ECON 326: Economics of Developing Countries
TA Session 6

Vaidehi Parameswaran (Northwestern Econ)

February 2025

Today's Agenda

- Qian (2008)
- Midterm Exam

Qian (2008)

Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance

Overview

- Estimates the effect of sex-specific income on sex-differential survival of children, using exogenous variation in the former
- Preview of findings:
  - Increasing female income improves survival rates for girls and educational attainment for all children
  - Increasing male income worsens survival rates for girls and decreases educational attainment for girls without increasing boys' educational attainment

What's the endogeneity issue?

What's the threat to identification in the cross-section?

What's the endogeneity issue?

What's the threat to identification in the cross-section?

- Omitted variable issue: other reasons may cause women's status to be higher in certain areas, which would affect both women's income share and girls' outcomes
- Calls for an instrumental variable approach!

Context

- Policy change: Post-Mao liberalization of agriculture under the "household responsibility system" led to an increase in the production of cash crops (tea, orchards, vegetables) relative to cereals after 1979
- Two main reforms:
  - Devolved responsibilities to households
  - Increased procurement prices
- Differential impact on sex-specific income: Tea is a crop where women have a comparative advantage, whereas orchard is a crop where men have a comparative advantage

Variation over time induced by policy change

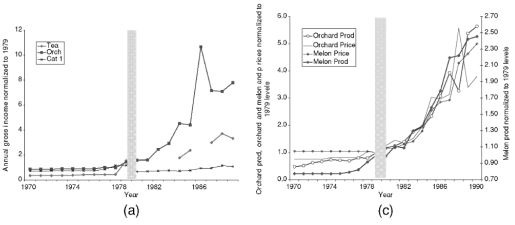


FIGURE II

(a) Gross agricultural incomes from producing tea and Category 1 crops. *Source*. FAO and Ministry of Agriculture of China. *Note*. The missing data points reflect years for when labor output data are missing. (b) Category 1 production: grains. *Source*. FAO. (c) Category 2 production: orchard and melon production and procurement prices. *Source*. FAO. (d) Tea yield and tea procurement price. *Source*. FAO.

Diff-in-diff Plot

- Suppose parents respond to expected returns of having boys and girls
- What would you expect to see if you plotted the gender ratio over time in regions suitable to tea production vs. other regions?

Diff-in-diff Plot

- Suppose parents respond to expected returns of having boys and girls
- What would you expect to see if you plotted the gender ratio over time in regions suitable to tea production vs. other regions?

Diff-in-diff Strategy

- Compare sex ratio:
  - Before and after the reform: Comparing within counties across cohorts differences out time-invariant community characteristics
  - Tea producing and non tea producing regions: Comparing within cohorts between tea-planting and non-tea-planting communities differences out changes over time that affect these regions similarly
- Then difference the differences
- Estimating equation:

$$s_{ik} = \alpha + \beta POST + \delta TEA + \gamma(POST \times TEA) + \epsilon_{ik}$$

, where γ is the difference-in-difference estimate of the effect of growing tea on sex ratios

Identification

- Parallel trends assumption
  - No change in sex preferences in tea counties during the same time as the reform
  - No other changes in tea counties during the same time of the reform
- The decision to plant tea not affected by sex preference. Why does she need this assumption?

Basic diff-in-diff results

TABLE III

OLS AND 2SLS ESTIMATES OF THE EFFECT OF PLANTING TEA AND ORCHARDS ON SEX RATIOS CONTROLLING FOR COUNTY LEVEL LINEAR COHORT TRENDS

	Dependent variables					
	Fraction of males			Tea × post	Fraction of males	
	(1) OLS	(2) OLS	(3) OLS	(4) 1st	(5) IV	(6) IV
Tea × post	−0.012	−0.013	−0.012		−0.072	−0.011
	(0.007)	(0.006)	(0.005)		(0.031)	(0.007)
Orchard × post	0.005					
	(0.002)					
Slope × post	−0.002			0.26		
	(0.002)			(0.057)		
Linear trend	No	No	Yes	Yes	No	Yes
Observations	28,349	37,756	37,756	37,756	37,756	37,756

Notes. Coefficients of the interactions between dummies indicating whether a cohort was born post-reform and the amount of tea planted in the county of birth. All regressions include county and birth year fixed effects and controls for Han, and cashcrop × post. All standard errors are clustered at the county level. In column (1), the sample includes all individuals born during 1970–1986. In columns (2)–(6), the sample includes all individuals born during 1962–1990. Post = 1 if birthyear > 1979. Data for land area sown are from the 1997 China Agricultural Census.

Generalise in two steps

- Step 1: Include one dummy per cohort, and one dummy per county

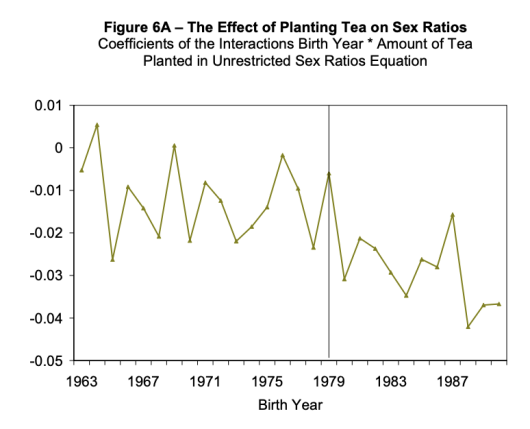
$$s_{ik} = \alpha + \sum_t \beta_t D_t + \sum_j \delta_j D_j + \gamma(POST_k \times TEA_i) + \epsilon_{ik}$$

  where $D_k = 1$ if $t = k$ and 0 otherwise, $D_j = 1$ if $i = j$ and 0 otherwise

- Step 2: Include an interaction per cohort

$$s_{ik} = \alpha + \sum_t \beta_t D_t + \sum_j \delta_j D_j + \sum_t \gamma_t (D_t \times TEA_i) + \epsilon_{ik}$$

Plot the γ_k



Introducing orchards as control group

- What if the effect is not about tea, but about cash crops in general?
- What would you expect to see if you plotted the gender ratio over time in regions suitable to orchard production and in other regions? If you repeated the diff-in-diff analysis but with orchards?

$$s_{ik} = \alpha + \sum_t \beta_t D_t + \sum_j \delta_j D_j + \lambda_t (D_t \times ORCHARD_i) + \epsilon_{ik}$$

Orchards

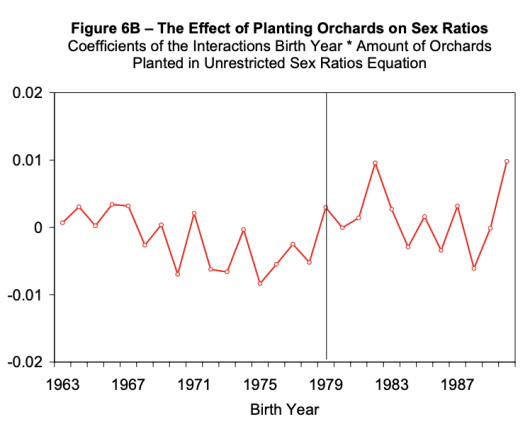
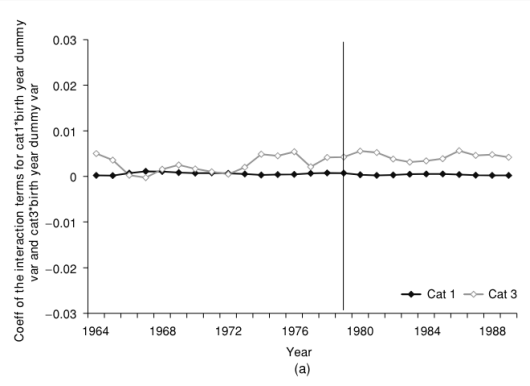


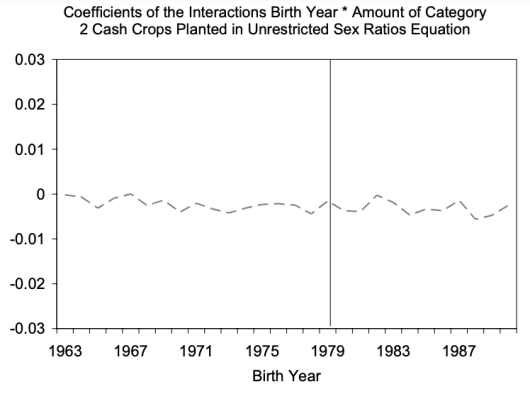
Figure 6B – The Effect of Planting Orchards on Sex Ratios
Coefficients of the Interactions Birth Year * Amount of Orchards Planted in Unrestricted Sex Ratios Equation
0.02
0.01
0
-0.01
-0.02
1963
1967
1971
1975
1979
1983
1987
Birth Year

Placebo test



- Category 1: grains, that did not become unregulated
- Category 3: vegetable patches that had never been regulated

Test of income effect



- All cash crops taken together have no effect

Introducing orchards and tea together in one regression

- For precise estimation, in the paper all the data are pooled together in one big regression with three sets of dummies

$$s_{ik} = \alpha + \beta POST_k + \sum_j \delta_j D_j + \gamma_t (D_t \times TEA_i) + \sum_t \lambda_t (D_t \times ORCHARD_i) + \epsilon_{ik}$$

Yearly coefficients

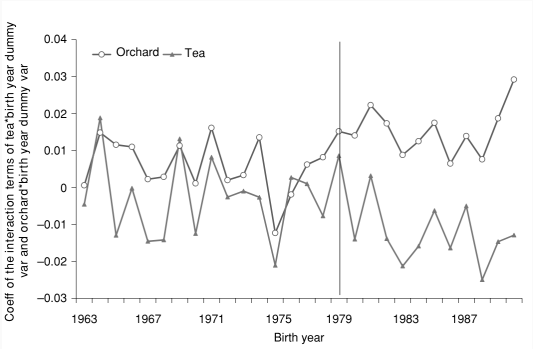


FIGURE V

The Effect of Planting Tea and Orchards on Sex Ratios

Coefficients of the interactions of birth year × amount of tea planted and birth year × amount of orchards planted controlling for year and county of birth FEs.

Differential effect on girls' education

Same regression but with education as the dependent variable

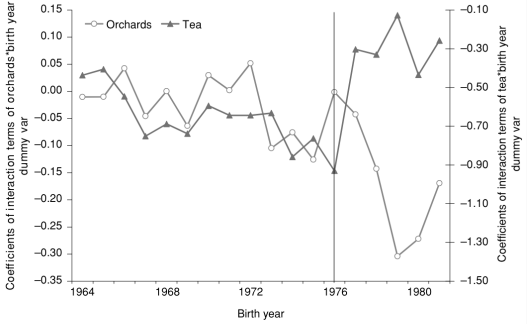


FIGURE VI

The Effect of Planting Tea and Orchards on Girls' Educational Attainment

Coefficients of the interactions birth year × amount of tea planted and birth year × amount of orchards planted controlling for year and county of birth FEs.

Dicussion and interpretation

- What if people who like girls grow tea? What is a potential strategy to deal with it?

Dicussion and interpretation

- What if people who like girls grow tea? What is a potential strategy to deal with it?
- Instrument for tea production using suitability of land for tea production (hilliness / slope of land)

Dicussion and interpretation

- What if people who like girls grow tea? What is a potential strategy to deal with it?
- Instrument for tea production using suitability of land for tea production (hilliness / slope of land)
- Common use case of instruments: providing exogenous variation for a variable which we worry is endogenous

TABLE III

OLS AND 2SLS ESTIMATES OF THE EFFECT OF PLANTING TEA AND ORCHARDS ON SEX RATIOS CONTROLLING FOR COUNTY LEVEL LINEAR COHORT TRENDS

	Dependent variables					
	Fraction of males			Tea × post	Fraction of males	
	(1) OLS	(2) OLS	(3) OLS	(4) 1st	(5) IV	(6) IV
Tea × post	−0.012	−0.013	−0.012		−0.072	−0.011
	(0.007)	(0.006)	(0.005)		(0.031)	(0.007)
Orchard × post	0.005					
	(0.002)					
Slope × post	−0.002			0.26		
	(0.002)			(0.057)		
Linear trend	No	No	Yes	Yes	No	Yes
Observations	28,349	37,756	37,756	37,756	37,756	37,756

Notes. Coefficients of the interactions between dummies indicating whether a cohort was born post-reform and the amount of tea planted in the county of birth. All regressions include county and birth year fixed effects and controls for Han, and cashcrop × post. All standard errors are clustered at the county level. In column (1), the sample includes all individuals born during 1970–1986. In columns (2)–(6), the sample includes all

Mechanisms

- Are we sure it is about higher returns to girls?

Mechanisms

- Are we sure it is about higher returns to girls?
- What else does an increase in womens' relative wages cause, immediately?

Midterm Exam

See you next time!