Nobel lecture
IP Paris 2021

Angrist's and Imbens' contributions to applied economics

Pauline Rossi (CREST - Ecole Polytechnique)

"The credibility revolution"

Angrist and Pischke (JEP 2010)

- Answering causal questions using observational data
- **Natural experiments**
  - Units of analysis are exposed to as good as random variation caused by nature, institutions, or policy changes
- **Design-based approach**
  - Merge standard econometrics framework (e.g. OLS) with the potential-outcomes framework for causal inference
- Transparent way of assessing the credibility of key assumptions

Illustration: the returns to education

- Correlation is not causation
- Introduction to instrumental variables (IV)
  - Spirit
  - Validity
- Important applications
  - Regression Discontinuity Designs
  - Encouragement Designs

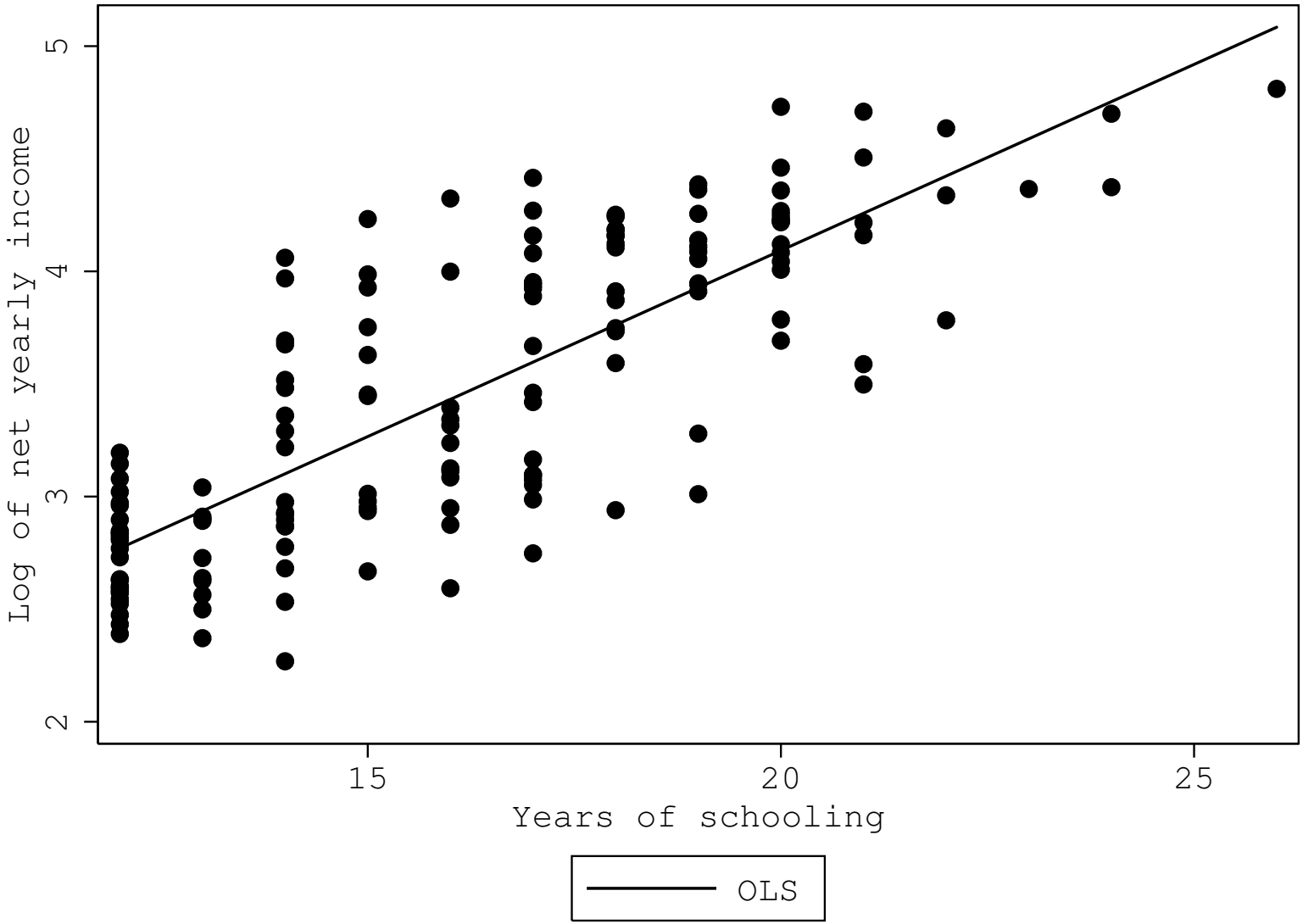
Part A

Correlation is not causation

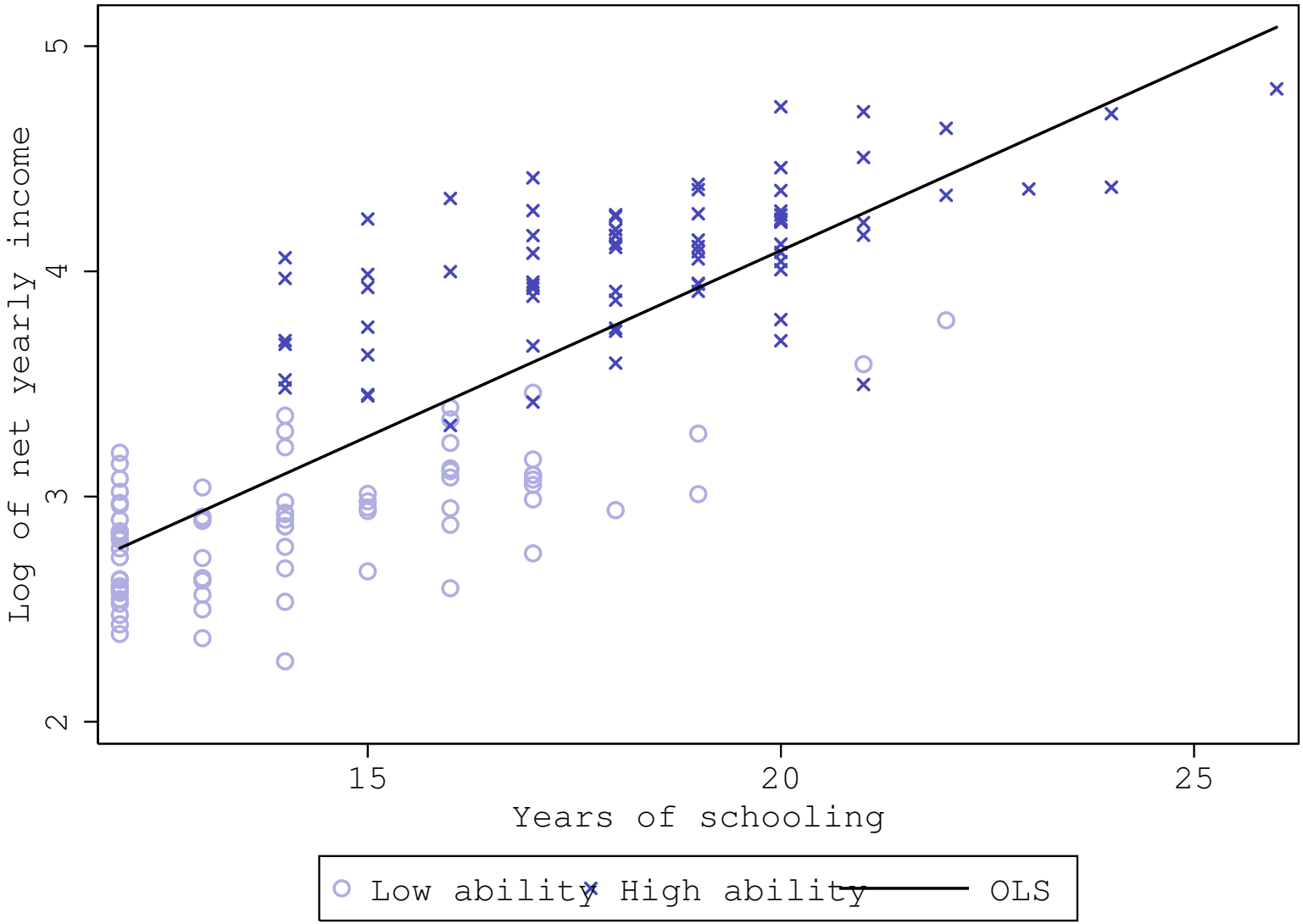
Estimating the returns to education

$$\text{Years of schooling} \overset{?}{\rightarrow} \text{Wage}$$

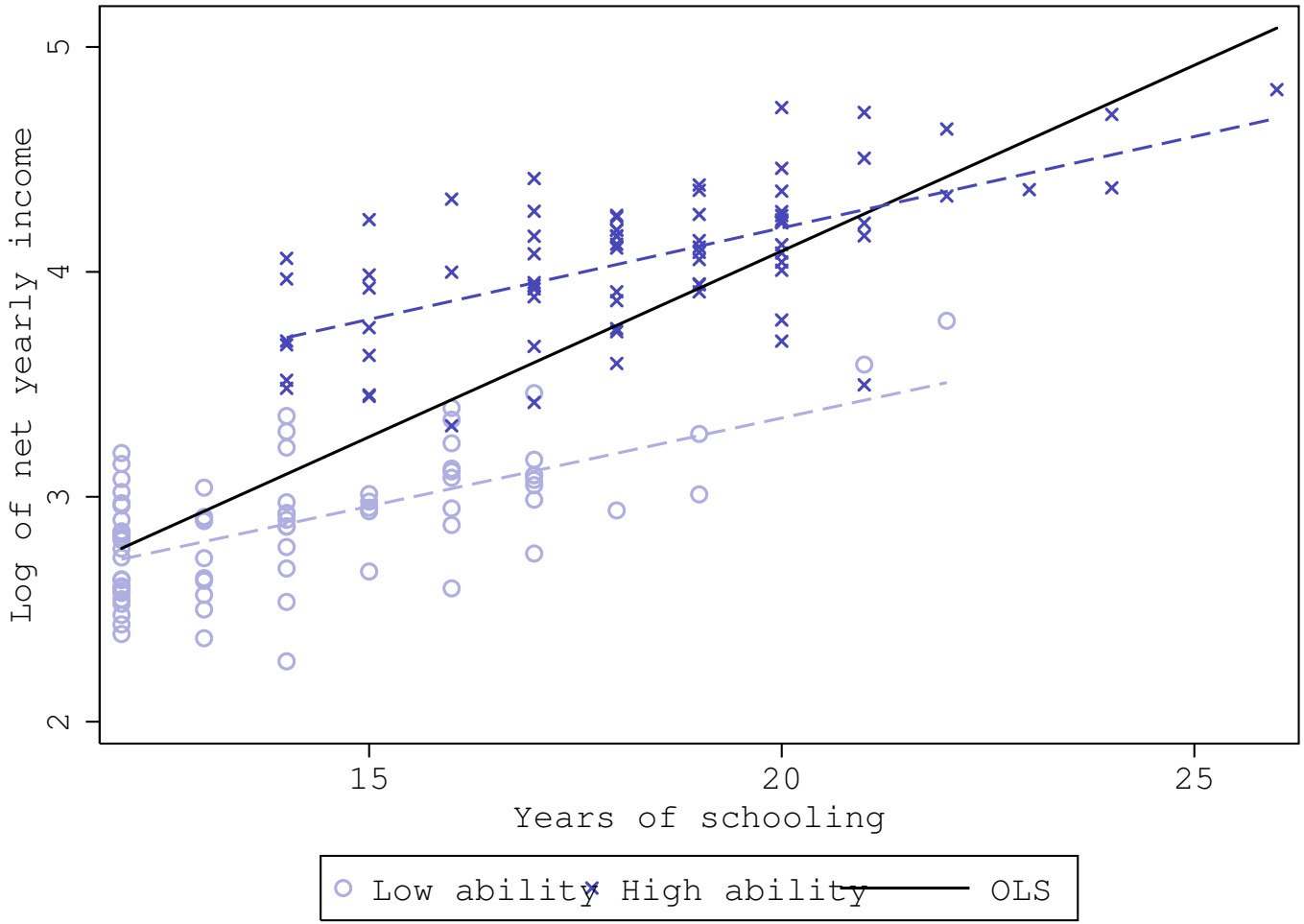
Association between schooling and wages



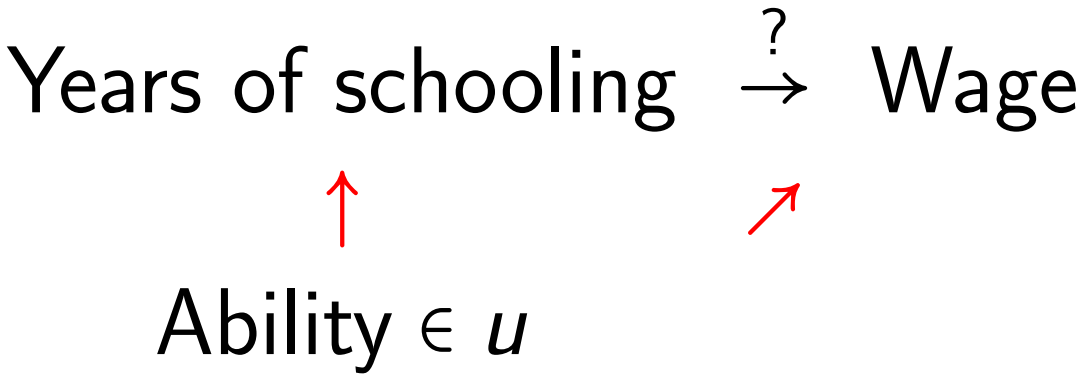
Suppose there are two types of individuals



OLS slope coefficient upward biased because of (unobserved) ability



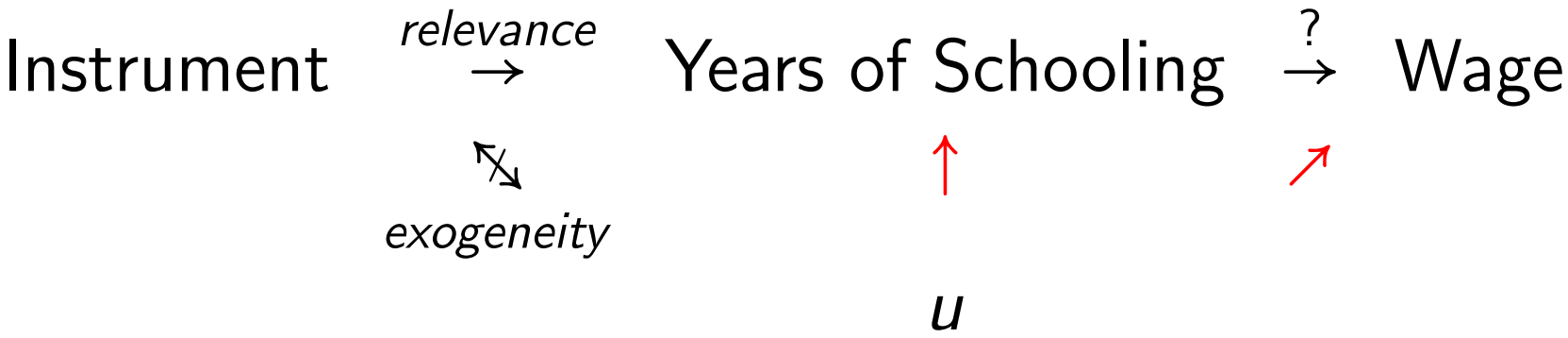
Omitted Variable Bias



- Ability is just one example: motivation, parental support, health issues etc.
- We do not observe these confounding variables: What to do?
- Potential solution: Instrumental Variables

Part B
Introduction to IV

IV spirit



Which variable can determine schooling and be unrelated to all other determinants of wages?

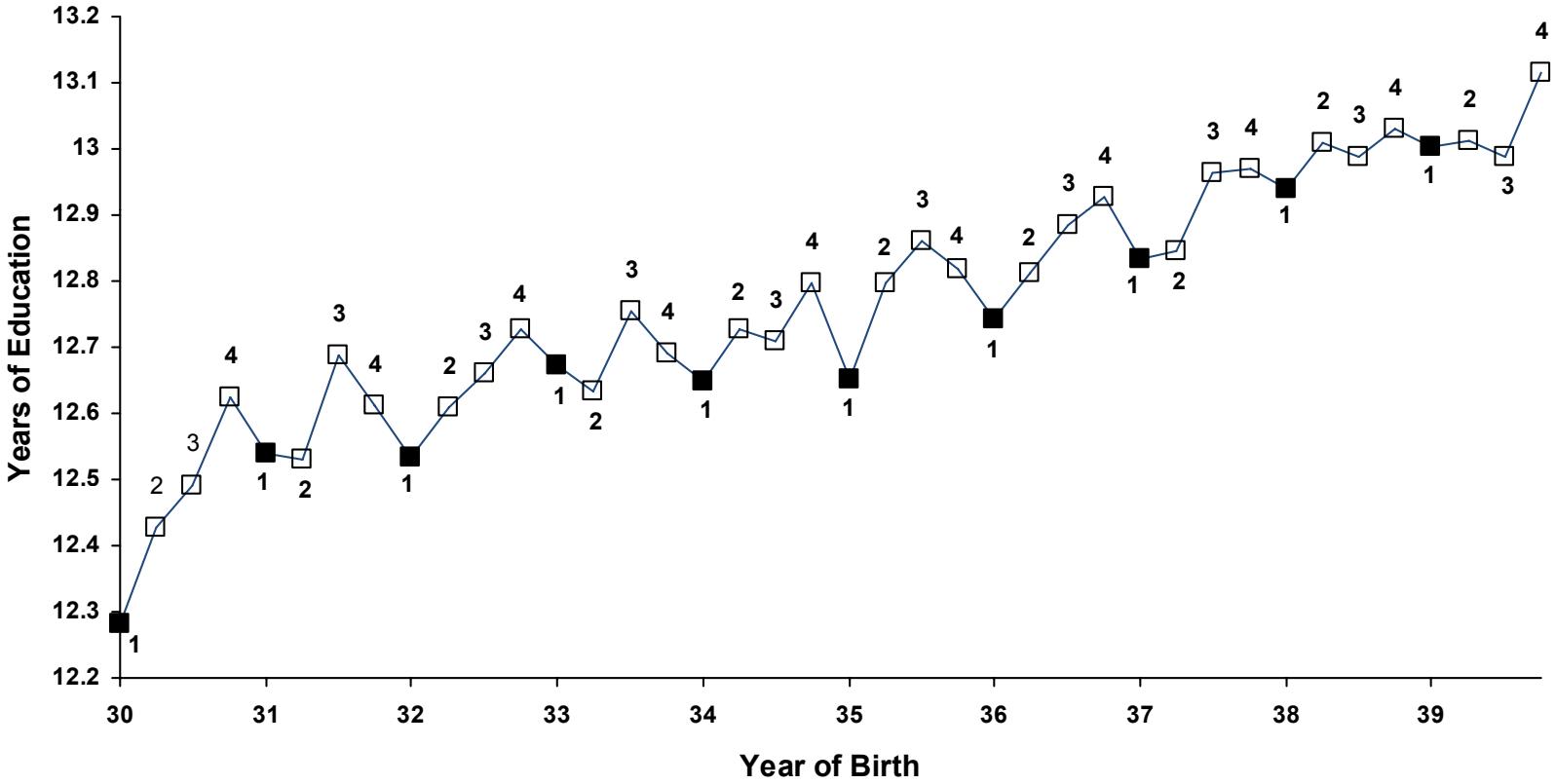
Quarter of birth (QoB) instrument

Angrist and Krueger (QJE 1991)

Compulsory schooling laws may provide a natural experiment:

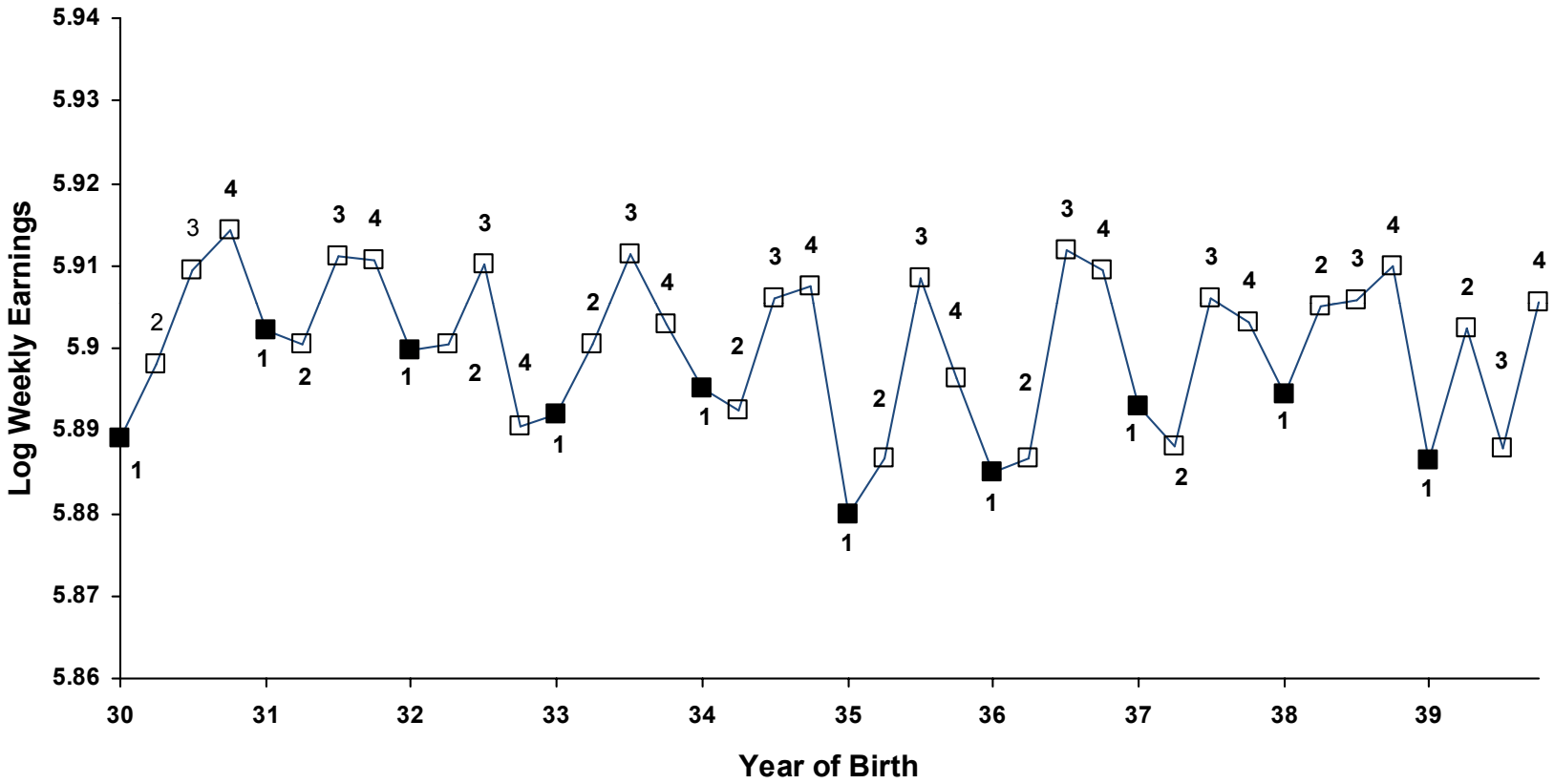
1. Children have to be in school until age 16. Children born later in the year have to attend school longer because they start at a younger age [*relevance*]
   - Born in Q1: enroll in grade 1 when they are 6.5 years old. Can drop out of school in the middle of grade 10.
   - Born in Q4: enroll in grade 1 when they are 5.5 years old. Can drop out of school in the middle of grade 11.
2. Children born later in the year should be similar to those born earlier [*exogeneity*]

First stage: Instrument relevance (coefficient= 0.1088)



Adapted from Angrist & Krueger (1991).

Reduced Form: Instrument effect (coefficient= 0.0111)



Adapted from Angrist & Krueger (1991)

Two-Stage Least Squares: intuition

Conceptually	Formally	Example Return to schooling
	First Stage	
	$\texttt{reg x z} \Rightarrow \hat{\beta}^{FS} = \frac{dx}{dz}$	$\hat{\beta}^{FS} = \frac{dS}{dQoB} = 0.1088$
$Z \rightarrow X \overset{?}{\rightarrow} Y$ $\uparrow \quad \nearrow$ u	**Reduced Form**	
	$\texttt{reg y z} \Rightarrow \hat{\beta}^{RF} = \frac{dy}{dz}$	$\hat{\beta}^{RF} = \frac{dln(W)}{dQoB} = 0.0111$
	Two-Stage Least Squares	
	$\hat{\beta}^{TSLS} = \frac{\hat{\beta}^{RF}}{\hat{\beta}^{FS}} = \frac{\frac{dy}{dz}}{\frac{dx}{dz}} = \frac{dy}{dx}$	$\hat{\beta}^{TSLS} = \frac{dln(W)}{dS} = \frac{0.0111}{0.1088} = 0.102$

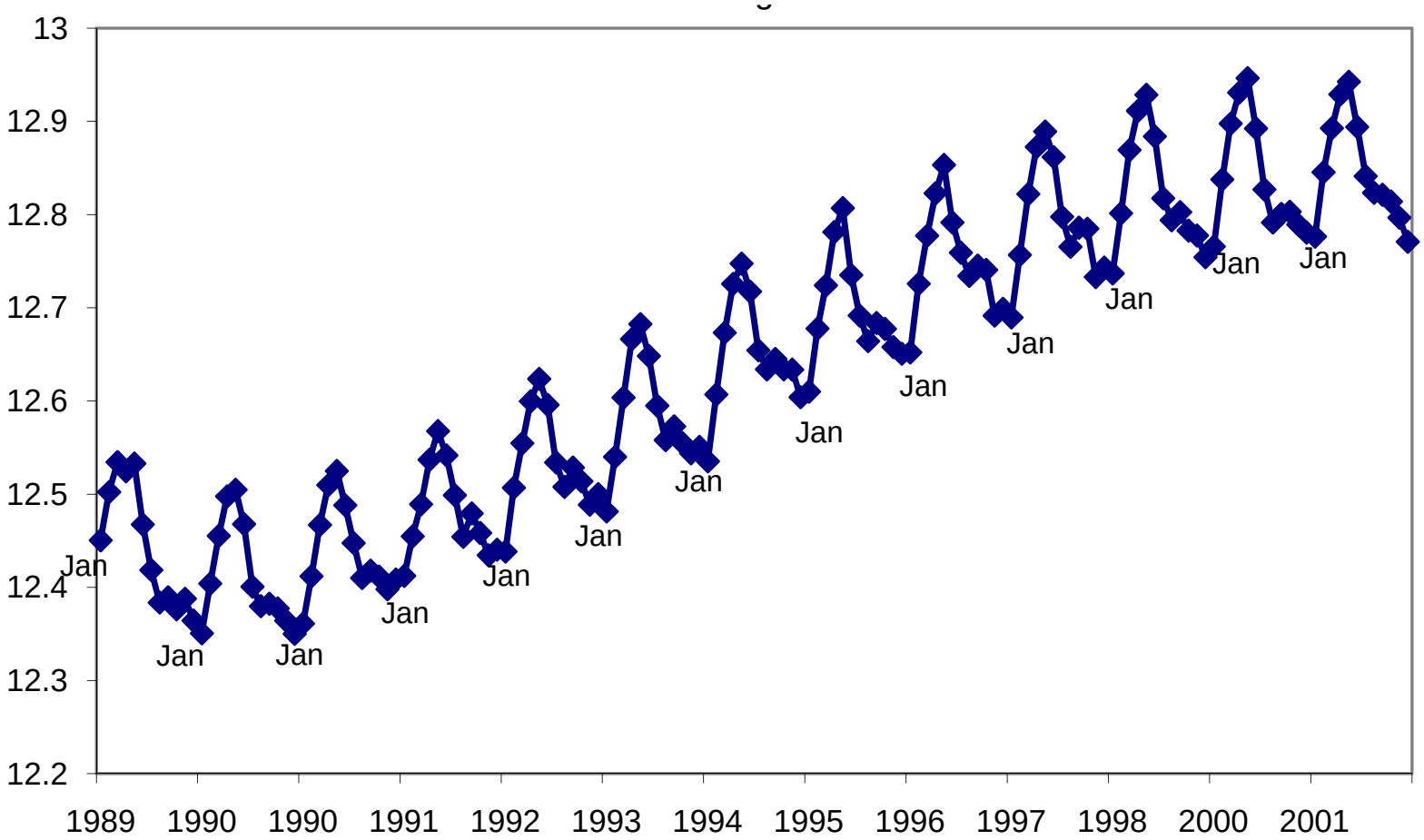
IV assumptions

The core IV assumptions are

1. **Instrument Relevance:** $\text{cov}(Z_i, X_i) \neq 0$
   - The instrument is correlated with the endogenous variable.
   - Testable: coefficient in the first stage regression should be significantly different from 0.
2. **Instrument Exogeneity:** $\text{cov}(Z_i, u_i) = 0$
   - **Independence:** Instrument not correlated to other factors
   - **Exclusion restriction**: Instrument has no direct effect on Y
   - Not directly testable but can be assessed by:
     - testing the relation with *other*, pre-determined, **W**
     - testing direct effect through *other,* post-determined, **W**

Exogeneity Assessment 1: Is Z correlated with u?

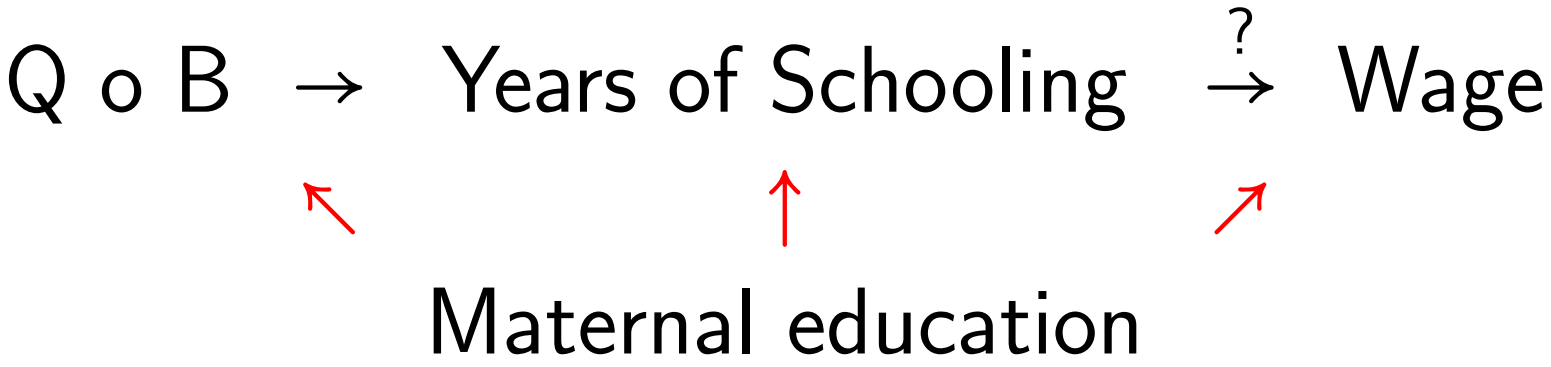
Example: QoB correlated with maternal education



Taken from Buckles & Hungerman (2013)

Exogeneity Assessment 1: Is Z correlated with u?

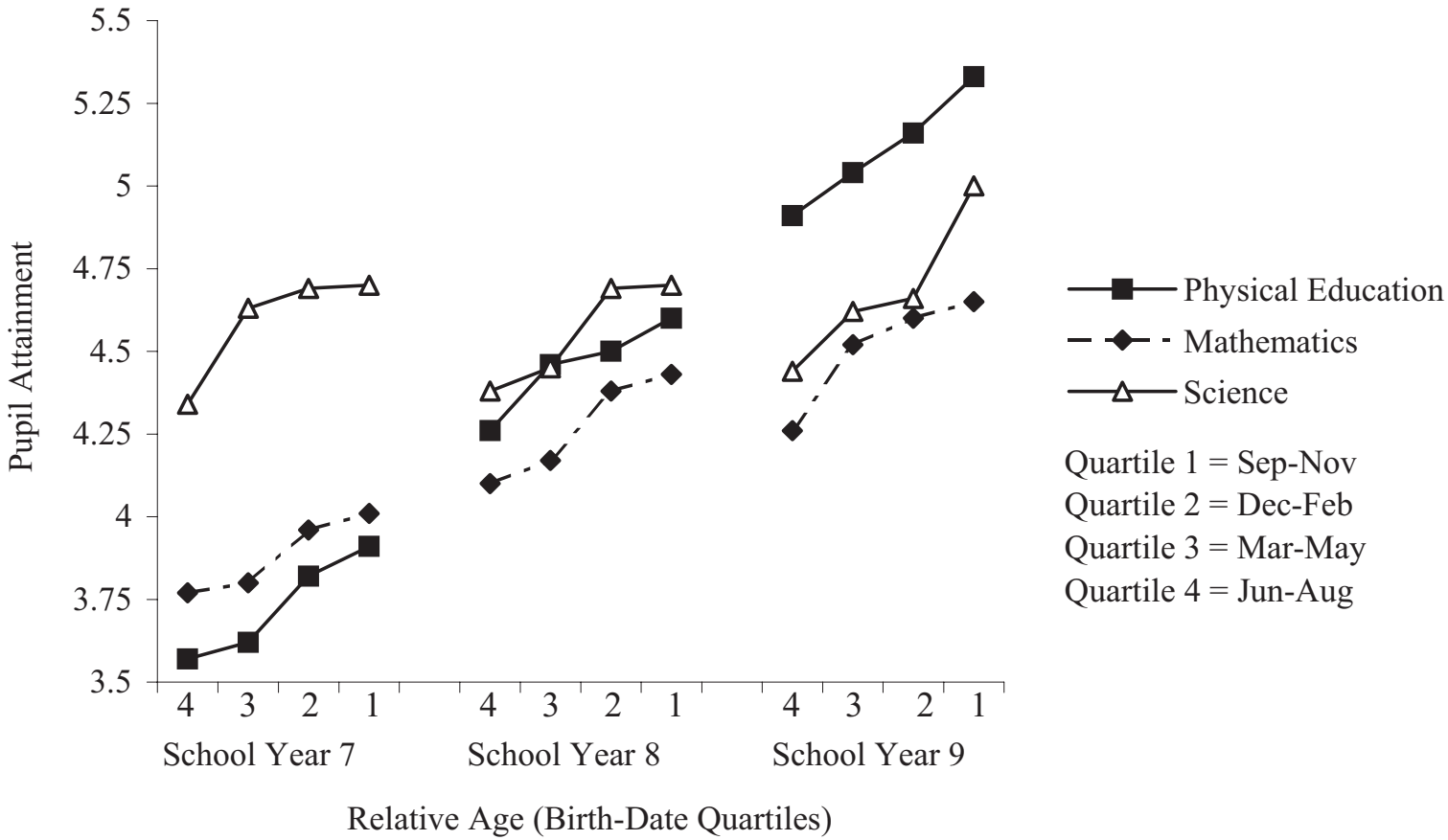
Example: QoB correlated with maternal education



Exogeneity fails through violation of independence

Exogeneity Assessment 2: Has Z direct effect through u?

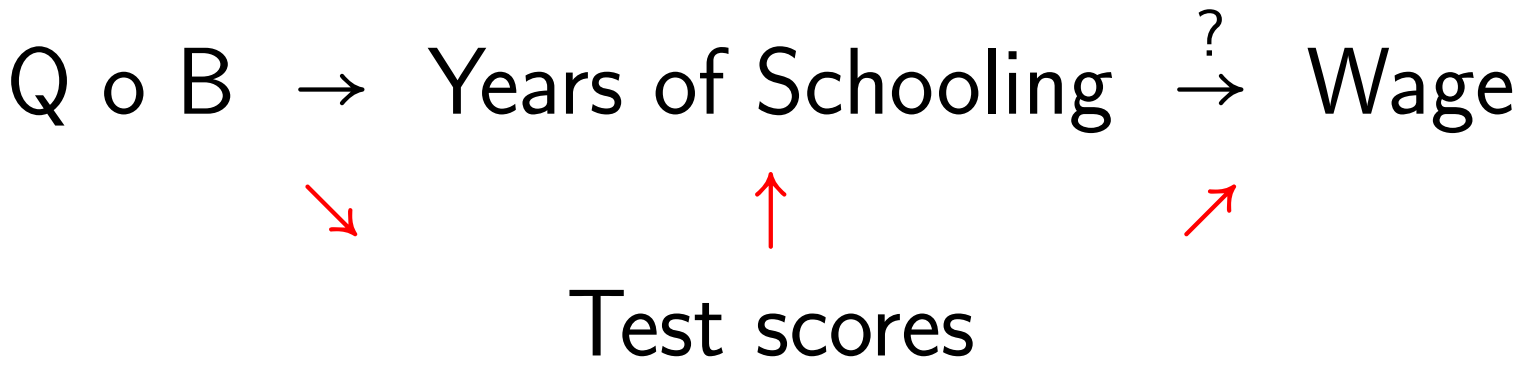
Example: QoB correlated with test scores



Taken from Cobley *et al*. (2009). Also see Bedard & Dhuey (2006)
Note: different context where the oldest in a class are born in Sep-Nov

Exogeneity Assessment 2: Has Z direct effect through u?

QoB correlated with test scores



Exogeneity fails through violation of the exclusion restriction

Part C
Important applications

Other famous instruments

- Angrist and Evans (AER 1998)
  - Number of children $\overset{?}{\rightarrow}$ female labor supply

Other famous instruments

- Angrist and Evans (AER 1998)
  - Number of children $\overset{?}{\rightarrow}$ female labor supply
  - IV: gender composition of first-born children

Other famous instruments

- Angrist and Evans (AER 1998)
  - Number of children $\overset{?}{\rightarrow}$ female labor supply
  - IV: gender composition of first-born children
- Angrist and Lavy (QJE 1999)
  - Class size $\overset{?}{\rightarrow}$ test scores

Other famous instruments

- Angrist and Evans (AER 1998)
  - Number of children $\overset{?}{\rightarrow}$ female labor supply
  - IV: gender composition of first-born children
- Angrist and Lavy (QJE 1999)
  - Class size $\overset{?}{\rightarrow}$ test scores
  - IV: Maimonides' rule (maximum number of children per class)

Regression Discontinuity Designs

- Take advantage of an institutional feature that generates a discontinuity in the exposure to a treatment.
  - E.g. conditions in terms of income to benefit from a program.
- Units just above and just below the cut-off are similar.
- Locally, RDD is a randomization: lucky vs. unlucky draws.
- Use the cut-off as an instrument for participation in treatment
- Returns to education example:
  - Admission cut-offs at elite schools
  - Abdulkadiroglu, Angrist and Pathak (ECTA 2014) "The elite illusion"

Encouragement Designs

IV widely used in experiments with imperfect compliance.

- X_i: actual treatment status (binary).
- Z_i: initial random assignment to treatment group (binary).
- In case of imperfect compliance, $X_i \neq Z_i$ and X_i is endogenous.
- Solution: use the initial assignment as an instrument for the actual status.
  - Z_i is correlated with X_i if some units comply with assignment.
  - Z_i is exogenous (random assignment).
- Returns to education example:
  - Admission lotteries at charter schools
  - Angrist, Pathak and Walters (AEJ: Applied 2013) "Explaining charter school effectiveness"

Conclusion

Key contributions

- Strengthen the credibility of applied work
  - Transparent "thought experiments"
  - Plausibility of the assumptions easy to assess
- Raise new questions for econometricians
  - Why do different instruments lead to different coefficients?
  - What if only some units respond to the instrument?
  - What if the treatment effect is heterogenous?
- Establish a general econometric framework to interpret coefficients in presence of heterogeneity and imperfect compliance.
  - Imbens and Angrist (ECTA 1994) "Identification and estimation of local average treatment effects"
  - Angrist and Imbens (JASA 1995) "TSLS estimation of average causal effect in models with variable treatment intensity"