

# Genetic Endowments, Educational Attainment, and Social Mobility

---

Jason M. Fletcher

Professor of Public Affairs, Sociology,  
Agricultural and Applied Economics, and  
Population Health Sciences



# Intergenerational Transmission Models

---

$$y_i^c = \beta_0 + \beta_1 y_i^p + \varepsilon_i$$

- The estimated intergenerational coefficient measures the strength of the intergenerational association
- It represents the combined effect of many different mechanisms,
  - Genetic inheritance
  - Environment
    - Prenatal
    - Postnatal: the environment in which the child grew up / investments



# Interpretation of coefficient

---

- “There may be **genetic differences** in ability that are transmitted from parent to child and that lead to intergenerational persistence in income or education.”
  - Potential for lower bound of persistence
- “To the extent that this is the underlying cause of the intergenerational correlation in income or education, it may suggest a more limited role for policy.”
- Not true with GxE

# Health and Retirement Study (HRS) Data

---

- >20,000 observations for individuals over 50 years old (and a spouse who might be younger)
- 1992-present, biennial survey
- 2006-2010 genetic data collection
  - ~12,000 observations
- Use Education Polygenic Score
  - Only European ancestry(~9400)
- Sample with non-missing parental education and state of birth (~7000), birth cohorts 1925-1960



# Issues and Solutions

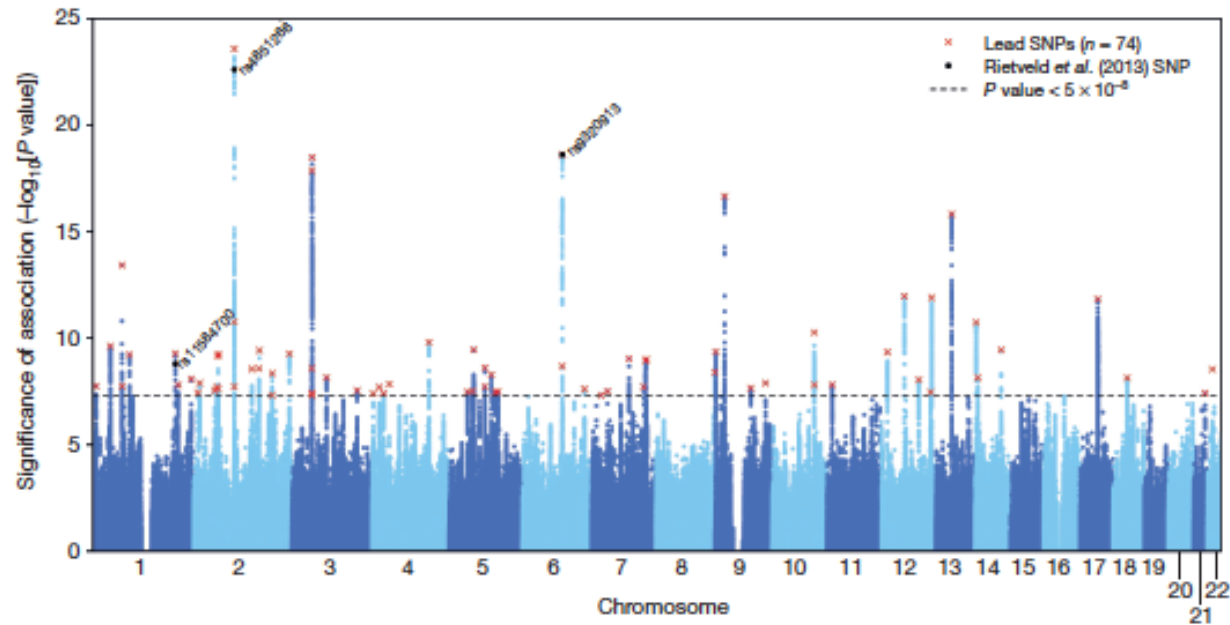
---

- Measuring genetic inheritance
  - Recently available through use of Polygenic Score
- Measuring environments that matter
  - Family/parent investments
  - Meso/macro environments
- Research design considerations:
  - gene environment interaction vs. gene environment correlation



# Genome-wide association study identifies 74 loci associated with educational attainment

26 MAY 2016 | VOL 533 | NATURE | 539



**Figure 1 | Manhattan plot for EduYears associations ( $n = 293,723$ ).** The  $x$  axis is chromosomal position, and the  $y$  axis is the significance on a  $-\log_{10}$  scale (two-tailed test). The black dashed line shows the genome-

wide significance level ( $5 \times 10^{-8}$ ). The red crosses are the 74 approximately independent genome-wide significant associations (lead SNPs). The black dots labelled with rs numbers are the three SNPs identified in ref. 1.

Notes:  $N=300,000$ ; Replication Sample is 110,000

Robert M. La Follette School of  
Public Affairs

# Polygenic scores

---

- Select some fraction of the results of the GWAS (up to 100%)
- Use the estimated betas as “weights”
- Apply these weights to a new sample (i.e. HRS) that has genome wide data to create a summary score
- Interpretation: genetic propensity for a trait
- Advantage—predictive power
  - Education Polygenic Score  $R^2 \sim 6-7\%$



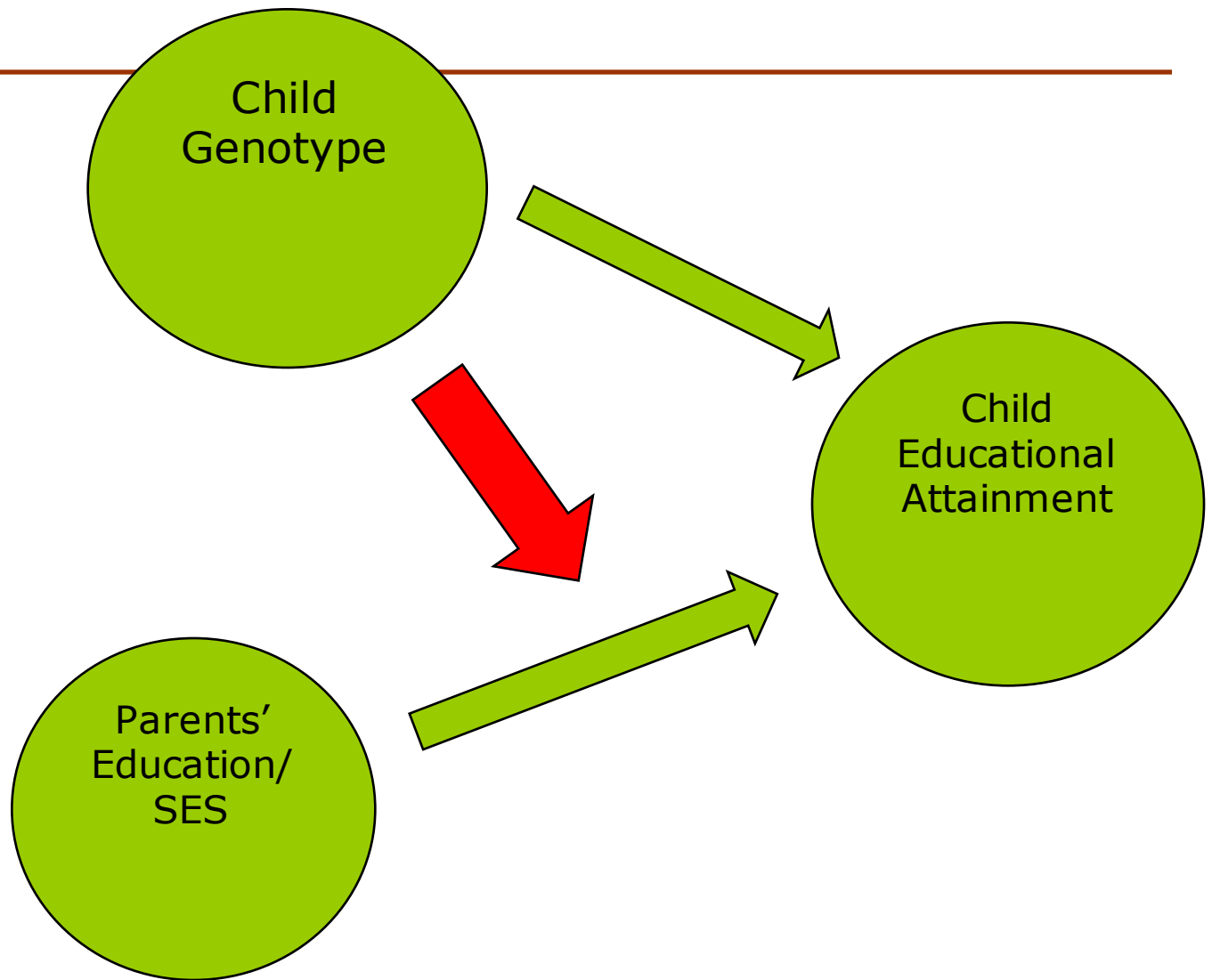
# Examples of GxE in predicting adult attainments

---

- Conley et al. (2015, Sociological Science)
  - HRS and FHS data
  - Child Polygenic Score X Maternal Education (“E”)
- Belsky et al. (2016, Psychological Science)
  - Dunedin Data
  - Child Polygenic Score X Family SES (“E”)
- Papageorge and Thom (2016, WP)
  - HRS data
  - Child Polygenic Score X Family SES (“E”)

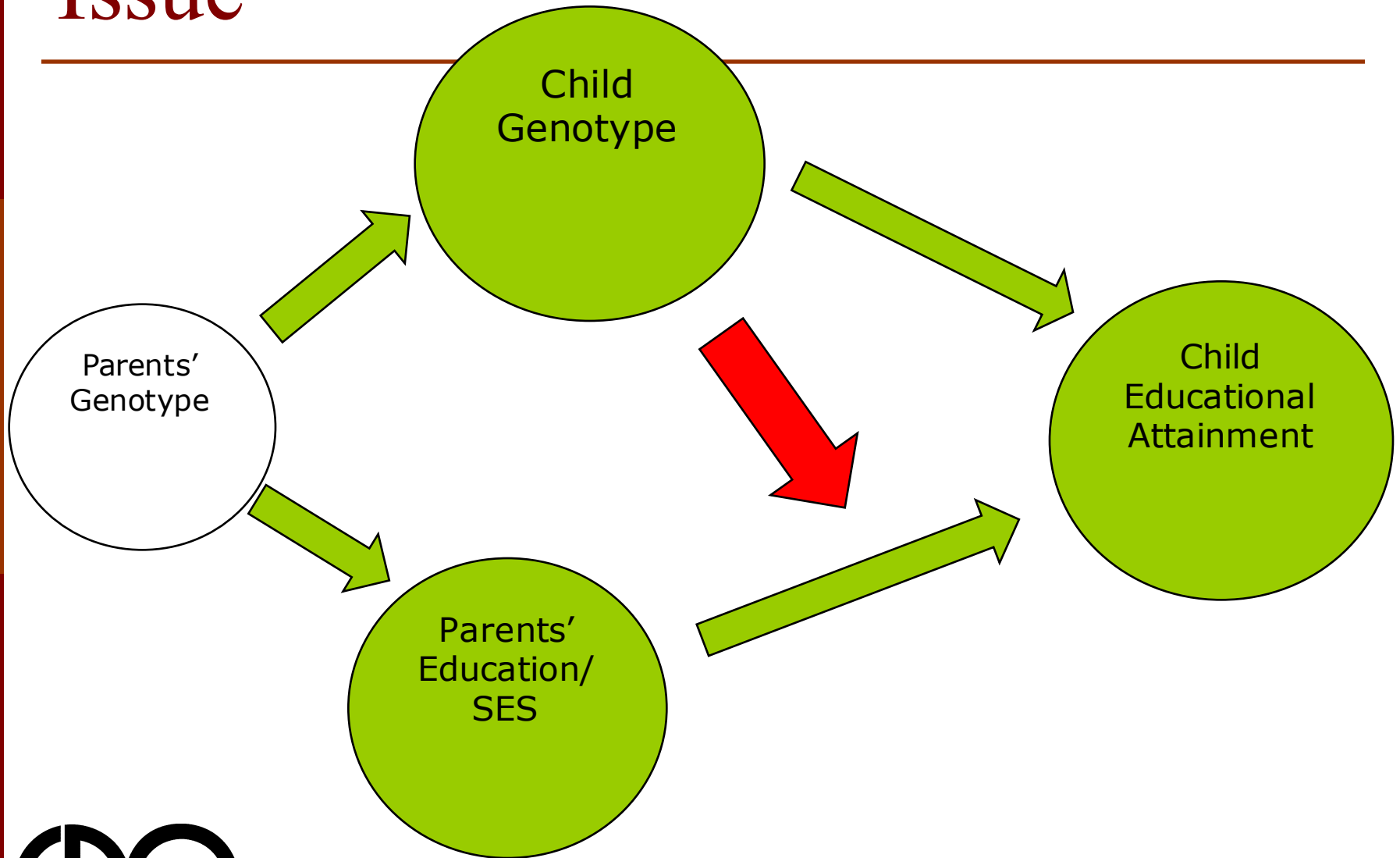




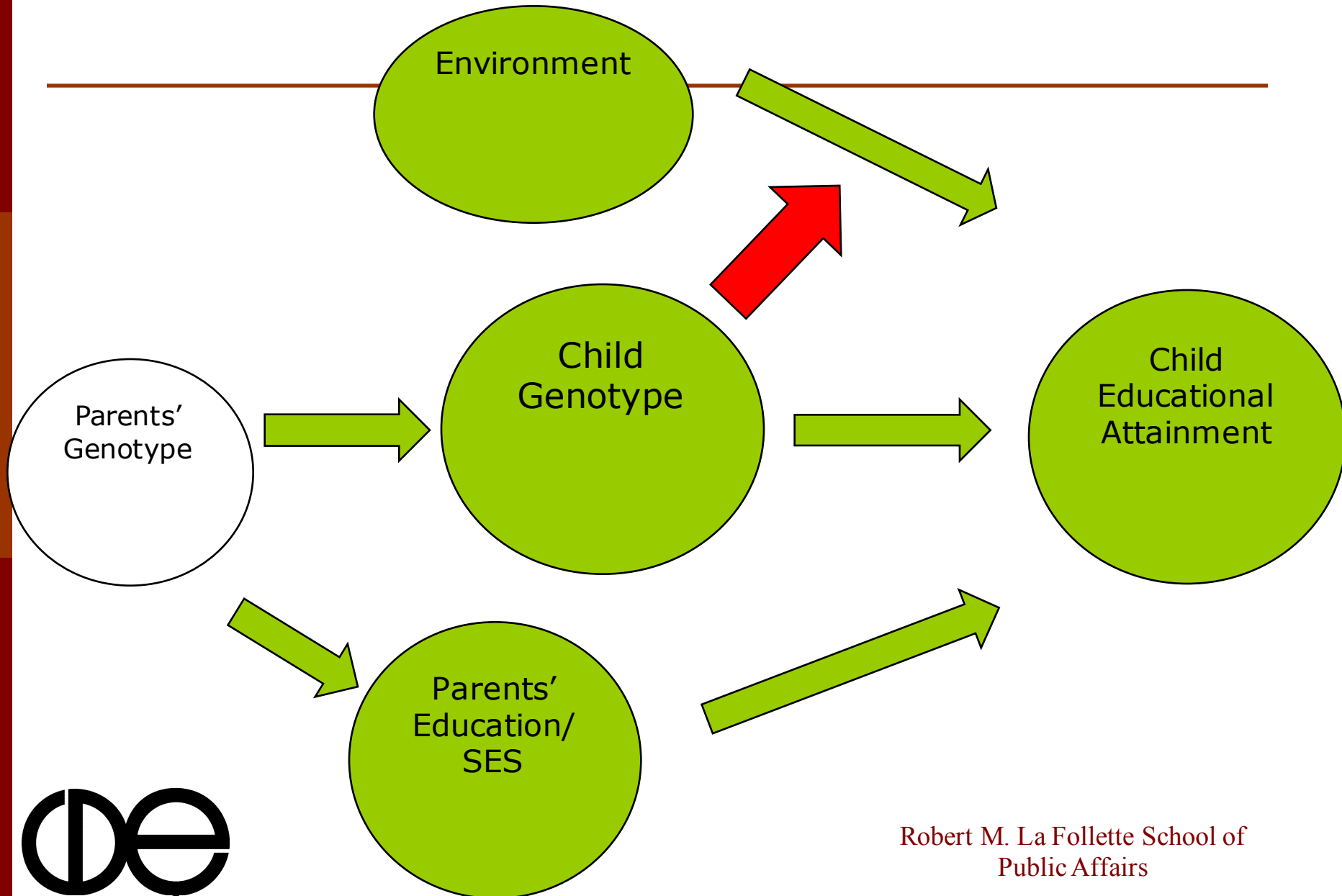


# Issue

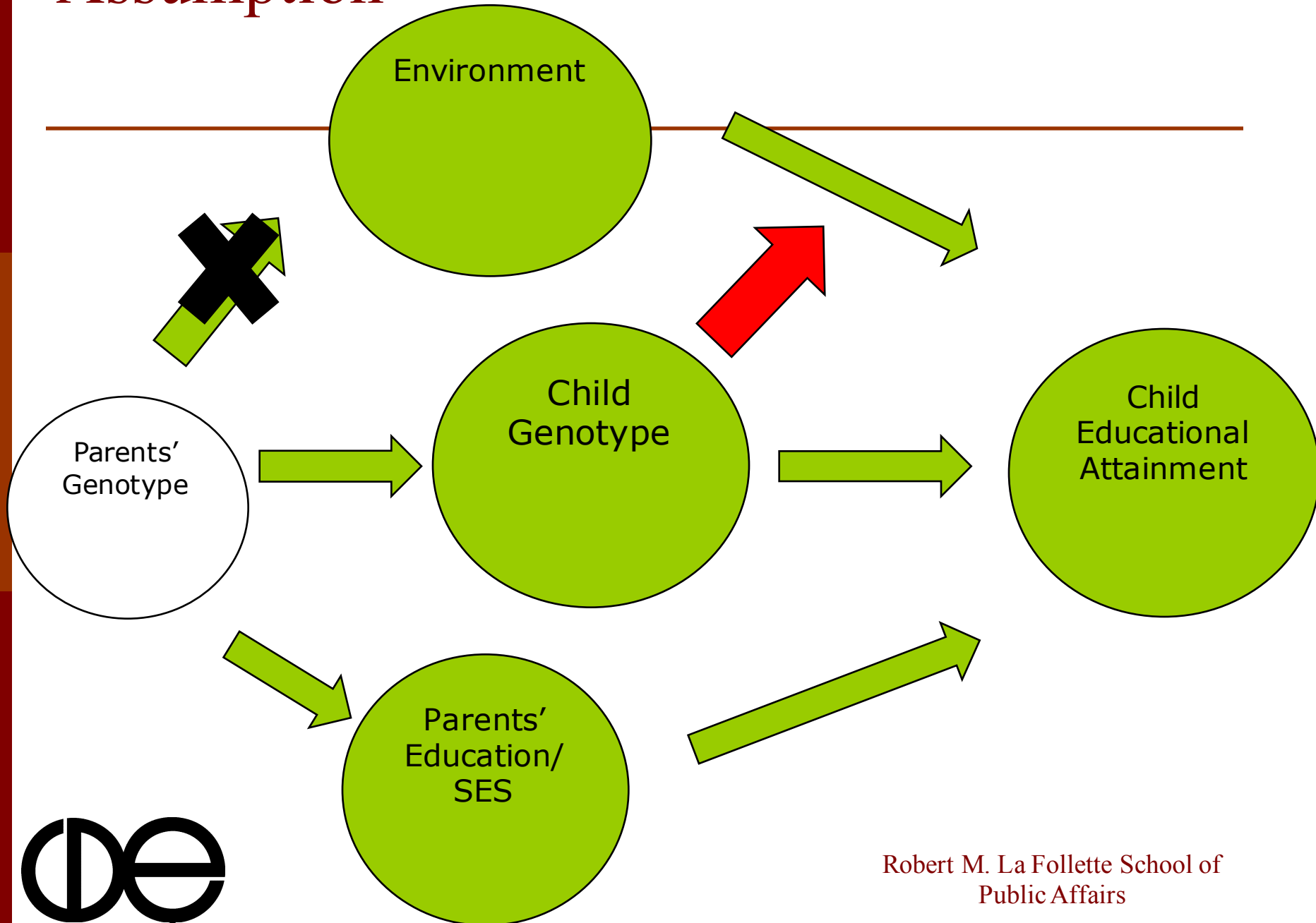
---



# Current contribution—add exogenous “E”

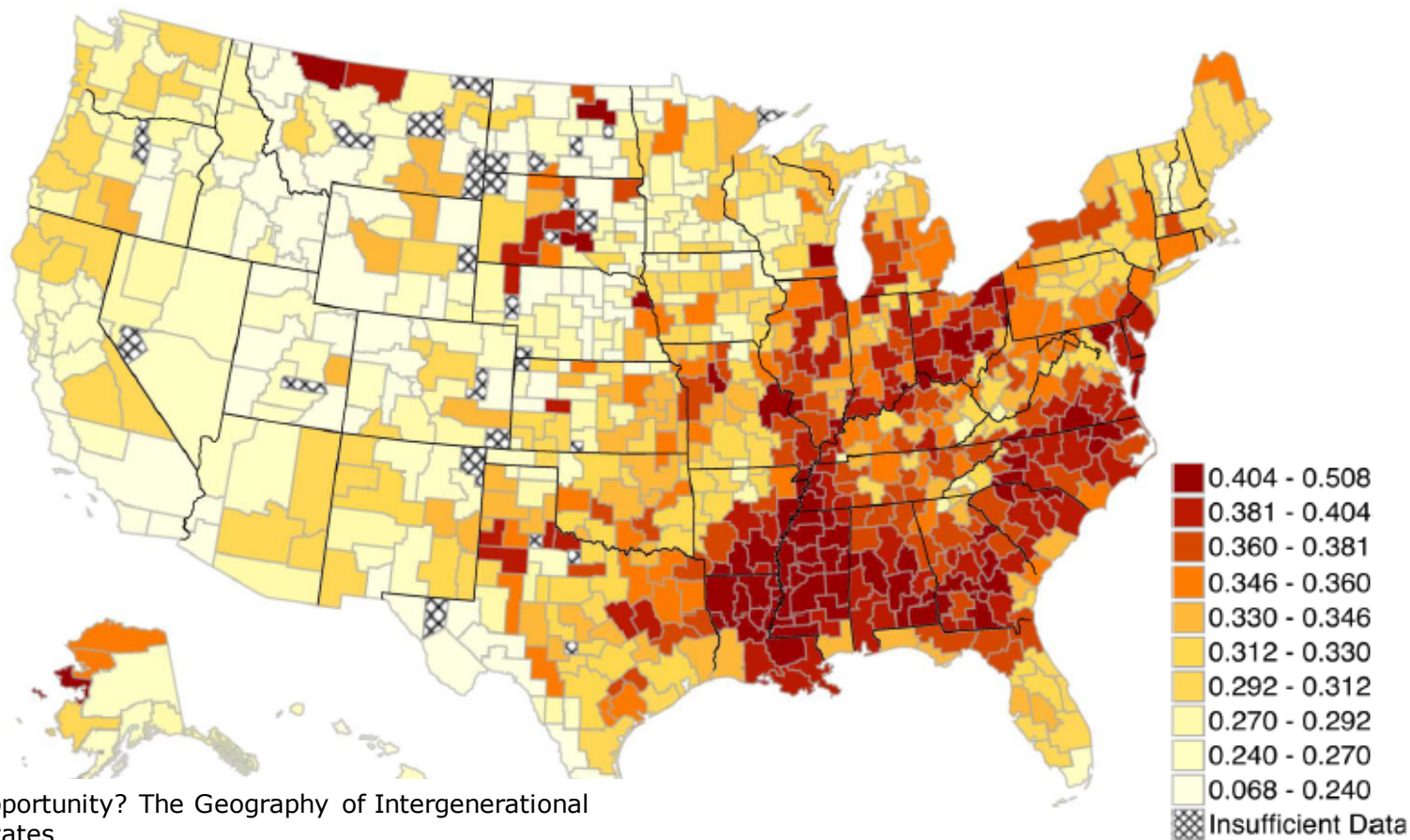


# Assumption



# Macro environmental context

Relative Mobility Across Areas in the U.S.  
Rank-Rank Slopes ( $Y_{100} - Y_0$ ) by Commuting Zone



Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States

Raj Chetty, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez  
*Quarterly Journal of Economics* 129(4): 1553-1623, 2014

$\bar{y}_{25} = -0.68$  (unweighted),  $-0.61$  (pop-weighted)

# Move to macro environmental context

---

- Ideally would have state x year exposure to mobility environments (e.g. Chetty)
- Measure the rank-rank correlation in child-parent years of schooling at the state-decade level in the HRS data
  - Estimate 20K regressions



# Gene-Environment Correlation?

VARIABLES	State Mobility	State Mobility	State Mobility	State Mobility
Education Genetic Score		0.000 (0.001)	-0.001 (0.001)	-0.001 (0.001)
Born in the South	-0.148*** (0.019)	-0.148*** (0.019)	-0.145*** (0.019)	-0.145*** (0.019)
Born in the West	-0.078*** (0.023)	-0.078*** (0.023)	-0.075*** (0.023)	-0.076*** (0.023)
Born in the Mid West	-0.065*** (0.015)	-0.065*** (0.015)	-0.063*** (0.015)	-0.063*** (0.015)
Born Other	-0.220*** (0.052)	-0.220*** (0.053)	-0.198*** (0.048)	-0.198*** (0.048)
Father Education	0.001 (0.002)			0.001 (0.002)
2nd Decade of Birth	0.059*** (0.020)	0.059*** (0.020)	0.059*** (0.020)	0.059*** (0.020)
3rd Decade of Birth	0.052** (0.023)	0.052** (0.023)	0.053** (0.022)	0.053** (0.022)
4th Decade of Birth	0.073*** (0.028)	0.073*** (0.028)	0.075*** (0.028)	0.075*** (0.028)
Male	-0.003 (0.002)	-0.003 (0.002)	-0.002 (0.002)	-0.002 (0.002)
Child Birth Year (Centered)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
PC			-0.111 (1.155)	-0.092 (1.138)
Constant	0.782*** (0.017)	0.784*** (0.017)	0.777*** (0.018)	0.775*** (0.018)
Observations	6,468	6,468	6,468	6,468
R-squared	0.392	0.392	0.396	0.396



# Gene Environment Interactions

	Education	Education	Education	Education
Mobility / Years-Years	1.621*** (0.588)	1.633*** (0.568)	0.476 (0.457)	0.487 (0.445)
Education Genetic Score	0.614*** (0.032)	1.172*** (0.232)	0.450*** (0.031)	0.880*** (0.239)
Father Years of Schooling			0.151*** (0.010)	0.151*** (0.010)
Mother Years of Schooling			0.175*** (0.013)	0.175*** (0.013)
Mobility X Genetic Score		-0.745** (0.316)		-0.574* (0.329)
Male	0.431*** (0.058)	0.432*** (0.058)	0.382*** (0.054)	0.383*** (0.055)
Constant	11.702*** (0.497)	11.707*** (0.478)	9.465*** (0.401)	9.471*** (0.389)
Observations	6,455	6,455	6,292	6,292
R-squared	0.108	0.109	0.253	0.254

Notes: Robust standard errors clustered at state-of-birth level in parentheses , \*\*\* p<0.01, \*\* p<0.05, \* p<0.1 Additional Controls: Decade of birth indicators, 20 Genetic Principal Components, Year of birth, region of birth indicators



# Conclusions

---

- Major shift in understanding of genetic influences underway
- A key point-of-integration with social science is in examining gene-environment interactions
- I focus on “E” that is (arguably) exogenous
- I find low mobility environments constrain genetic endowments
- Much evidence shows importance of gene-environment interplay, which suggests a need to incorporate this richness more generally, and enriches interpretations of intergenerational mobility coefficients.





---

Robert M. La Follette School of  
Public Affairs

# Associations with Education

	Education	Education	Education	Education	Education
Mobility / Years-Years	1.621*** (0.588)	0.678 (0.494)	0.741 (0.537)	0.424 (0.498)	0.476 (0.457)
Education Genetic Score	0.614*** (0.032)				0.450*** (0.031)
Father Years of Schooling		0.261*** (0.011)		0.166*** (0.010)	0.151*** (0.010)
Mother Years of Schooling			0.300*** (0.013)	0.184*** (0.013)	0.175*** (0.013)
Male	0.431*** (0.058)	0.388*** (0.054)	0.398*** (0.058)	0.389*** (0.055)	0.382*** (0.054)
Constant	11.702*** (0.497)	9.829*** (0.420)	9.201*** (0.471)	9.088*** (0.430)	9.465*** (0.401)
Observations	6,455	6,455	6,292	6,292	6,292
R-squared	0.108	0.191	0.184	0.221	0.253

Notes: Robust standard errors clustered at state-of-birth level in parentheses , \*\*\* p<0.01, \*\* p<0.05, \* p<0.1 Additional Controls: Decade of birth indicators, 20 Genetic Principal Components, Year of birth, region of birth indicators



# HRS Analysis Sample

Variable	Obs	Mean	Std Dev	Min	Max
Years of Education	6,945	13.41	2.45	0	17
Education Rank	6,945	0.57	0.29	0	1
Log Wealth in 2008	6,774	11.89	3.60	-5	17
Father Years of Education	6,958	10.14	3.67	0	17
Father Education Rank	6,958	0.54	0.28	0	1
Mother Years of Education	6,783	10.63	3.09	0	17
Mother Education Rank	6,783	0.54	0.28	0	1
Polygenic Score (std)	6,958	0.00	1.00	-4	3
Male	6,958	0.44	0.50	0	1
Birth Year	6,958	1940	8	1924	1963
Born in Northeast	6,958	0.22	0.42	0	1
Born in South	6,958	0.28	0.45	0	1
Born in West	6,958	0.09	0.29	0	1
Born in Midwest	6,958	0.39	0.49	0	1
Born other	6,958	0.02	0.14	0	1
Mobility (yrs/yrs)	6,958	0.75	0.10	0	1
Mobility (rank/rank)	6,958	0.64	0.11	0	1
Mobility(yrs std/ yrs std)	6,958	0.66	0.13	0	1
Mobility (yrs std state/decade)	6,958	0.65	0.11	0	1

Sample of genotyped respondents with parental education data and state of birth data

Robert M. La Follette School of Public Affairs

