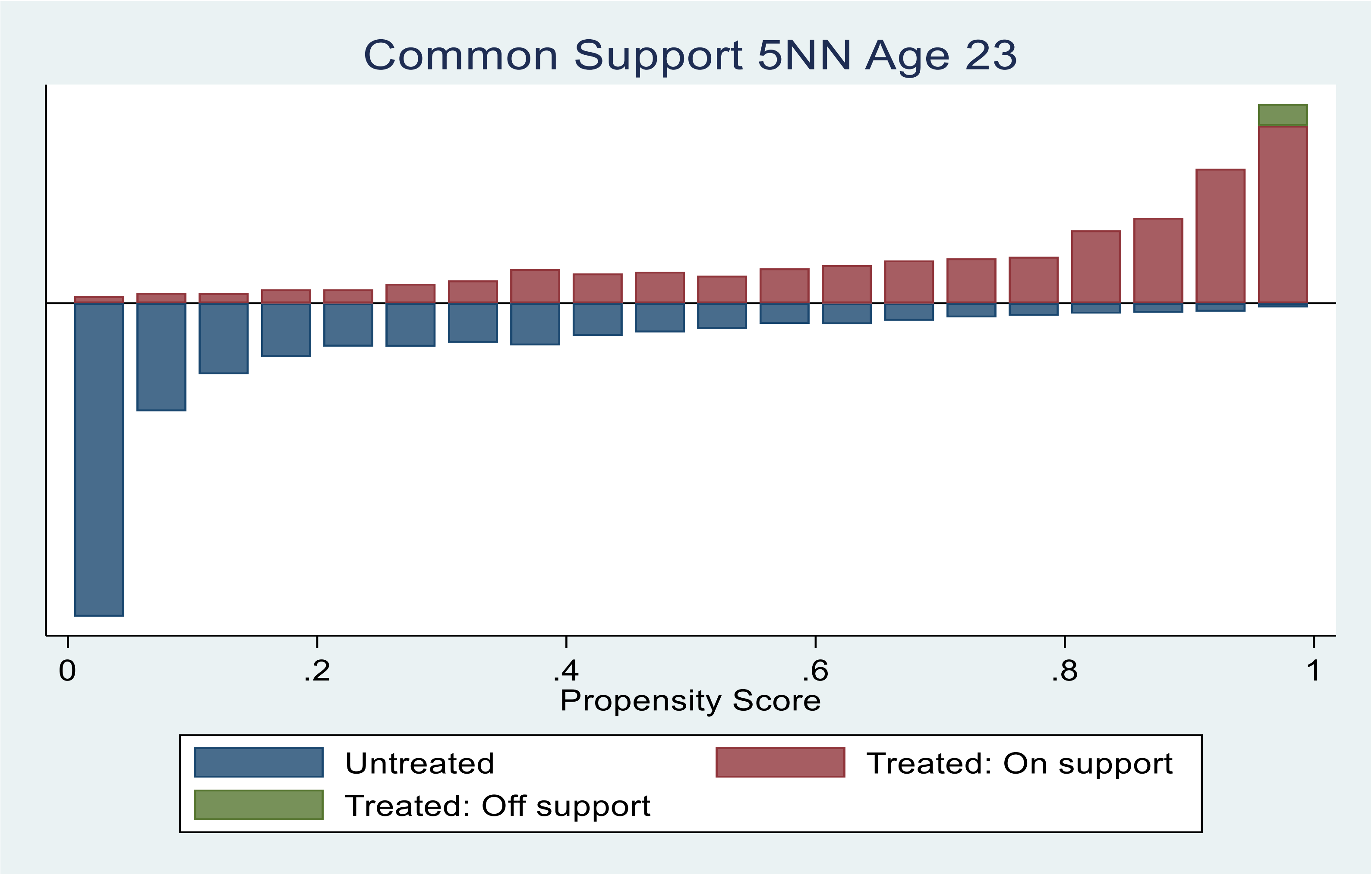
	23	33	42	50	55
Pseudo r-sq: **U** **M**	0.446 0.032	0.440 0.046	0.429 0.041	0.437 0.042	0.450 0.060
Rubin's B	41.4*	51.4*	48.1*	48.8*	58.6*
Rubin's R	0.83	1.13	0.95	1.05	1.08

Rubin's B: absolute standardised differences of means of linear index of propensity score in treated and match non-treated groups (B<25 is ok)

Rubin's R: ratio of treated to matched non-treated variances of propensity score index (R between 0.5 and 2 is deemed balanced)

- means falls outside tolerable balance limits
- Problems in relation to reading and maths scores at age 11, Rutter, BSAG, illness

Common Support



87 cases off common support. Zero at other ages

Results:
Log Hourly Wages, Raw Diffs and ATT from Matching

	23	33	42	50	55
Fem	1.536	1.843	1.908	2.080	2.022
Male U	1.704	2.209	2.354	2.455	2.359
Male M	1.711	2.211	2.402	2.504	2.437
Dif Raw	*-.168*	*-.367*	*-.446*	*-.355*	*-.337*
Dif Matched	*-.176*	*-.368*	*-.495*	*-.424*	*-.415*
N	8011	6881	7175	6031	4992

U = unmatched
M = matched

At age 23: raw = -.168 OLS = -.184 PSM = -.176 OLS with CS and match weights = -.173 ebalance = -.181

What Next?

1. Regression using match weights having enforced CS
    - Compare regression-adjusted GWG with standard regression
    - Compare decomposition with decomp from standard reg-adj GWG
2. Specification for probit
    - Have we got the right covariates?
    - More flexible specification
3. Alternative matching estimators
    - NN, kernel; combine exact matching with PSM; entropy balancing
4. Tackline participation decision
    - Bringing in the zeros results in a much larger GWG
    - Is this the right thing to do?
5. Attrition
6. Check out wages at ages 61 and 63 (smaller Ns)
7. Run on BCS 1970