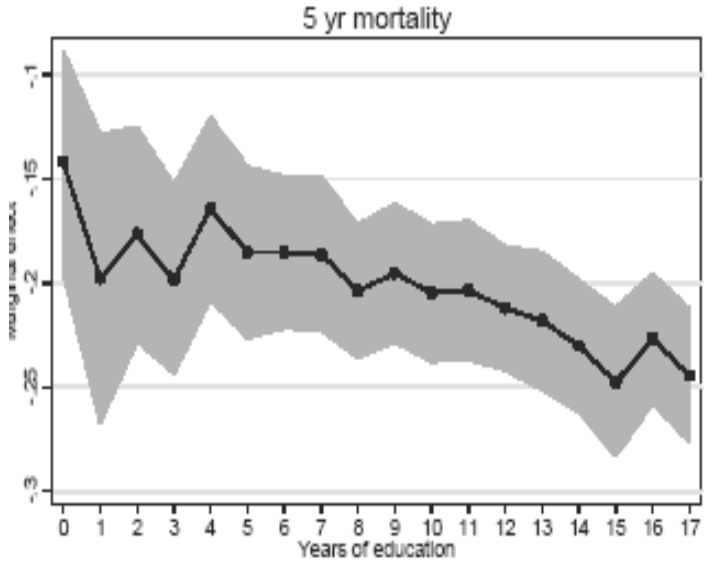
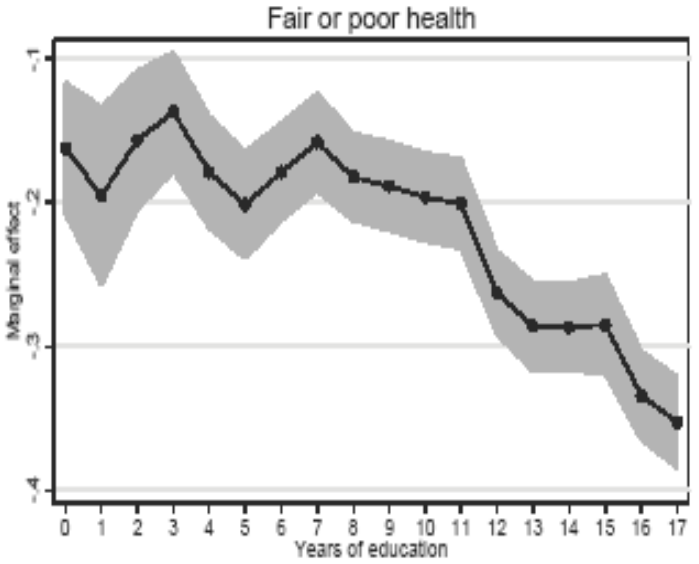
Education, Social Capital, and Health: An Empirical Framework

Don Kenkel
Cornell University & NBER
Prepared for the Workshop on Social Capital and Health, IRDES & OECD, Paris, 10 -11 October 2008

Goal

- Assess the feasibility of estimating the marginal effect of increases in the level of Education on Health and Social Capital
- Show how this can be done based on available data (esp. outside US, in OECD)
- Suggest a way to estimate a function or schedule showing the causal relationship between Health/ Social Capital and years/ level of education
  - Analogous to Figure 2, p.36 of the Cutler & Lleras-Muney, NBER working paper #12352 (on next slide)


5 yr mortality
-.1
-.15
-.2
-.25
-.3
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Years of education


Fair or poor health
Marginal effect
-.1
-.2
-.3
-.4
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Years of education

Challenge #1: Identifying Causal Effects

- Research documents strong correlations:
  - Between years of education and Health outcomes and behaviors.
  - Between years of education and behaviors and outcomes related to Social Capital/ Civic & Social Engagement
- Correlation is NOT causation

Challenge #2: Non-linear Relationships

- Marginal effect of an additional year varies across level of education
- Standard *assumption*: diminishing marginal effect
  - With 3 years of formal schooling, marginal effect of 1 more year is to add a lot of Health, Social Capital
  - With 16 years of formal schooling, marginal effect smaller

Which challenge is more important?

- Econometric challenge of identifying causality attracts academic interest
- How the marginal effect varies might be more relevant for policy making
- Policy economics is harder than academic economics

Aside: terminology

- "Education" or "Schooling"
  - Usually we measure schooling, not education
  - Policies can↑schooling
- "Social Capital" or "Civic and Social Engagement"
  - I want a broad, umbrella term to conveniently summarize a lot of individual-level outcomes (CSE)

Canonical equation

- **$\text{Outcome}_i = \alpha + \beta\ \text{Schooling}_i + \gamma\ X_i + \varepsilon_i$**
  - Outcome_i = earnings, health, CSE outcome for individual *i*
  - β is the marginal effect of an additional year of schooling on the outcome (linear)
  - X are control variables
  - Unobservable influences captured by ε_i

Interpretation of β

- Shows causal effect of schooling on outcomes
  - In an earnings function, β is an estimate of the private rate of financial returns from investing in more schooling
  - In other functions, β estimates health or CSE returns from investing in more schooling

“Structural” interpretation of β

- What does β mean in a structural economic model of individual decision-making?
- What are the channels through which more schooling leads to higher earnings, better health, and more CSE?
  - “Channels” ↔ “Structural relationships”

"Structural" interpretation of β, cont.

- "In many economic models of health, education is seen as enhancing a person's efficiency as a producer of health—a suggestive phrase, but not one that is very explicit about the mechanisms involved." (Deaton 2002)
- Allocative efficiency: schooling leads to different set of health inputs (e.g. less smoking, more exercise)
  - Schooling → information → health behaviors
- Parallel ideas for CSE?

Interpretation of β: Causal

- Identify causal link between schooling & outcome
- Do not necessarily identify channels (structural parameters)

Will β be a good estimate of the causal effect of schooling?

- Reverse causality: poor Health/ low CSE reduces educational attainment.
- Hard-to-observe “hidden third variable” or variables that are the true causes of both educational attainment and Health/ CSE (unobservable heterogeneity)
  - individual rate of time preference
  - attitudes related to self-efficacy
  - ability

Solutions

- Good data

- Fancy econometrics
  - This solution really relies on good data, too

The “best” data

- Randomized controlled trial
  - Assign some people to the control group that receives standard schooling
  - Assign others to a treatment group (or groups) that receive more schooling
  - Compare outcomes of treatments vs. controls
- In observational data, instead of random assignment people choose schooling levels
  - Same type of people may also choose to invest in more health, CSE

Good data

- Include controls for past health, CSE
  - Reduce bias in β due to reverse causality from past health/ past CSE to schooling
- Possibilities
  - Longitudinal data from childhood on (rare)
  - Longitudinal data on adults (may not solve problem)
  - Retrospective data on health problems in childhood
  - Family background measures proxy for differences in past health, CSE

Good data, cont.

- Include controls for hidden third variables
  - Some surveys try to measure risk, time preference, self-efficacy
  - Some surveys include ability measures (cognitive & noncognitive skills)
- Include proxies for hard-to-observe characteristics
  - Savings & consumer debt
  - Smoking status proxies for risk preferences
  - Is the “cure” (including proxies that are themselves endogenous) worse than disease?

Fancy Econometrics: IVs Based on Educational Reforms

- Use econometric method of Instrumental Variables (IV) to identify causal effects of education on Health/ CSE outcomes and behaviors
- IVs based on educational reforms: These provide a "natural" or "quasi-experiment" where people "treated" with the reform receive more education than untreated "control" group (so technique really relies on good data again)
- Method widely used in labor economics to identify earnings returns to education (Card, Econometrica 2001)

Key Ingredients for Empirical Framework

- Surveys that Measure Health/CSE Outcomes and Behaviors
  - Country-specific surveys
    - Examples: Danish panel survey, British Election Surveys
  - European Community Household Panel measures:
    - Physical and mental health outcomes
    - Social relations
  - WHO Multi-Country Survey Study measures:
    - Health Outcomes
    - Alcohol consumption
    - Depression
- Suitable IVs based on educational reforms available in a number of countries

Country	Educational Policy Used as IV for Education	Reference IV Study
Austria	school disruptions due to Word War II	Ichino and Winter-Ebmer (2004)
Canada	variation in school-leaving ages	Oreopoulos (2006)
	child labour laws	
Denmark	1958 reform: lowered educational barriers	Arendt (2005)
	1975 reform: raised school-leaving age from 7 to 9 years, and removed distinction between two tracks during 8^{th} to 10^{th} forms	
France	1968: educational reforms after student riots	Maurin and McNally (2008)
	Zay reform (increased school-leaving age to 14) and Bethoin reform (increased leaving age to 16)	Albouy and Lequien (2008)

Germany	school disruptions due to Word War II	Ichino and Winter-Ebmer (2004)
Ireland	mid 1960s: introduction of free secondary education 1972: school-leaving age increased from 14 to 15	Callan and Harmon (1999)
Italy	Law 910 of December 1969: possible for individuals who completed secondary education to enroll in college, regardless of curriculum chose in secondary school	Brunello and Miniaci (1999)
Korea	Expansion of high school in mid-1970s	Park and Kang (2008)
the Netherlands	1982: duration of university education decreased from five to four years	Webbink (2007)

Norway	1960s: compulsory education increased from seven to nine years	
Portugal	1956: compulsory education increased from three to four years 1964: compulsory education increased from four to six years	Vieira (1999)
Sweden	1960s compulsory education increased from seven or eight to nine years	Meghir and Palme (2005)
Taiwan	1968: compulsory education increased from six to nine years large expansion in junior high school construction (intensity varied across regions of Taiwan)	Chou et al. (2007)
United Kingdom	1947: minimum school leaving age increased from 14 to 15 1973: school reform	Harmon and Walker (1995) Oreopoulos (2006)

Example: "When Compulsory Schooling Laws Really Matter"

- Oreopoulos (2006) studies compulsory schooling reforms in Britain & Northern Ireland
- He estimates that the average increase in earnings in Northern Ireland from raising the school-leaving age from 14 to 15 is 13.5% - 20%

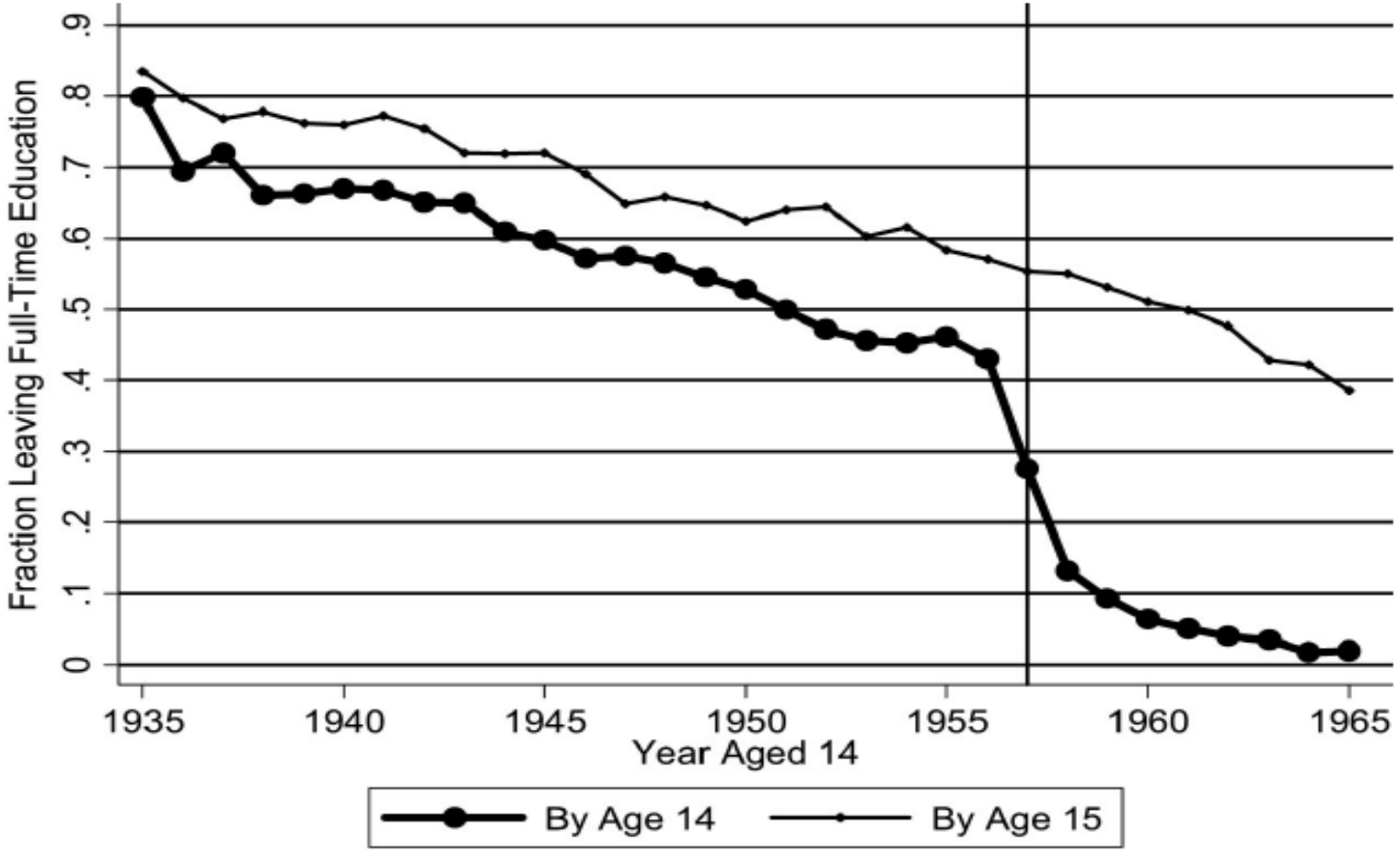


FIGURE 2. FRACTION LEFT FULL-TIME EDUCATION BY YEAR AGED 14 AND 15
(*Northern Ireland*)

Note: The lower line shows the proportion of Northern Irish adults aged 32 to 64 from the 1985 to 1998 General Household Surveys who report leaving full-time education at or before age 14 from 1935 to 1965. The upper line shows the same, but for age 15. The minimum school-leaving age in Northern Ireland changed in 1957 from 14 to 15.

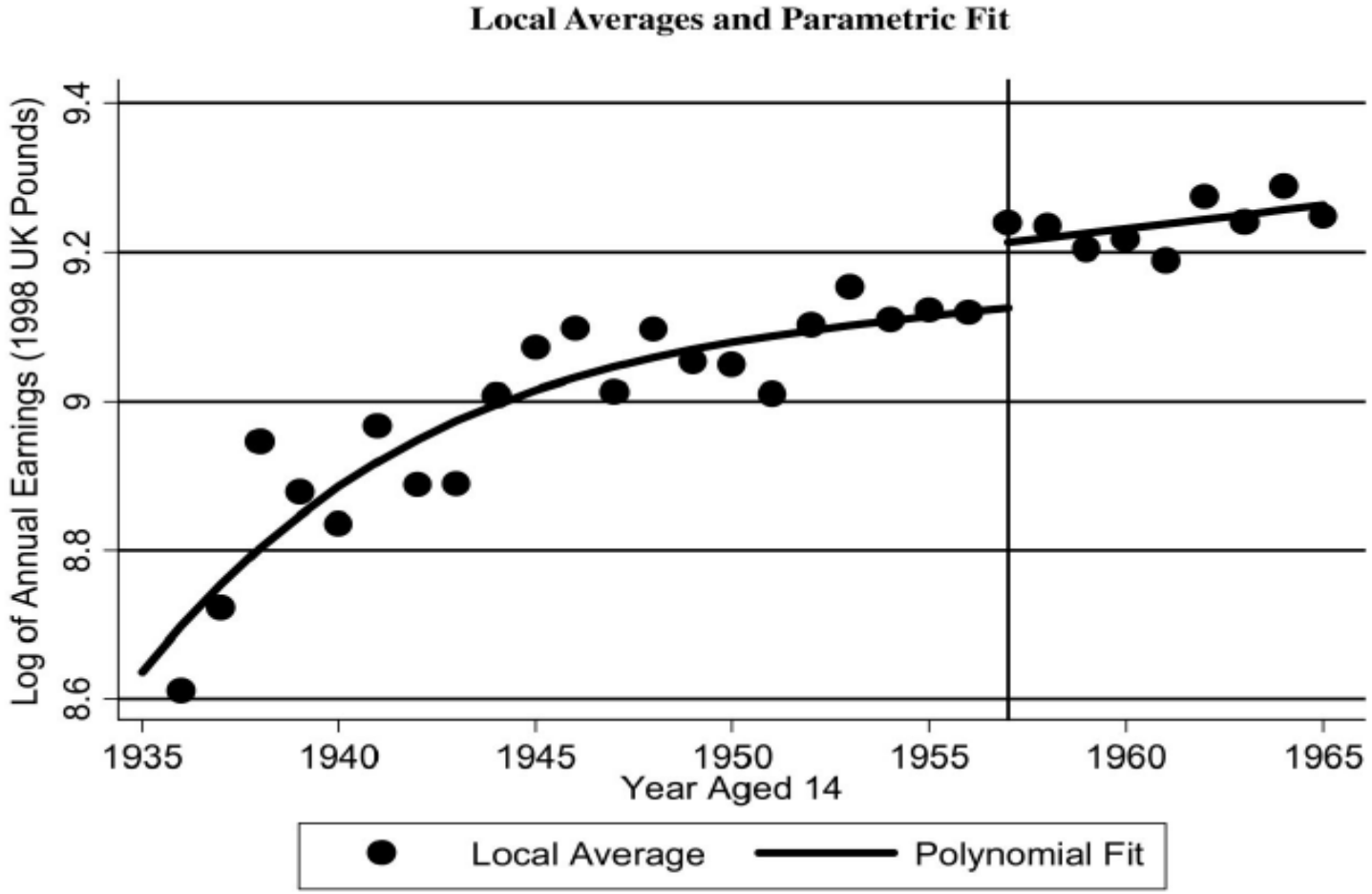


FIGURE 7. AVERAGE ANNUAL LOG EARNINGS BY YEAR AGED 14
(*Northern Ireland*)

Note: Local averages are plotted for Northern Irish adults aged 32 to 64 from the 1985 to 1998 General Household Surveys. The curved line shows the predicted fit from regressing average log annual earnings on a birth cohort quartic polynomial and an indicator for the school-leaving age faced at age 14. The school-leaving age increased from 14 to 15 in 1957, indicated by the vertical line. U.K. pounds using the U.K. retail price index.

Extensions

- Comparison of IV and OLS estimates
  - Bias $\rightarrow \beta_{OLS} > \beta_{IV}$
  - Often find $\rightarrow \beta_{OLS} < \beta_{IV}$
- Non-linear functional form
- Heterogeneous treatment effects (LATEs)
- Cross-country comparisons
- General equilibrium effects

Non-linearities

- Non-linear relationship:

$$\text{Outcome} = \alpha + \beta_1 Y1 + \beta_2 Y2 + \ldots \beta_{18} Y18 + \gamma X + \varepsilon$$

(Y1 indicates 1 year of schooling, etc.)

- More flexible functional forms demand more from the data
  - May lack sample size for precision

Non-linearities, cont.

- Using IV approach to estimate non-linear relationship is at cutting edge
  - Moffitt (2007) NBER working paper 13534
- Need IVs that identify different margins of education

Heterogeneous Treatment Effects

- **$\text{Outcome}_i = \alpha + \beta_i \text{ Education}_i + \gamma X_i + \varepsilon_i$**

- Each individual *i* faces a different marginal effect β_i
  - Focus on distribution of treatment effects β_i , for example the average treatment effect (ATE)

IV estimates a LATE

- IV estimate is a weighted average of the causal effect of a year of schooling within a subgroup
  - Weights depend on how much the subgroup is affected by the IV
- Equally valid IVs relying on different subgroups generate different results corresponding to different LATEs

Concluding Comments

- "In my view one of the most important empirical developments in the past two decades has been the application of instrumental variables techniques to the relationship between schooling and earnings. There are many fewer examples of the application of this technique to the relationship between schooling and nonmarket outcomes. Such research deserves high priority on an agenda for future research ....." (Grossman, Handbook of the Economics of Education)

Concluding Comments, cont.

- “The perils of invalid and weak instruments open all instrumental variable estimates to skepticism. Although instrumental variable estimation can be a powerful tool for avoiding the biases that ordinary least squares estimation suffers....applying instrumental variables persuasively requires imagination, diligence, and sophistication.” (Murray 2006, *J. Econ. Persp*.)
- “In many cases the IV estimates are relatively imprecise, and none of the empirical strategies is based on true randomization. Thus, no individual study is likely to be decisive....” (Card 2001)

The sequel: estimating causal effects of social capital

- Health (or other outcome) = $\alpha + \beta$ SC + …
- Community-level Social Capital
  - Exogenous shocks/ natural experiments
- Individual-level Social Capital
  - Suitable IVs less obvious