Education and Innovation

Ufuk Akcigit
University of Chicago

ECA Talks - World Bank

18 January 2023

Introduction

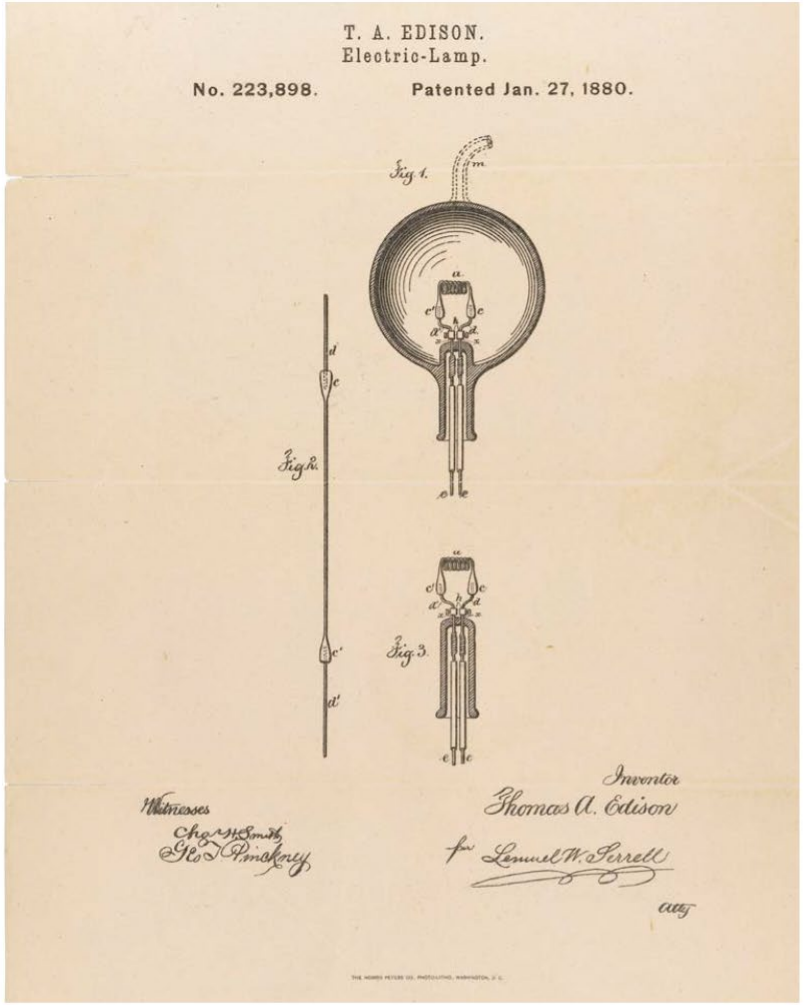
- Education has long been seen as an essential input for economic well-being.
- Education has mainly been considered as driver of labor productivity.
- What about its impact on growth through **“innovation”**?
- Major challenge: micro data on inventors and their educational backgrounds.

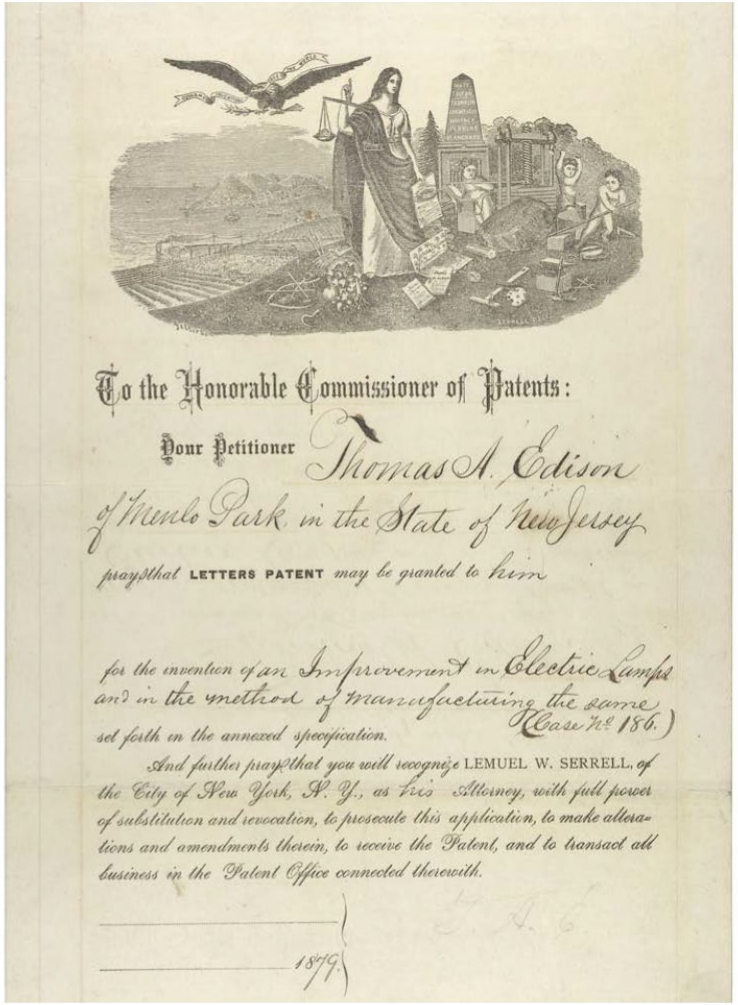
United States

Akcigit, Grigsby, and Nicholas (2017)

Major data collection effort. We generate novel **microdata** to study regional performance as well as the background of the Inventors of the Golden Age.

- Digitize the USPTO patents (OCR + hand entry).
- Newly-released decennial census data (1880-1940) and merge.
- Present key facts about innovation at regional and individual levels.





To the Honorable Commissioner of Patents:

Your Petitioner Thomas A. Edison of Menlo Park in the State of New Jersey prays that LETTERS PATENT may be granted to him

for the invention of an Improvement in Electric Lamps and in the method of manufacturing the same set forth in the annexed specification. (Case No 186.)

And further pray that you will recognize LEMUEL W. SERRELL, of the City of New York, N. Y., as his Attorney, with full power of substitution and revocation, to prosecute this application, to make alterations and amendments therein, to receive the Patent, and to transact all business in the Patent Office connected therewith.

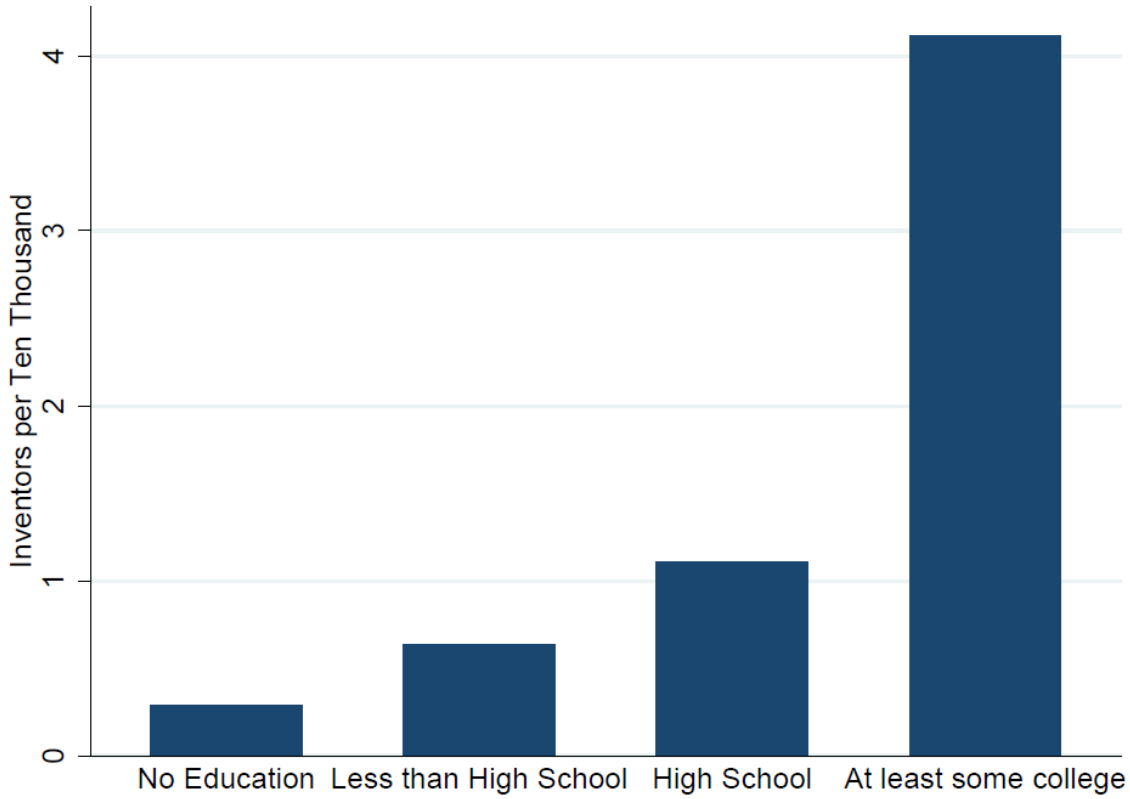
1879.

Historical patent files

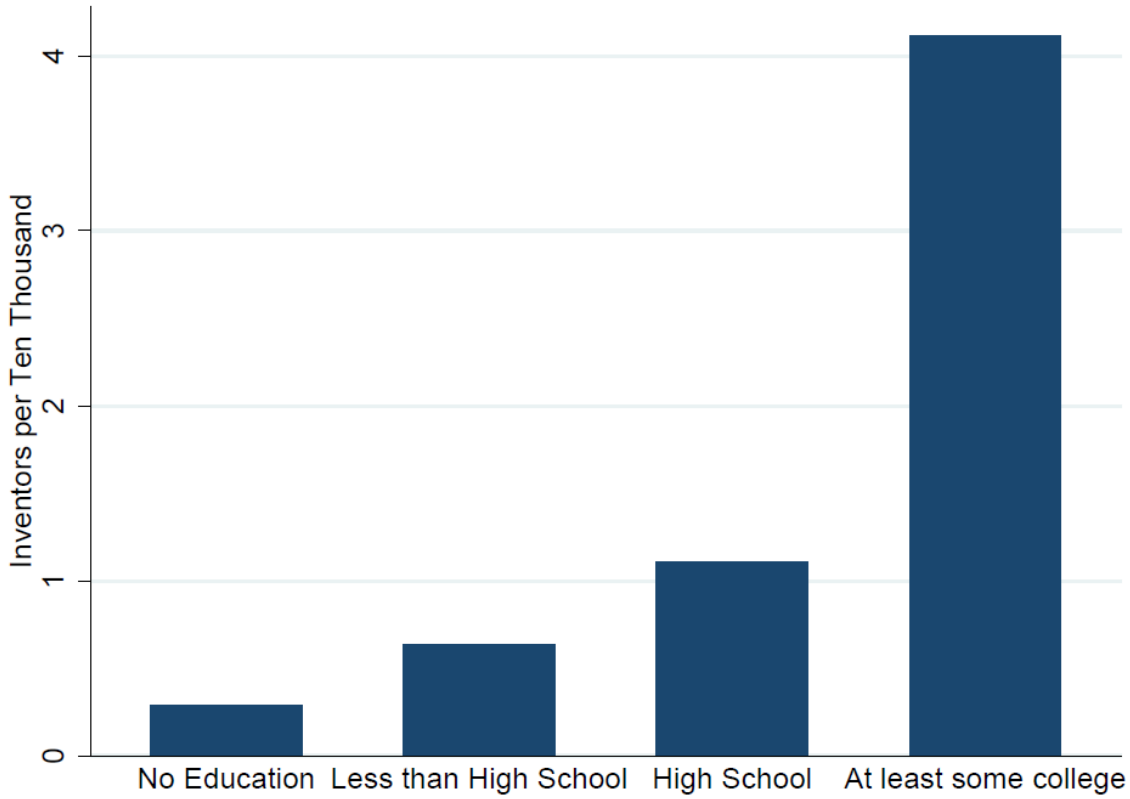
fname	sname	year	age	marital_st~s	birthplace	city
THOMAS	EDISON	1880	32	Married	OHIO	MENLO PARK
THOMAS	EDISON	1900	52	Married	OHIO	MENLO PARK
WILLIAM	WINE	1920	38	Married	VIRGINIA	TOLEDO WARD 4
ADIEL	DODGE	1940	48	Married	MISSOURI	ROCKFORD

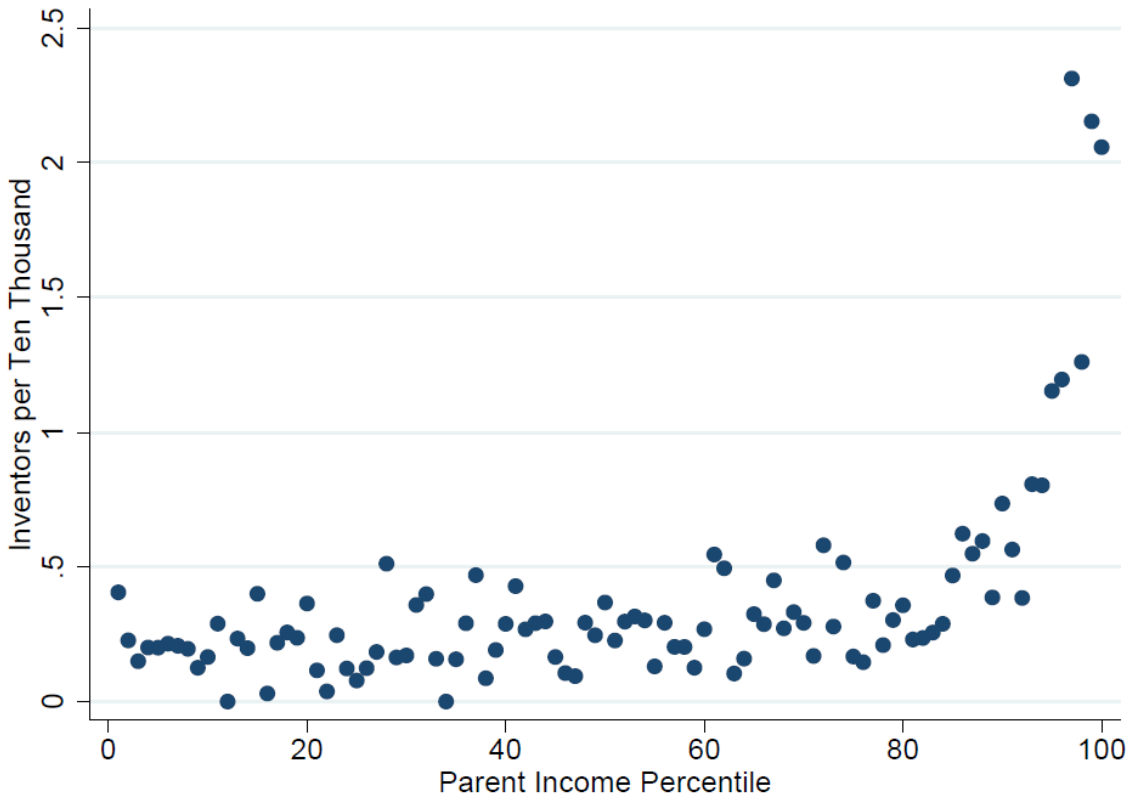
Historical Census records (1880-1940)

Who Becomes an Inventor?



Who Becomes an Inventor?





Who Becomes Regressions

Table: Who Becomes an Inventor?

	(1)	(2)	(3)	(4)
Father Income $90^{th} - 95^{th}$ %ile	0.411***	0.409***	0.297**	-0.070
	(0.119)	(0.120)	(0.124)	(0.127)
Father Income 95th %ile and above	1.084***	1.061***	0.770***	0.009
	(0.227)	(0.228)	(0.193)	(0.147)
Father Inventor		16.074**	15.859**	15.464**
		(7.545)	(7.544)	(7.552)
Father: High School Graduate			0.563***	-0.173
			(0.150)	(0.144)
Father: At least Some College			1.034***	-0.250**
			(0.165)	(0.102)
Self: High School Graduate				0.841***
				(0.111)
Self: At least Some College				3.558***
				(0.499)
Observations	82810258	82810258	82810258	82810258
Mean of Dep. Var.	1.091	1.091	1.091	1.091

Notes. Standard errors clustered at the state-level reported in parentheses. All regressions include state fixed effects, and controls for race, sex, migration status, and a quadratic in age. Columns (2) through (5) include indicators for father being between the 50th and 75th percentile of income, and between the 75th and 90th percentile of income as independent variables. The omitted categories are below median income and less than high school eduction.

Who Becomes Regressions

Table: Who Becomes an Inventor?

	(1)	(2)	(3)	(4)
Father Income $90^{th} - 95^{th}$ %ile	0.411***	0.409***	0.297**	-0.070
	(0.119)	(0.120)	(0.124)	(0.127)
Father Income 95th %ile and above	1.084***	1.061***	0.770***	0.009
	(0.227)	(0.228)	(0.193)	(0.147)
Father Inventor		16.074**	15.859**	15.464**
		(7.545)	(7.544)	(7.552)
Father: High School Graduate			0.563***	-0.173
			(0.150)	(0.144)
Father: At least Some College			1.034***	-0.250**
			(0.165)	(0.102)
Self: High School Graduate				0.841***
				(0.111)
Self: At least Some College				3.558***
				(0.499)
Observations	82810258	82810258	82810258	82810258
Mean of Dep. Var.	1.091	1.091	1.091	1.091

Notes. Standard errors clustered at the state-level reported in parentheses. All regressions include state fixed effects, and controls for race, sex, migration status, and a quadratic in age. Columns (2) through (5) include indicators for father being between the 50th and 75th percentile of income, and between the 75th and 90th percentile of income as independent variables. The omitted categories are below median income and less than high school eduction.

Who Becomes Regressions

Table: Who Becomes an Inventor?

	(1)	(2)	(3)	(4)
Father Income $90^{th} - 95^{th}$ %ile	0.411***	0.409***	0.297**	-0.070
	(0.119)	(0.120)	(0.124)	(0.127)
Father Income 95th %ile and above	1.084***	1.061***	0.770***	0.009
	(0.227)	(0.228)	(0.193)	(0.147)
Father Inventor		16.074**	15.859**	15.464**
		(7.545)	(7.544)	(7.552)
Father: High School Graduate			0.563***	-0.173
			(0.150)	(0.144)
Father: At least Some College			1.034***	-0.250**
			(0.165)	(0.102)
Self: High School Graduate				0.841***
				(0.111)
Self: At least Some College				3.558***
				(0.499)
Observations	82810258	82810258	82810258	82810258
Mean of Dep. Var.	1.091	1.091	1.091	1.091

Notes. Standard errors clustered at the state-level reported in parentheses. All regressions include state fixed effects, and controls for race, sex, migration status, and a quadratic in age. Columns (2) through (5) include indicators for father being between the 50th and 75th percentile of income, and between the 75th and 90th percentile of income as independent variables. The omitted categories are below median income and less than high school eduction.

Who Becomes Regressions

Table: Who Becomes an Inventor?

	(1)	(2)	(3)	(4)
Father Income $90^{th} - 95^{th}$ %ile	0.411***	0.409***	0.297**	-0.070
	(0.119)	(0.120)	(0.124)	(0.127)
Father Income 95th %ile and above	1.084***	1.061***	0.770***	0.009
	(0.227)	(0.228)	(0.193)	(0.147)
Father Inventor		16.074**	15.859**	15.464**
		(7.545)	(7.544)	(7.552)
Father: High School Graduate			0.563***	-0.173
			(0.150)	(0.144)
Father: At least Some College			1.034***	-0.250**
			(0.165)	(0.102)
Self: High School Graduate				0.841***
				(0.111)
Self: At least Some College				3.558***
				(0.499)
Observations	82810258	82810258	82810258	82810258
Mean of Dep. Var.	1.091	1.091	1.091	1.091

Notes. Standard errors clustered at the state-level reported in parentheses. All regressions include state fixed effects, and controls for race, sex, migration status, and a quadratic in age. Columns (2) through (5) include indicators for father being between the 50th and 75th percentile of income, and between the 75th and 90th percentile of income as independent variables. The omitted categories are below median income and less than high school eduction.

Finding:

Father's income was correlated with becoming an inventor.

This effect disappears once child's education is controlled for.

Who Becomes Regressions and State Characteristics

Table: Who Becomes an Inventor, Father Income, and State Characteristics

	(1)	(2)	(3)	(4)	(5)
Father's Incomec	0.036*	0.047***	0.052**	0.053**	0.042**
	(0.018)	(0.016)	(0.021)	(0.020)	(0.019)
Father's Income		-0.028***			
× Deposits/Cap		(0.009)			
Father's Income			0.048***		
× Gini Coefficient			(0.012)		
Father's Income				0.046***	
× 90/10 Ratio				(0.009)	
Father's Income					0.030**
× Top 10 Share					(0.012)
Observations	82810258	82650789	82810258	82810258	82810258
Mean of Dep. Var.	1.091	1.092	1.091	1.091	1.091
S.D. of Dep. Var.	104.430	104.479	104.430	104.430	104.430

Notes: Standard errors clustered at the state level reported in parentheses. Dependent variable in all regressions is an indicator equal to 100 if the individual is granted at least one patent in 1940. All regressions control for race, sex, international migrant status, father's age, occupation skill, and a quadratic in age. Father's income and state characteristics standardized to have zero mean and unit standard deviation. *Source*: FDIC, 1940 Census, USPTO Historical Patent Records.

Who Becomes Regressions and State Characteristics

Table: Who Becomes an Inventor, Father Income, and State Characteristics

	(1)	(2)	(3)	(4)	(5)
Father's Incomec	0.036*	0.047***	0.052**	0.053**	0.042**
	(0.018)	(0.016)	(0.021)	(0.020)	(0.019)
Father's Income		-0.028***			
× Deposits/Cap		(0.009)			
Father's Income			0.048***		
× Gini Coefficient			(0.012)		
Father's Income				0.046***	
× 90/10 Ratio				(0.009)	
Father's Income					0.030**
× Top 10 Share					(0.012)
Observations	82810258	82650789	82810258	82810258	82810258
Mean of Dep. Var.	1.091	1.092	1.091	1.091	1.091
S.D. of Dep. Var.	104.430	104.479	104.430	104.430	104.430

Notes: Standard errors clustered at the state level reported in parentheses. Dependent variable in all regressions is an indicator equal to 100 if the individual is granted at least one patent in 1940. All regressions control for race, sex, international migrant status, father's age, occupation skill, and a quadratic in age. Father's income and state characteristics standardized to have zero mean and unit standard deviation. *Source*: FDIC, 1940 Census, USPTO Historical Patent Records.

Who Becomes Regressions and State Characteristics

Table: Who Becomes an Inventor, Father Income, and State Characteristics

	(1)	(2)	(3)	(4)	(5)
Father's Incomec	0.036*	0.047***	0.052**	0.053**	0.042**
	(0.018)	(0.016)	(0.021)	(0.020)	(0.019)
Father's Income × Deposits/Cap		-0.028***			
		(0.009)			
Father's Income × Gini Coefficient			0.048***		
			(0.012)		
Father's Income × 90/10 Ratio				0.046***	
				(0.009)	
Father's Income × Top 10 Share					0.030**
					(0.012)
Observations	82810258	82650789	82810258	82810258	82810258
Mean of Dep. Var.	1.091	1.092	1.091	1.091	1.091
S.D. of Dep. Var.	104.430	104.479	104.430	104.430	104.430

Notes: Standard errors clustered at the state level reported in parentheses. Dependent variable in all regressions is an indicator equal to 100 if the individual is granted at least one patent in 1940. All regressions control for race, sex, international migrant status, father's age, occupation skill, and a quadratic in age. Father's income and state characteristics standardized to have zero mean and unit standard deviation. *Source*: FDIC, 1940 Census, USPTO Historical Patent Records.

Who Becomes Regressions and State Characteristics

Table: Who Becomes an Inventor, Father Income, and State Characteristics

	(1)	(2)	(3)	(4)	(5)
Father's Incomec	0.036*	0.047***	0.052**	0.053**	0.042**
	(0.018)	(0.016)	(0.021)	(0.020)	(0.019)
Father's Income × Deposits/Cap		-0.028***			
		(0.009)			
Father's Income × Gini Coefficient			0.048***		
			(0.012)		
Father's Income × 90/10 Ratio				0.046***	
				(0.009)	
Father's Income × Top 10 Share					0.030**
					(0.012)
Observations	82810258	82650789	82810258	82810258	82810258
Mean of Dep. Var.	1.091	1.092	1.091	1.091	1.091
S.D. of Dep. Var.	104.430	104.479	104.430	104.430	104.430

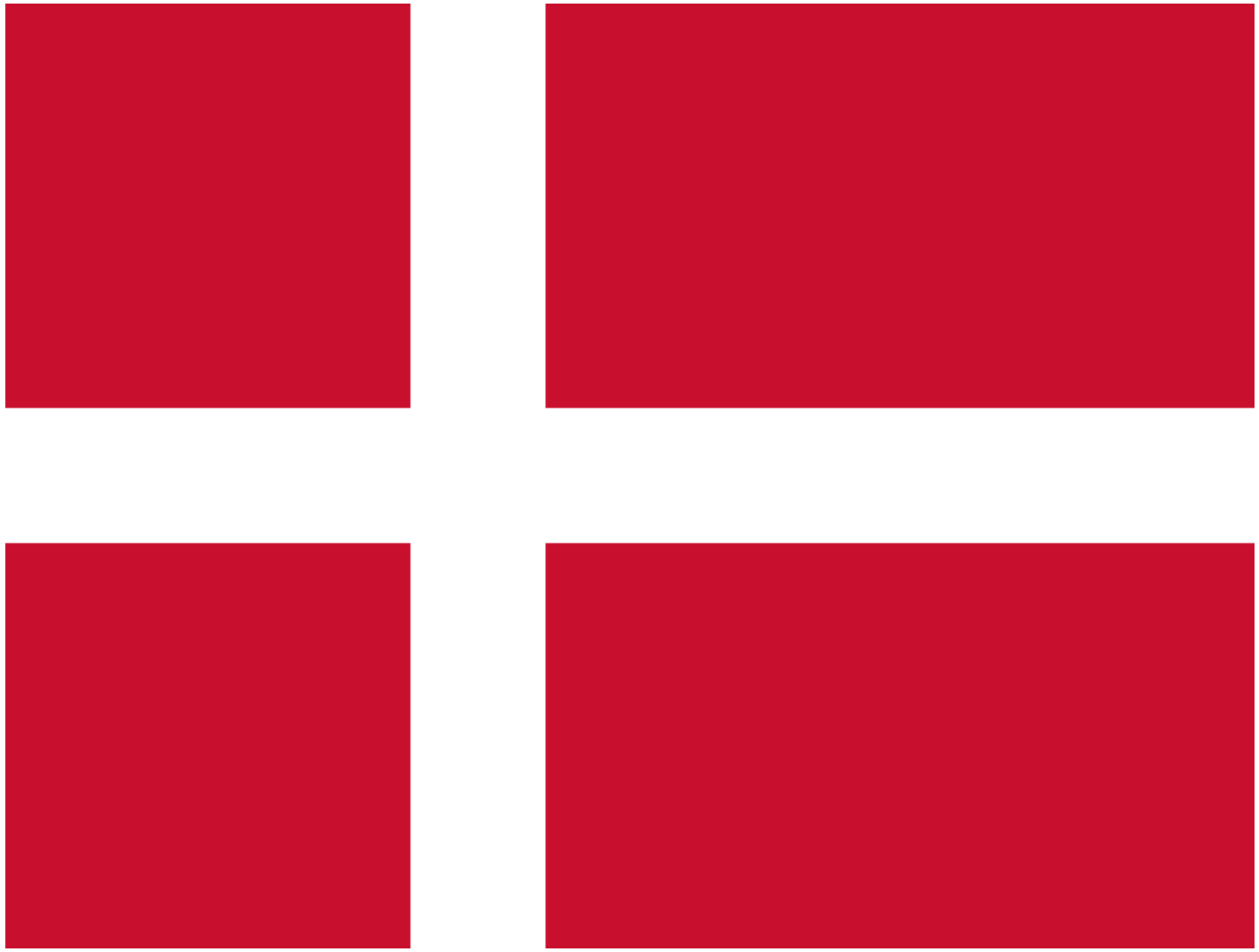
Notes: Standard errors clustered at the state level reported in parentheses. Dependent variable in all regressions is an indicator equal to 100 if the individual is granted at least one patent in 1940. All regressions control for race, sex, international migrant status, father's age, occupation skill, and a quadratic in age. Father's income and state characteristics standardized to have zero mean and unit standard deviation. *Source*: FDIC, 1940 Census, USPTO Historical Patent Records.

Who Becomes Regressions and State Characteristics

To sum up:

- Parental income is strongly correlated with becoming an inventor but less so
  1. in more financially-developed states.
  2. in more equal states.

Denmark



Data

Micro Data from Danmarks Statistik and Treasury Department (DST) with:

- Detailed information on innovation policy
- Individual level income, IQ and education data
- Linked parental information
- Matched employer-employee data
- Patent data from European Patent Office (EPO)

Education Matters for Innovation

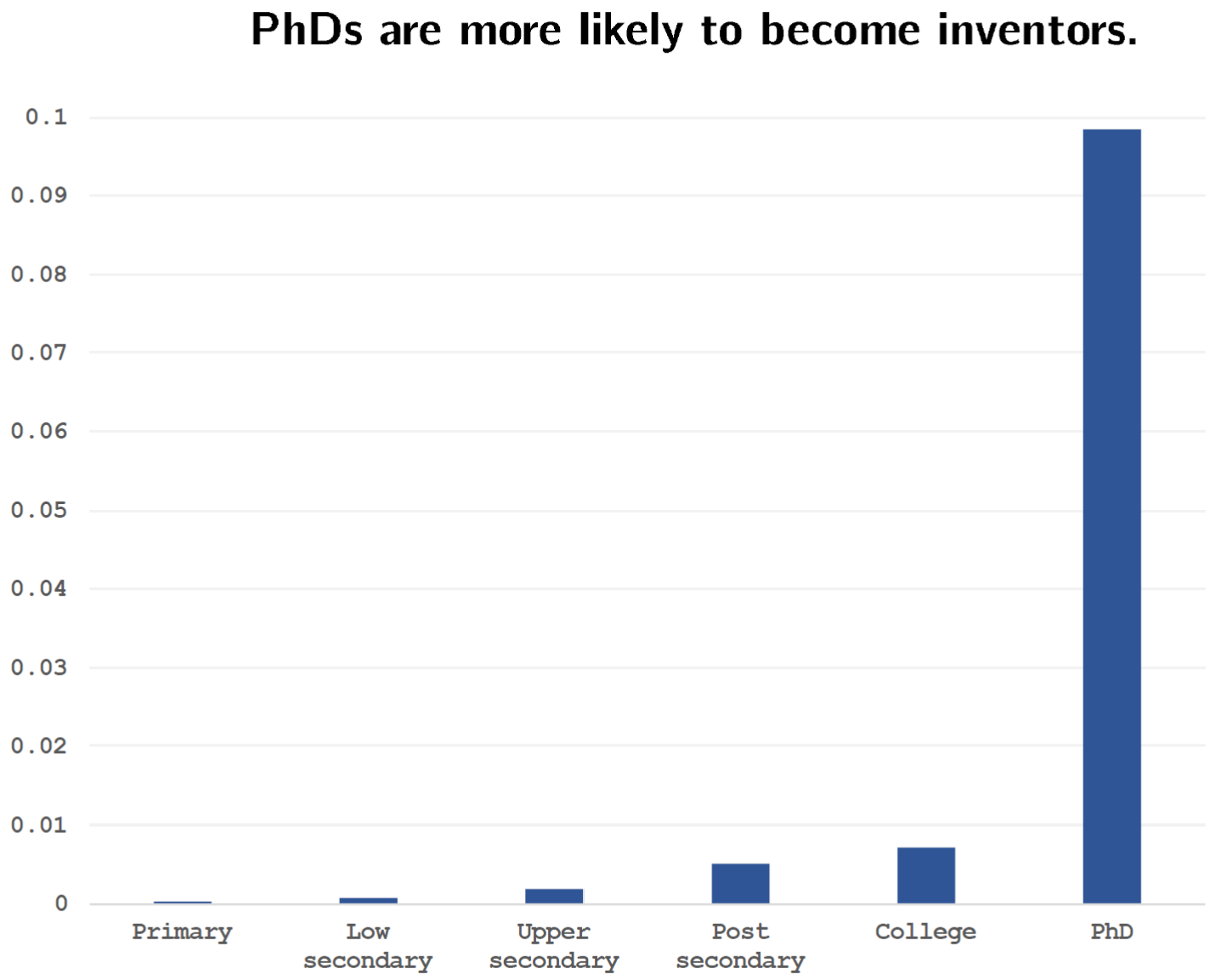


Figure: Fraction of individuals with at least one patent by education group.

Financial Frictions

Individuals with higher father's income are more likely to get a PhD.

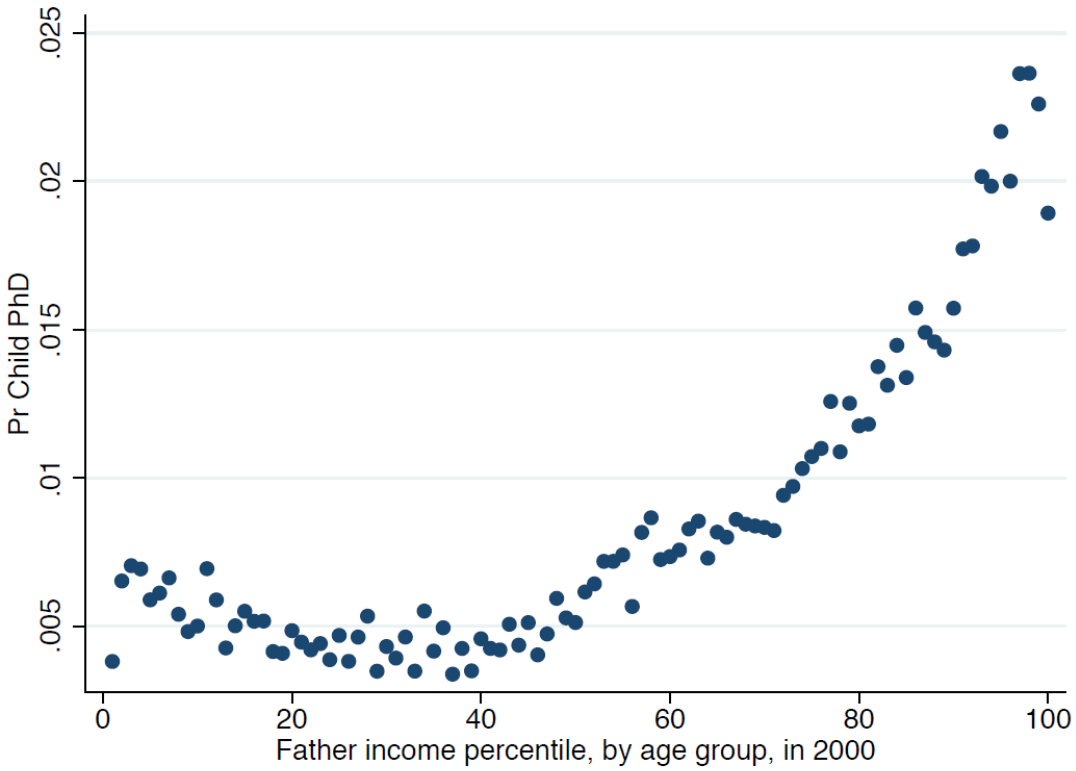


Figure: Proportion of individuals with a PhD by father's income percentile.

Talent Heterogeneity Matters

Individuals with higher IQ more likely to obtain a PhD.

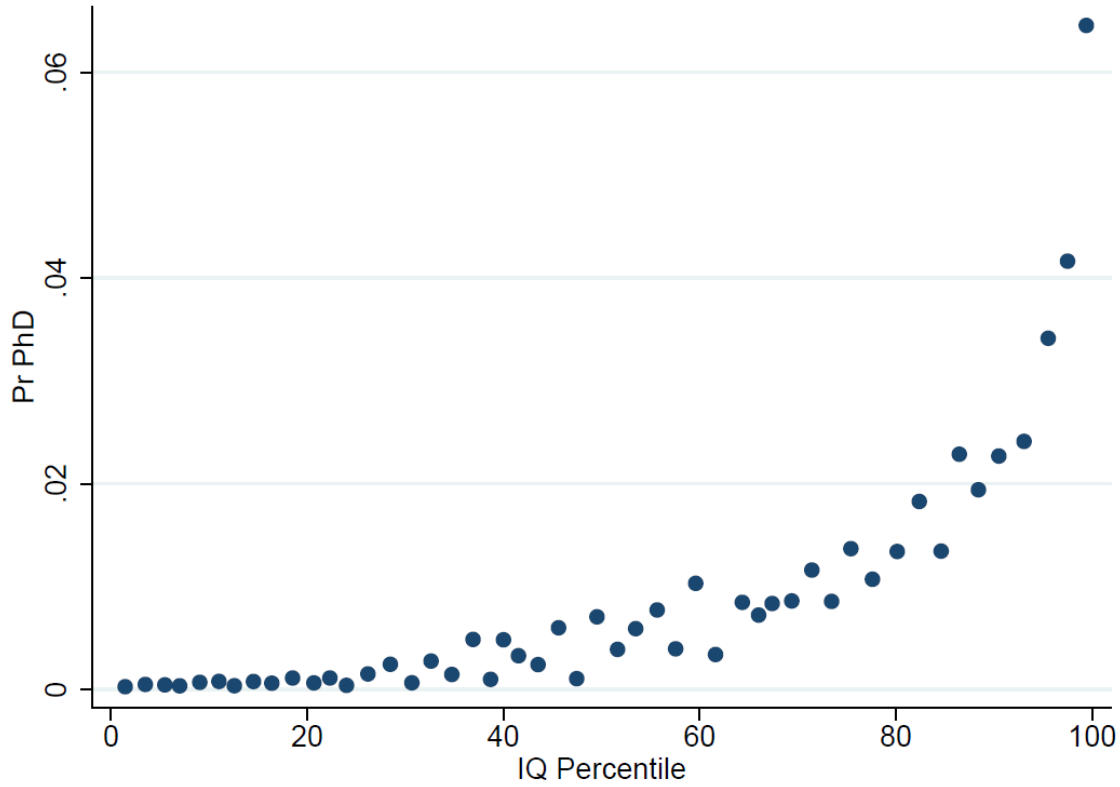
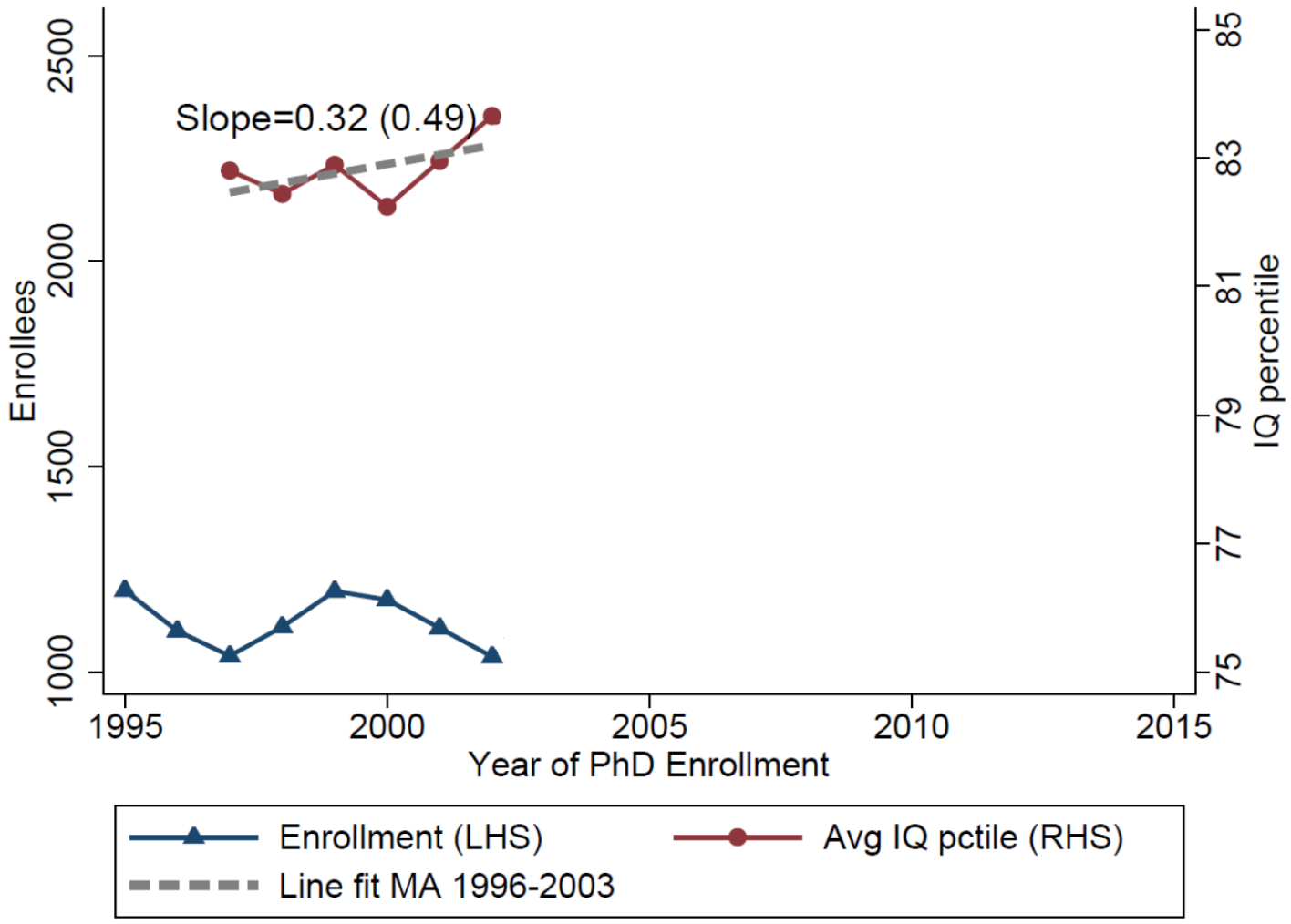
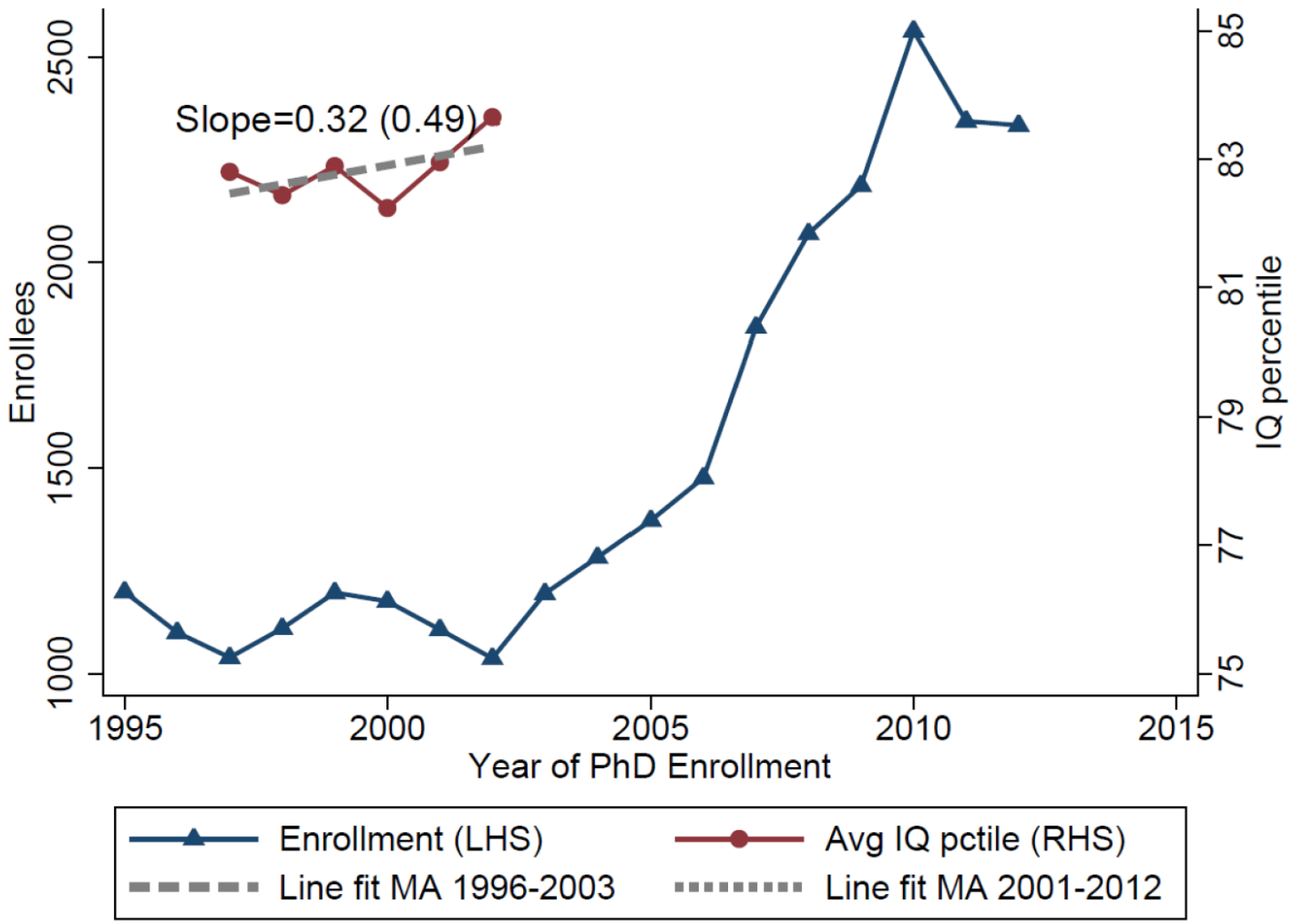


Figure: Fraction of individuals with a PhD by IQ percentile.

Tradeoff Size vs. Quality of Researchers

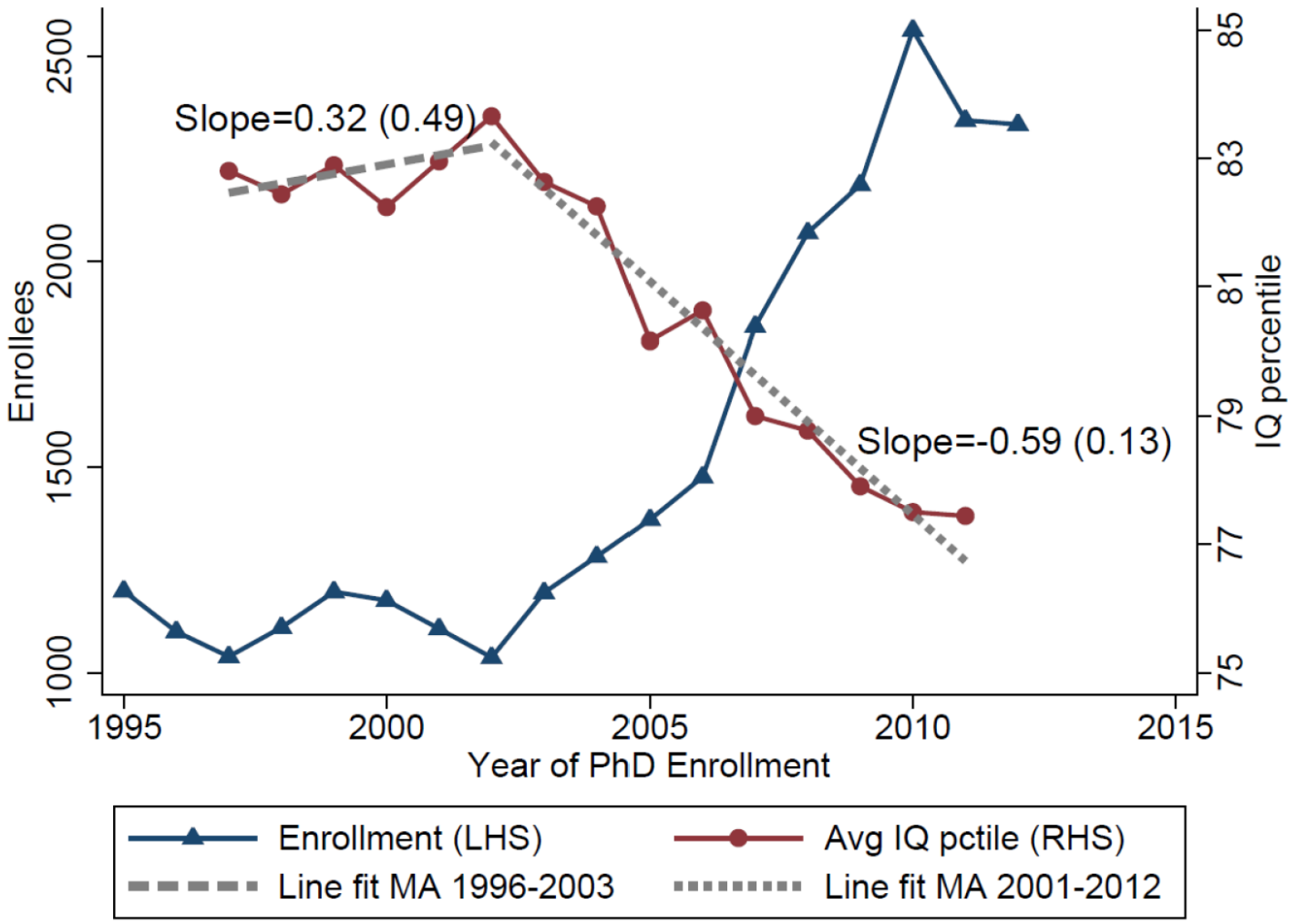


Tradeoff Size vs. Quality of Researchers



Tradeoff Size vs. Quality of Researchers

As the number of PhD enrollees increases, avg. IQ of enrollees declines.

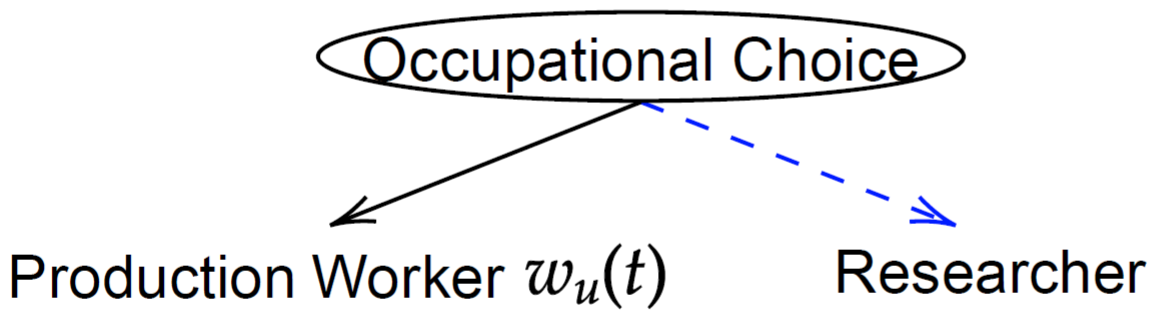


- What do these facts tell us about the talent (mis)allocation?

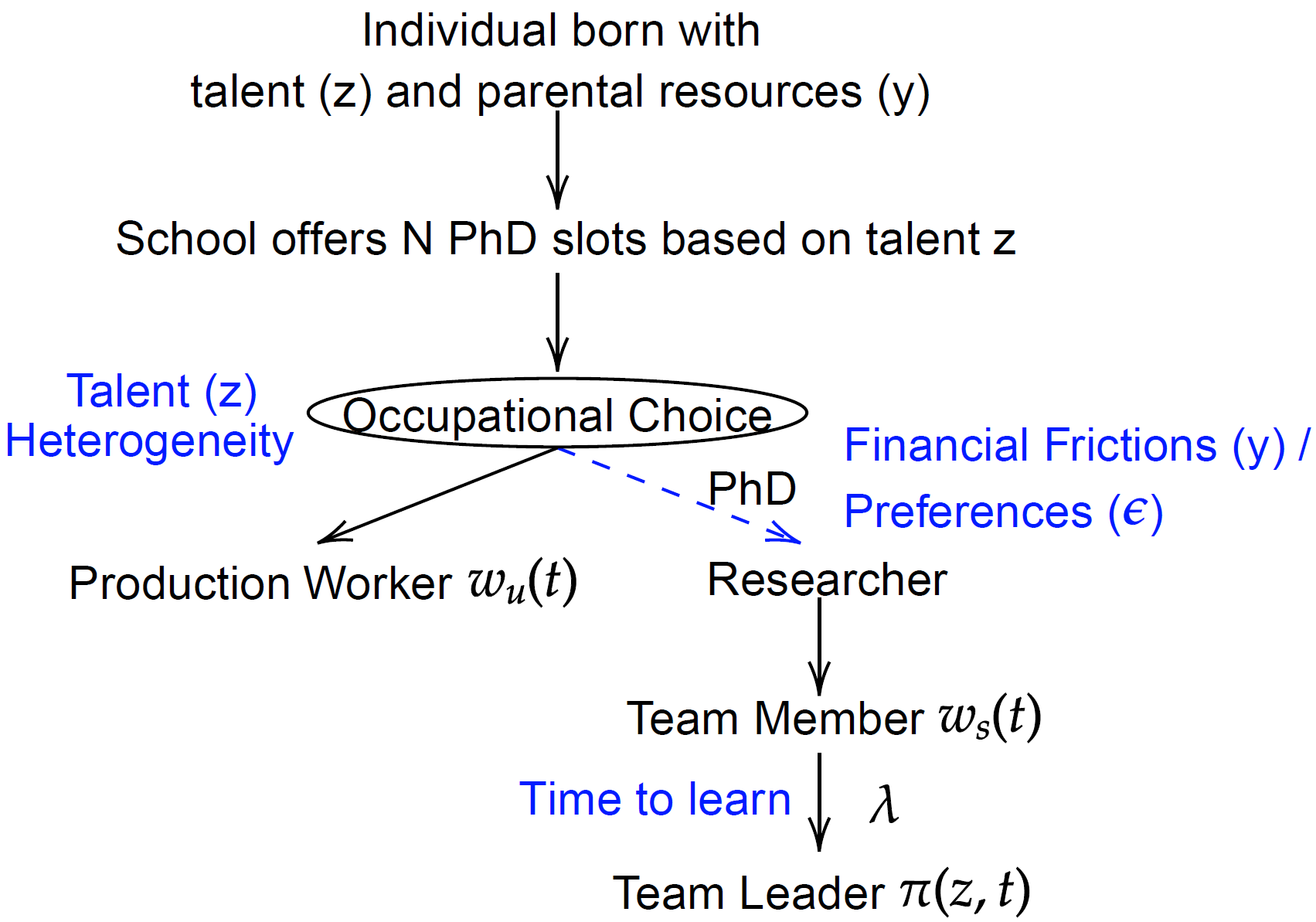
- What are the impacts of different policies on economic growth?
  - R&D subsidy
  - Education subsidy
  - Funding more education slots

- Need for a new theoretical framework for education and innovation.

Occupational Sorting - Models in the Literature

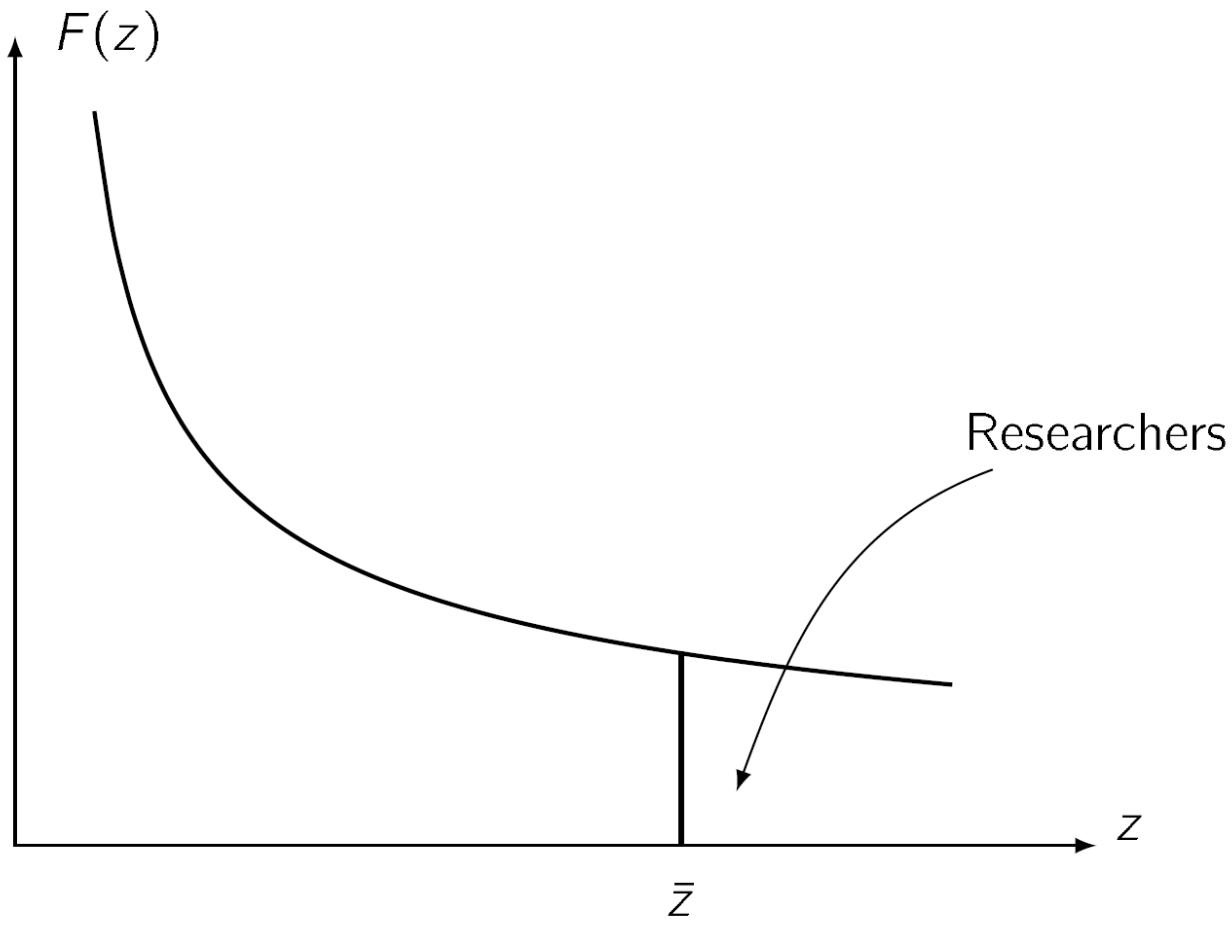


Occupational Sorting - Our Framework



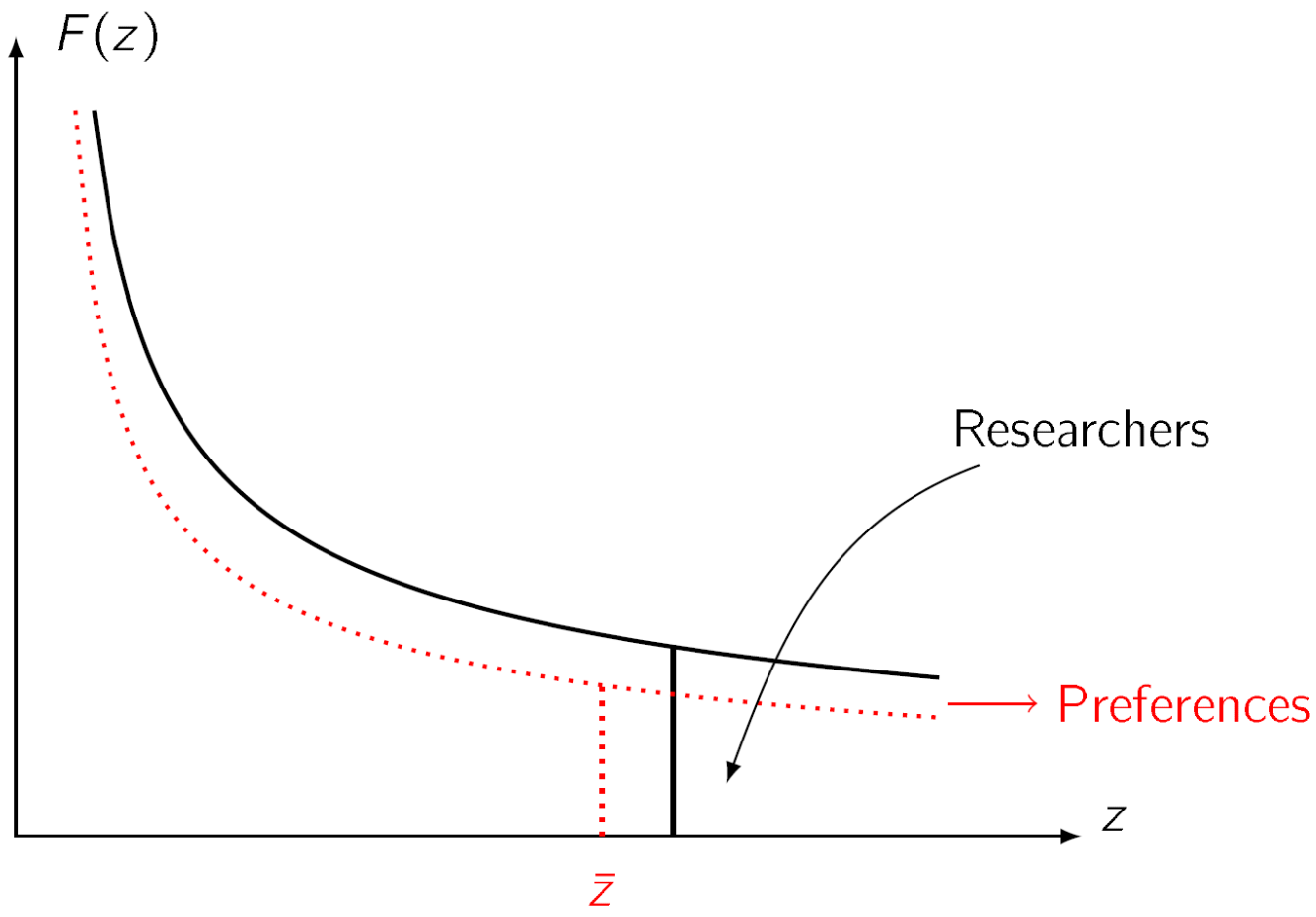
Occupational Choice: Production Worker or Researcher?

- Allocation of talent in absence of frictions



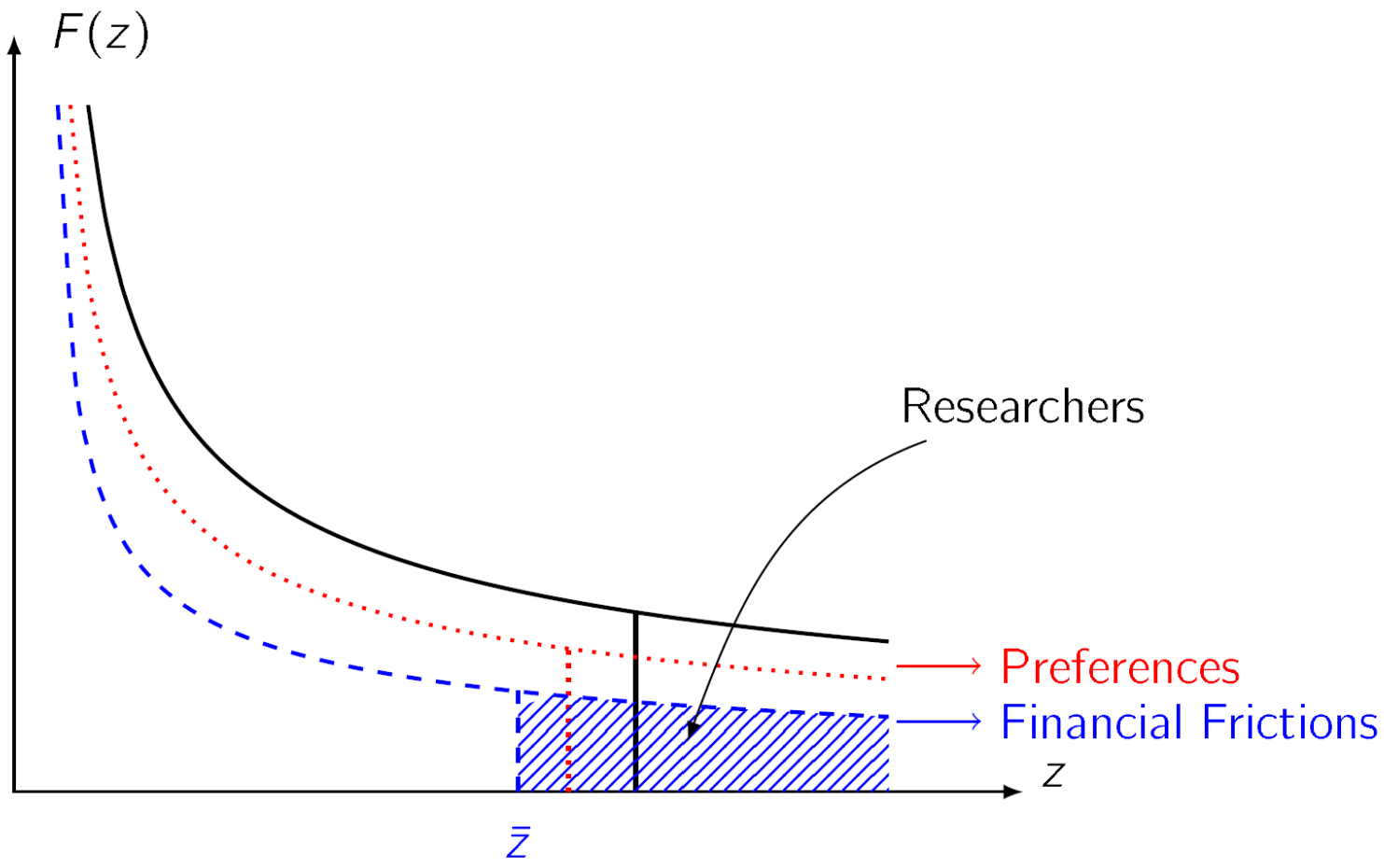
Occupational Choice: Production Worker or Researcher?

- Due to taste shock, some talented individuals do not want to become researchers



Occupational Choice: Production Worker or Researcher?

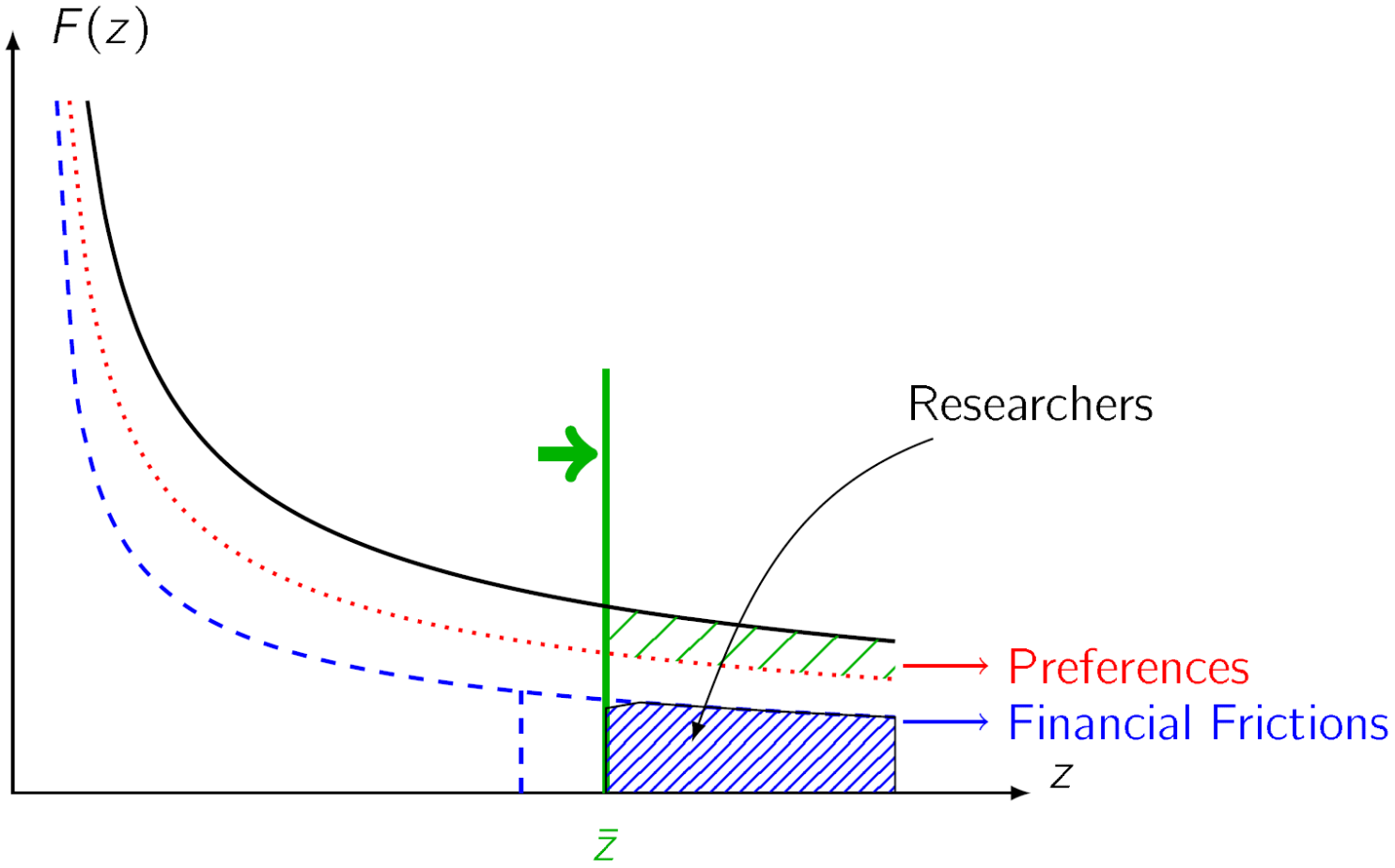
- Due to financial frictions, some talented individuals cannot afford to become researchers



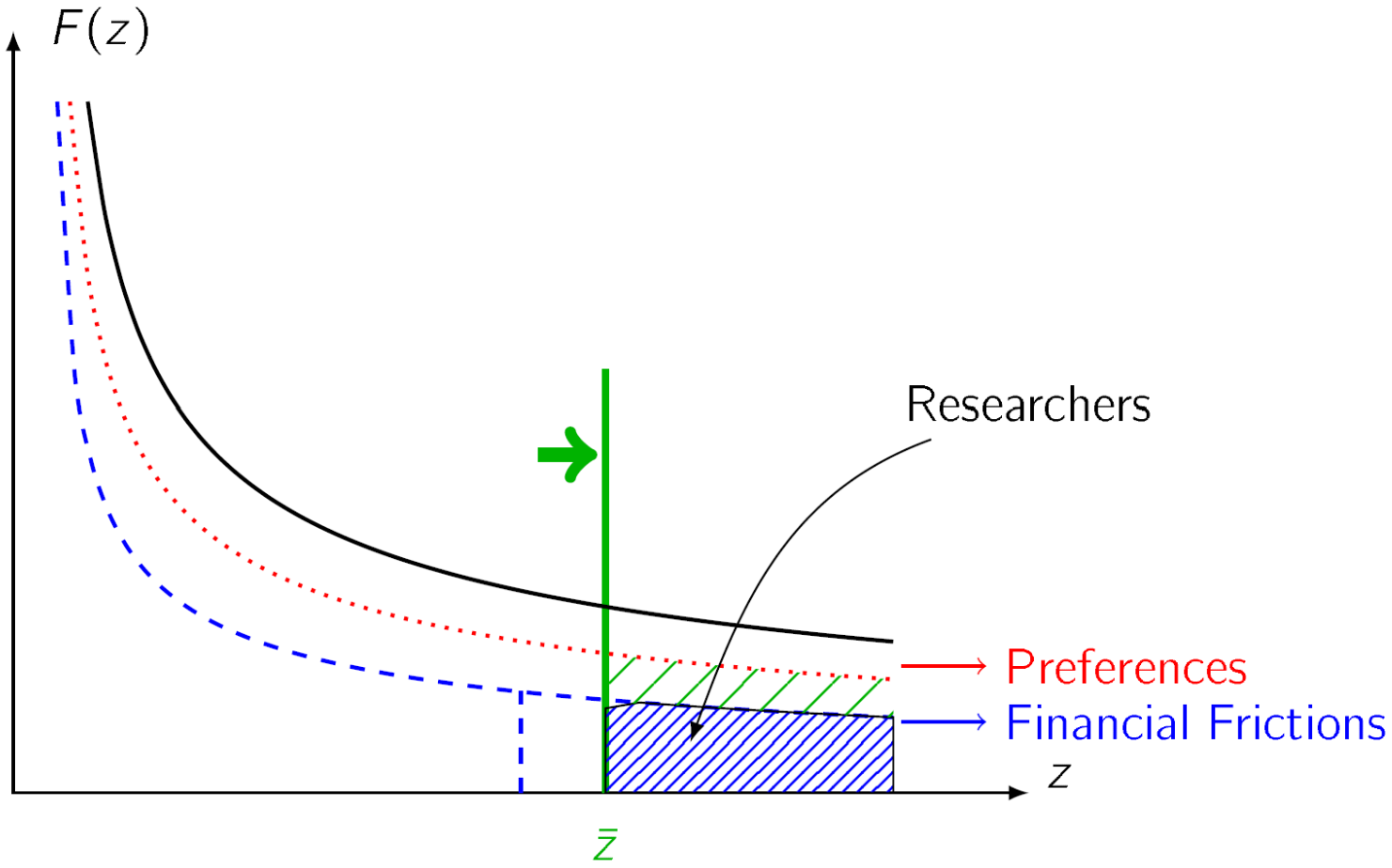
Impact of Different Policies

- R&D subsidy
- Education subsidy
- Funding more education slots

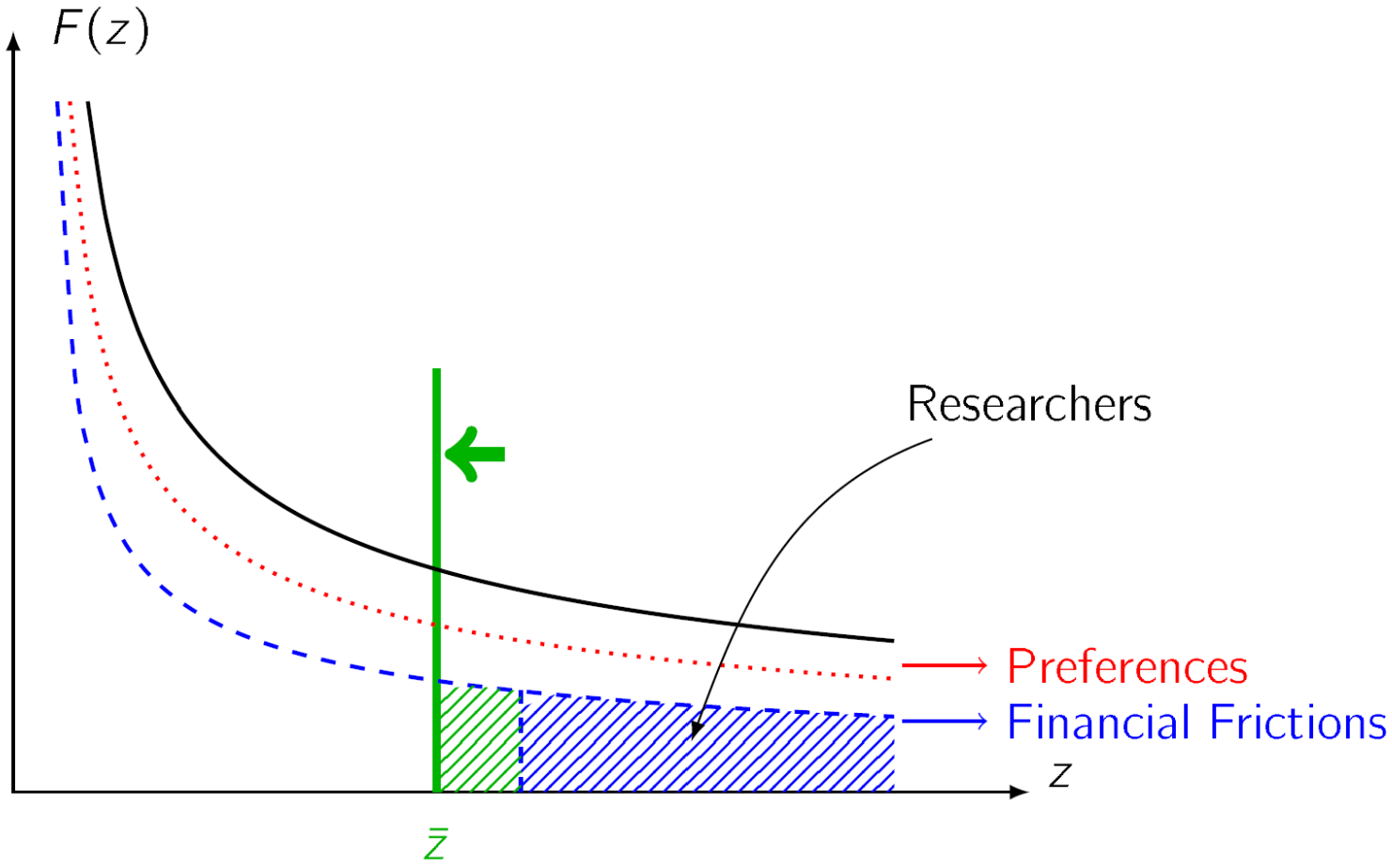
Policy Exercise 1: R&D subsidy



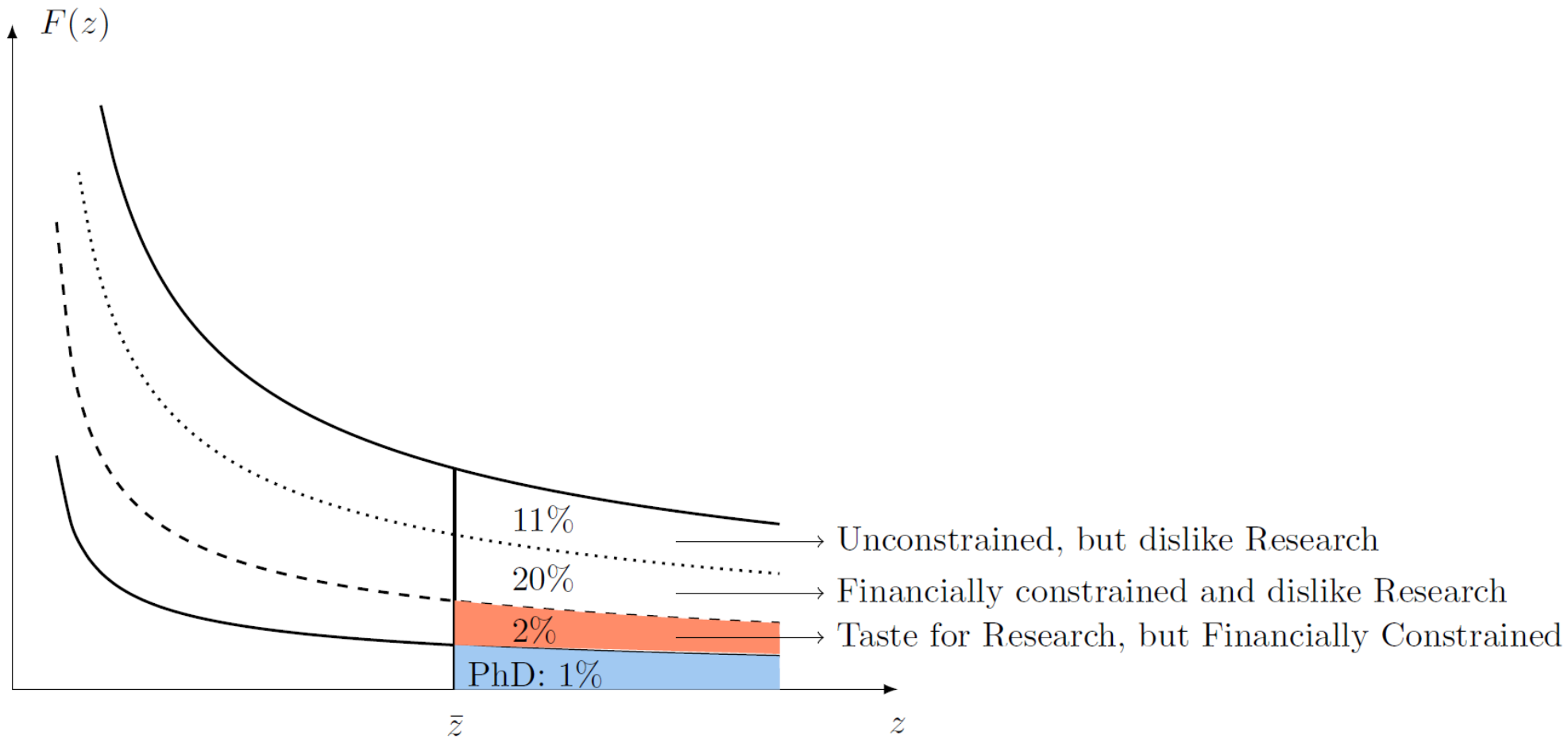
Policy Exercise 2: Subsidy to Cost of Education



Policy Exercise 3: Increase PhD Slots

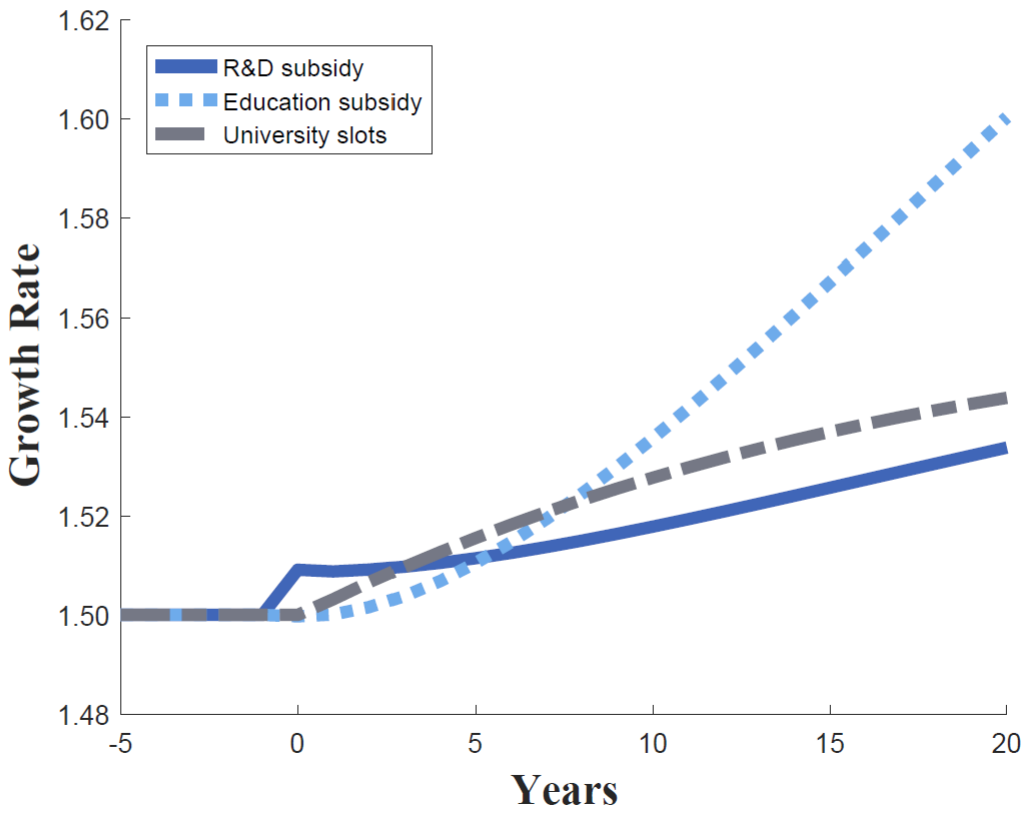


The allocation of talent in the economy



Policy Analysis

- Subsidy of 0.1% GDP
- Short Run: R&D policy most effective
- Long Run: Education subsidy most effective

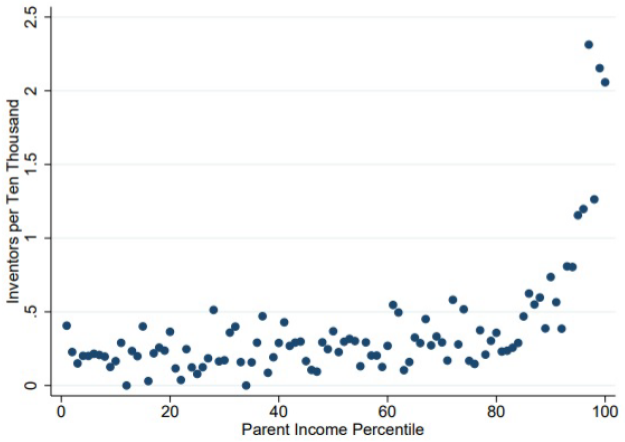


Finland

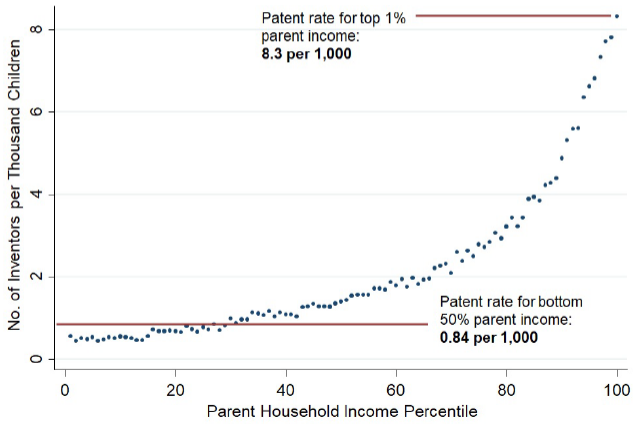
Parental Income vs Becoming Innovator: US, Finland

FIGURE 1: PARENTAL INCOME AND PROB(*Inventor*)

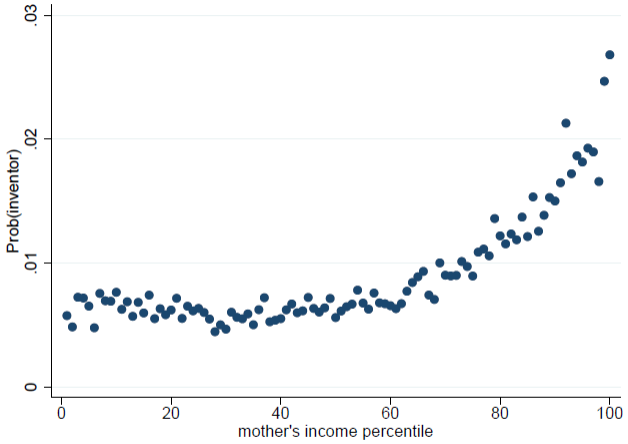
1A. 1930s U.S.



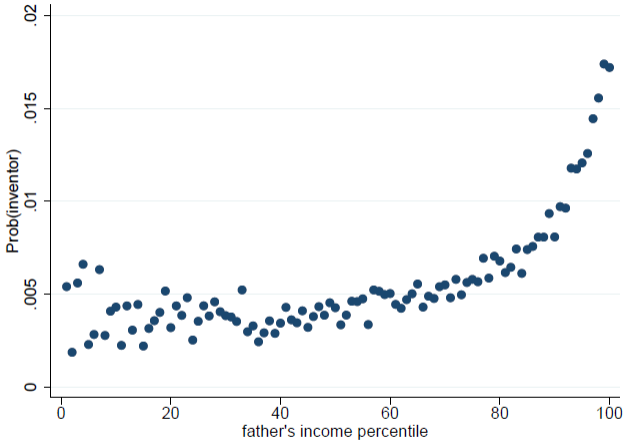
1B. 1980s U.S



1C. Maternal income and Prob(invent): Finland



1D. Paternal income and Prob(invent): Finland

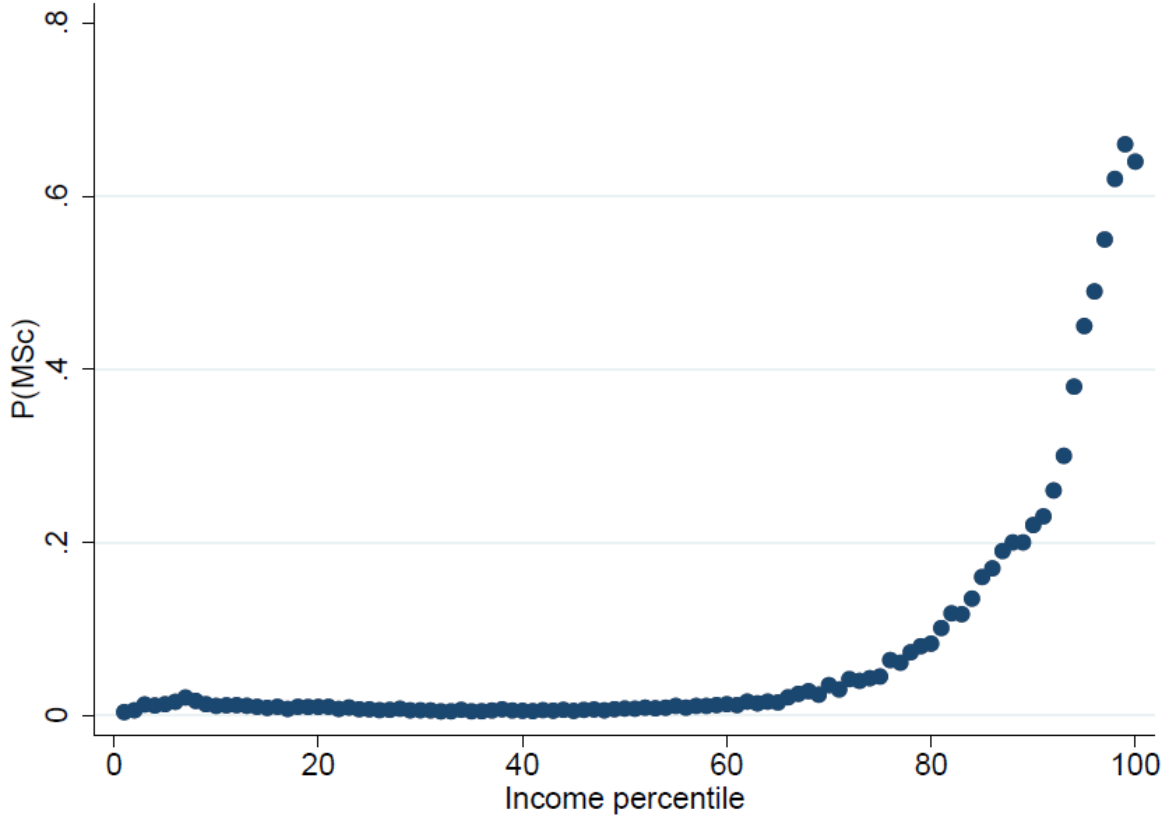


Sources: Figure 1A: Akcigit et al. (2017), Figure 1B: Bell et al. (2019), Figures 1C & 1D: own calculations.

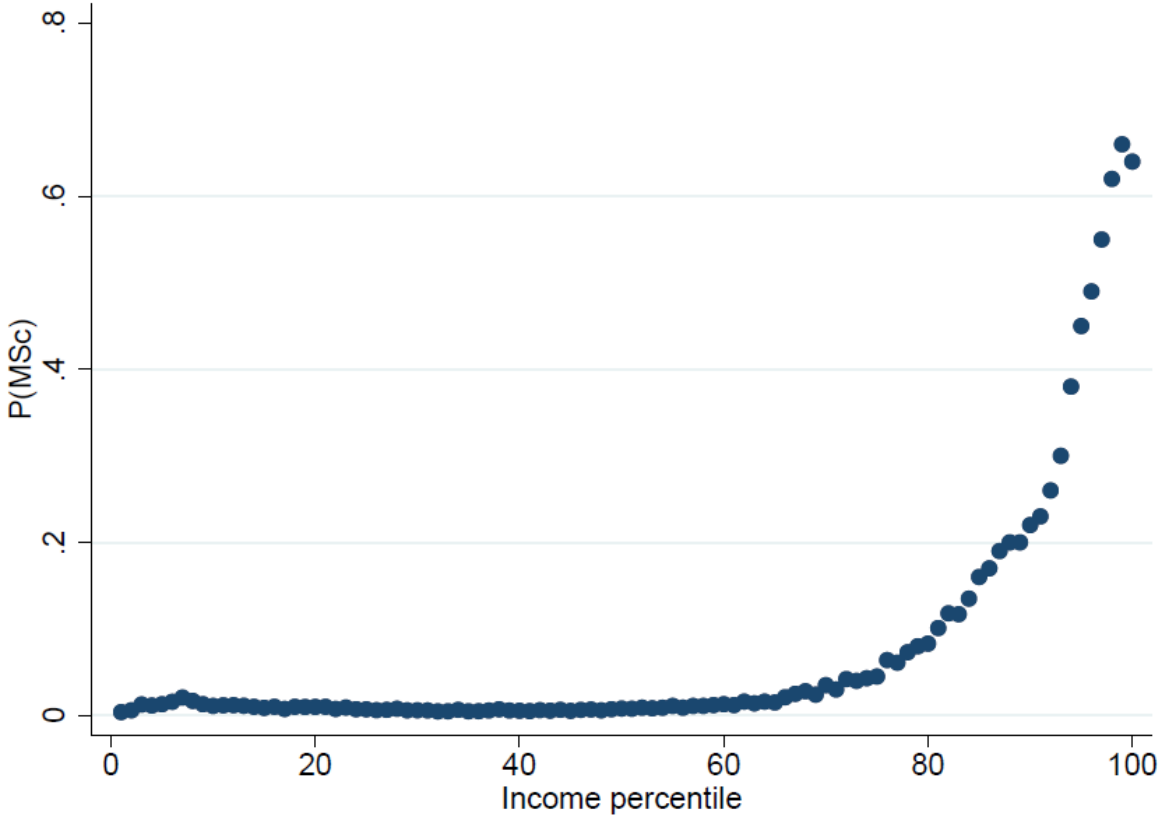
Parental Education and Parental Income

FIGURE 2: PARENTAL INCOME AND PROB(MSc_{parent})

2A. Maternal income & Prob(MSc_{mother})

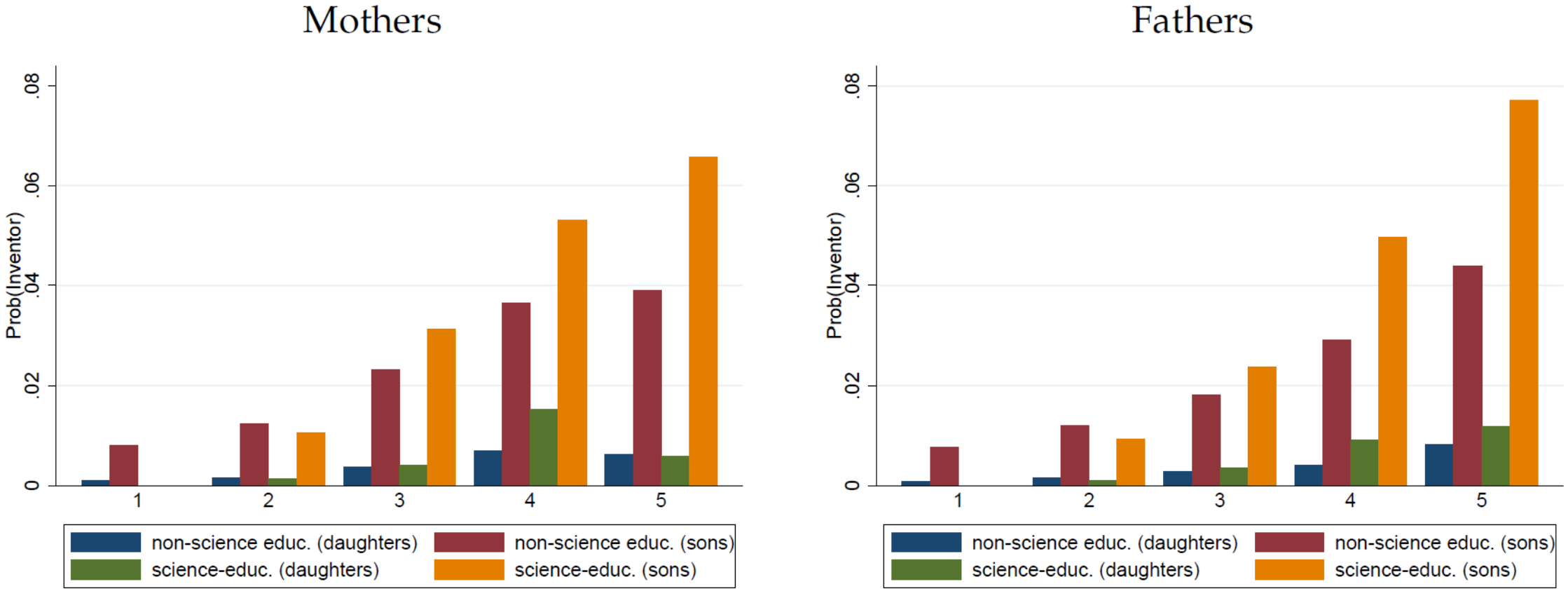


2B. Paternal income & Prob(MSc_{father})



Parental Education and Becoming Inventor

FIGURE 3: PROB(*Inventor*) AND PARENTAL EDUCATION



Notes: 1 = base education; 2 = secondary education; 3 = BSc; 4 = MSc; 5 = PhD. Non-science and science refer to the field of education of the parent.

Table 1: Estimation results

	Panel A. All Children			
	(1)	(2)	(3)	(4)
	OLS	IV	IV	IV
$D(MScparents)$	0.0159***	0.0506***	0.0328***	0.0327***
	(0.00132)	(0.0110)	(0.009)	(0.0049)
F	-	55.73	140.73	508.87
Nobs	1 450 789			
Dep. var. mean	0.0067			
	Panel B. Daughters			
$D(MScparents)$	0.0049***	0.0100	0.0203**	0.0160***
	(0.0005)	(0.0085)	(0.0086)	(0.0034)
F	-	47.25	75.87	326.80
Nobs	709 117			
Dep. var. mean	0.0016			
	Panel C. Sons			
$D(MScparents)$	0.0261***	0.0866***	0.0430**	0.0487***
	(0.0023)	(0.0193)	(0.0205)	(0.0092)
F	-	35.95	94.12	264.76
Nobs	741 671			
Dep. var. mean	0.0118			
	Instruments			
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

Table 1: Estimation results

	(1)	(2)	(3)	(4)
	OLS	IV	IV	IV
	Panel A. All Children			
$D(MScparents)$	0.0159***	0.0506***	0.0328***	0.0327***
	(0.00132)	(0.0110)	(0.009)	(0.0049)
F	-	55.73	140.73	508.87
Nobs	1 450 789			
Dep. var. mean	0.0067			
	Panel B. Daughters			
$D(MScparents)$	0.0049***	0.0100	0.0203**	0.0160***
	(0.0005)	(0.0085)	(0.0086)	(0.0034)
F	-	47.25	75.87	326.80
Nobs	709 117			
Dep. var. mean	0.0016			
	Panel C. Sons			
$D(MScparents)$	0.0261***	0.0866***	0.0430**	0.0487***
	(0.0023)	(0.0193)	(0.0205)	(0.0092)
F	-	35.95	94.12	264.76
Nobs	741 671			
Dep. var. mean	0.0118			
	Instruments			
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

Sample mean = 0.006. This implies MSc parent is 2.7 times more likely to have an inventor child.

Table 1: Estimation results

	(1)	(2)	(3)	(4)
	OLS	IV	IV	IV
Panel A. All Children				
$D(MScparents)$	0.0159***	0.0506***	0.0328***	0.0327***
	(0.00132)	(0.0110)	(0.009)	(0.0049)
F	-	[illegible]	140.73	508.87
Nobs	1 450 789			
Dep. var. mean	0.0067			
Panel B. Daughters				
$D(MScparents)$	0.0049***	0.0100	0.0203**	0.0160***
	(0.0005)	(0.0085)	(0.0086)	(0.0034)
F	-	47.25	75.87	326.80
Nobs	709 117			
Dep. var. mean	0.0016			
Panel C. Sons				
$D(MScparents)$	0.0261***	0.0866***	0.0430**	0.0487***
	(0.0023)	(0.0193)	(0.0205)	(0.0092)
F	-	35.95	94.12	264.76
Nobs	741 671			
Dep. var. mean	0.0118			
Instruments				
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

IV: (log of) distance to the nearest university from the birth municipality of mother.

Table 1: Estimation results

	(1)	(2)	(3)	(4)
	Panel A. All Children			
	OLS	IV	IV	IV
$D(MScparents)$	0.0159***	0.0506***	0.0328***	0.0327***
	(0.00132)	(0.0110)	(0.009)	(0.0049)
F	-	55.73	140.73	508.87
Nobs	1 450 789			
Dep. var. mean	0.0067			
	Panel B. Daughters			
$D(MScparents)$	0.0049***	0.0100	0.0203**	0.0160***
	(0.0005)	(0.0085)	(0.0086)	(0.0034)
F	-	47.25	75.87	326.80
Nobs	709 117			
Dep. var. mean	0.0016			
	Panel C. Sons			
$D(MScparents)$	0.0261***	0.0866***	0.0430**	0.0487***
	(0.0023)	(0.0193)	(0.0205)	(0.0092)
F	-	35.95	94.12	264.76
Nobs	741 671			
Dep. var. mean	0.0118			
	Instruments			
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

For daughters, the impacts are 6 - 13 folds (sample mean = 0.0016).

Table 1: Estimation results

	(1)	(2)	(3)	(4)
	Panel A. All Children			
	OLS	IV	IV	IV
$D(MScparents)$	0.0159***	0.0506***	0.0328***	0.0327***
	(0.00132)	(0.0110)	(0.009)	(0.0049)
F	-	55.73	140.73	508.87
Nobs	1 450 789			
Dep. var. mean	0.0067			
	Panel B. Daughters			
$D(MScparents)$	0.0049***	0.0100	0.0203**	0.0160***
	(0.0005)	(0.0085)	(0.0086)	(0.0034)
F	-	47.25	75.87	326.80
Nobs	709 117			
Dep. var. mean	0.0016			
	Panel C. Sons			
$D(MScparents)$	0.0261***	0.0866***	0.0430**	0.0487***
	(0.0023)	(0.0193)	(0.0205)	(0.0092)
F	-	35.95	94.12	264.76
Nobs	741 671			
Dep. var. mean	0.0118			
	Instruments			
Maternal dist.	NO	YES	NO	YES
Paternal dist	NO	NO	YES	YES

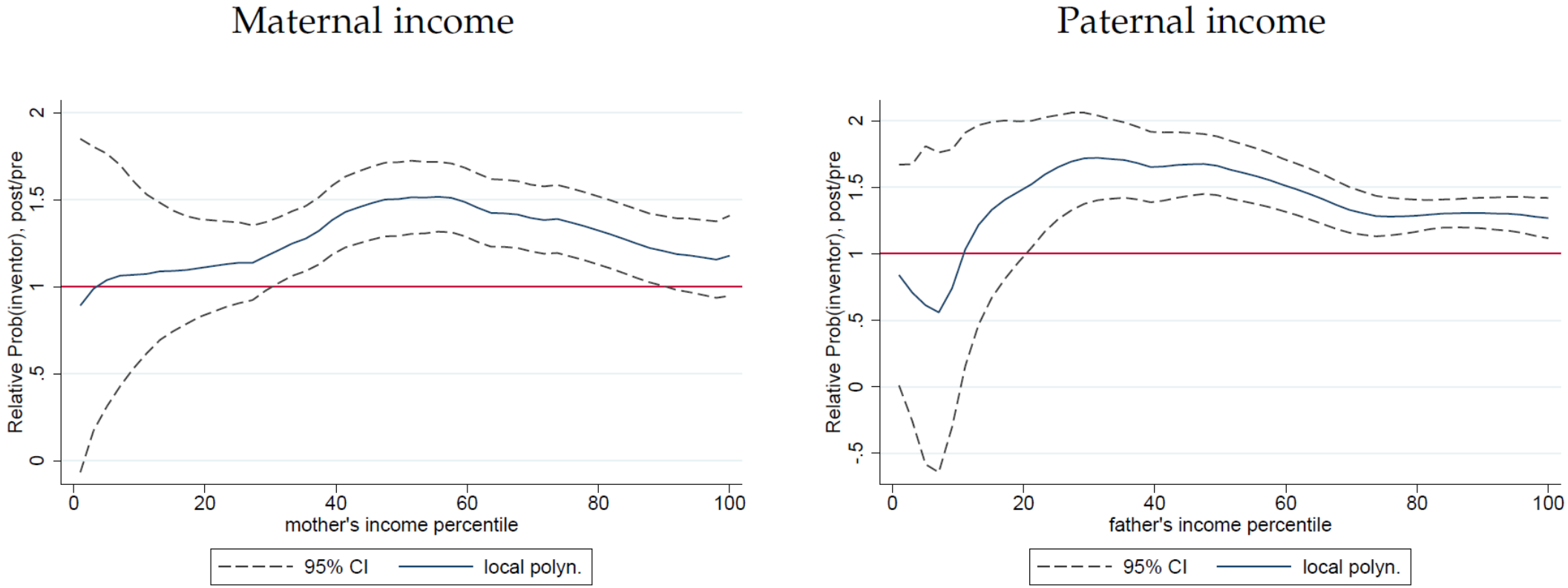
For sons, the impacts are 4 - 9 folds (sample mean = 0.0118).

1972 Schooling Reform

- Before 1968, students were sorted into academic grammar schools and work-oriented civic schools after their four years of primary education.
- In 1972 (transition until 1977), a new legislation merged these two tracks into a 9-year comprehensive school including 6 years of basic education and 3 years of lower secondary education.
- During this time, teacher education was reformed radically to equip teachers with competences needed to take greater responsibilities, such as curriculum development and student assessment.
- First affected cohorts were born in 1961-1966 (taking transition into account.)

1972 Schooling Reform

FIGURE 7: RELATIVE $P(Inventor)$ 1966-1970 COMPARED TO 1956-1960



Notes: On the x-axis is the maternal (left hand figure) or paternal (right hand figure) income percentile. On the y-axis is the parental income percentile-specific ratio of $P(Inventor)$ calculated for the 1966-1970 cohorts and divided by $P(Inventor)$ calculated for the 1966-1960 cohorts. The curve shows a local polynomial estimate and the shaded are the 95% confidence interval. A ratio of one indicates no change in $P(Inventor)$, a ratio above one an increase and a ratio of less than one a decrease from 1956-1960 to 1966-1970.

Conclusion

- Education can increase the innovative capacity of a society.
- PhD education has the strongest impact on innovation.
- Parental resources seem to play a big role for innovation through access to education.
  - Less severe when credit is available.
  - More severe when inequality is higher.
- Talent is crucial for innovation. Educating the right talents must be a priority.
- Pecking order of innovation policies: Education subsidy > R&D subsidy.
- Educated parents (causal) raise innovative kids.
- Finnish comprehensive schooling reform is associated with higher probability of becoming an inventor.