Early Life Exposure to Tap Water and the Development of Cognitive Skills

Chen Jie Yvonne[a] Li Li[b] Xiao Yun[c]

[a]Lee Kuan Yew School of Public Policy, National University of Singapore

[b]China Eastern Normal University

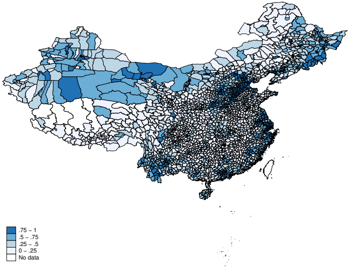
[c]Tinbergen Institute, the Netherlands

June 2, 2018

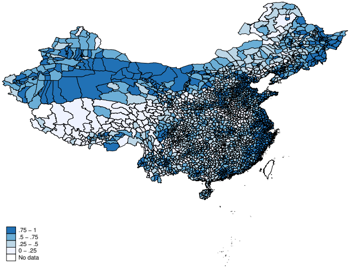
Tap water availability in China

- Water accessibility and treatment are known to have significant impacts on public health (Jalan and Ravallion, 2003; Galiani et al., 2005; Mangyo, 2008; Gamper-Rabindran et al., 2010; Zhang, 2012; Kosec, 2014).
- However, many countries around the world have low access to on-premise treated water (tap water) including China.

1) Tap water coverage rate from 2000 Census county assembly



2) Tap water coverage rate from 2010 Census county assembly



Preview

- Objective of this study
  - To examine the causal effect of early life access to tap water on children's cognitive development in rural China
  - We explore the variation in tap water exposure induced by the Rural Drinking Water Program in China
- Summary of findings
  - One additional year of tap water exposure during early life raises cognitive test score at ages 10-15 by 0.132 standard deviations.
    - A child receiving full exposure in early life would correctly answer four more questions in math test (24) than a counterpart without any exposure.
  - Early life proves to be the crucial time period for cognitive development.
  - Possible channels: tap water exposure during early life
    - Improves childhood health
    - Increases mothers' time for family care

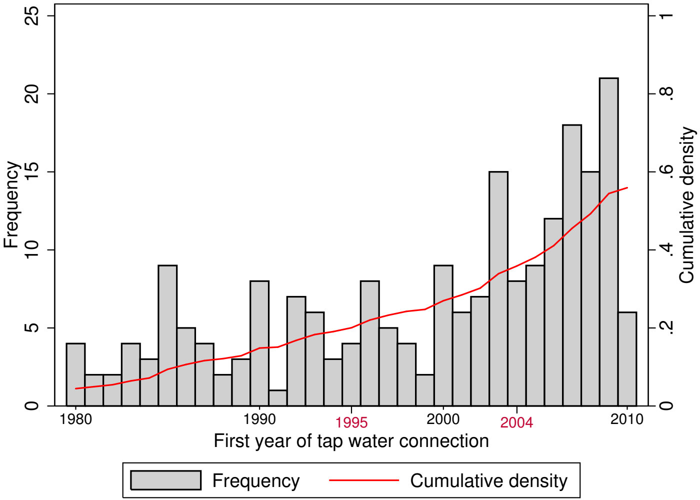
Contributions to literature

- The *determinants of cognitive skills* over the life-cycle
  - Family income (Dahl and Lochner, 2012), schooling (Carlsson et al., 2015)
  - Interventions that target cognitive development (Gertler et al., 2014; Bierman et al., 2017)
- The importance of *early life environment* on later human capital development
  - Early life interventions (Bleakley, 2007; Maluccio et al., 2009; Barham, 2012; Bharadwaj et al., 2013; Spears and Lamba, 2016)
  - Early childhood events, such as drought (Hoddinott and Kinsey, 2001), disease exposure (Cutler et al., 2010) and income shocks (Adhvaryu et al., 2017)
- The long-lasting effects of *improved access to treated water* on human capital development
  - Contemporaneous impact on children's health and educational outcomes (Jalan and Ravallion, 2003; Galiani et al., 2005; Gamper-Rabindran et al., 2010; Zhang, 2012; Koolwal and van de Walle, 2013; Kosec, 2014; Zhang and Xu, 2016)
- The microeconomic effects of *public infrastructure* construction and expansion on human capital development in developing countries
  - Such as access to schools (Duflo, 2001, 2004), electricity (Dinkelman, 2011), health facilities (Gruber et al., 2014), and roads (Banerjee et al., 2012)

Data

- Rural Drinking Water Program in China (1984-)
  - Tap water coverage rate in rural area increased from 11% in 1990 to 55% in 2015 (UNICEF and WHO, 2015).
- Data source: China Family Panel Studies (CFPS 2010)
- Sample: 2,168 children, aged 10-15 in 2010, from 404 rural communities in 25 provinces in China
- Cognitive skills: standardized word recognition test score, math test score, and the average
- Key variable: years of exposure to tap water during early life
  - Early life: from one year before birth to the first five years of life (spanning 6 years)
  - Constructed using year of birth and first year of tap water connection at community level

Program phase-in



Tap water penetration Probability and timing

Summary statistics

	Ratio of missing values (1)	Mean (2)	s.d. (3)
Panel A: Individual characteristics (2,168 children)			
- Outcomes			
Average cognitive test score	0	18.142	6.210
Word score	0	20.913	7.334
Math score	0	10.850	4.514
- Treatment variable			
Exposure to tap IU-5	0	1.326	2.380
- Control variables			
Boy	0	0.498	0.500
Mother's age at birth	0.009	25.702	4.511
Number of siblings	0.044	1.250	0.956
Birth order	0.044	1.761	0.906
Number of household members	0	5.100	1.633
Father's years of schooling	0.017	5.915	3.934
Mother's years of schooling	0.020	4.003	3.932
Exposure to electricity IU-5	0.036	5.362	1.735
Exposure to cable/satellite TV IU-5	0.037	1.688	2.407
Exposure to roads IU-5	0.022	3.451	2.831
Exposure to health facilities IU-5	0.011	2.705	2.930
Exposure to landline phones IU-5	0.036	3.917	2.552

Notes:The main sample for analysis consists of 2,168 children. We account for missing data by imputing the mean value for the sample, and include a dummy variable for imputed responses in the regressions. Column (1) reports the ratio of missing values for each variable.

Summary statistics - cont'd

	Mean (1)	s.d. (2)
Panel B: Community characteristics (382 rural communities)		
Suburban village (0/1)	0.217	0.413
Distance to nearest town or city (hours)	27.882	22.445
Population	2,063.757	1,616.015
Area	51.582	412.411
Average years of schooling of 25-55 year-old	5.068	1.846
Net income per capita (yuan)	3,612.587	3,096.639
Hills	0.317	0.466
Mountains	0.157	0.364
Plateaus	0.045	0.206
Plains	0.325	0.469
Others	0.157	0.364
Having tap water connection	0.550	0.498
Having electricity	0.914	0.281
Having cable/satellite TV	0.707	0.456
Having road connection	0.856	0.352
Having health facilities	0.584	0.494
Having landline phone	0.872	0.335

Notes: The main sample for analysis consists of 2,168 children. We account for missing data by imputing the mean value for the sample, and include a dummy variable for imputed responses in the regressions. Column (1) reports the ratio of missing values for each variable.

Empirical strategy

- We use the difference-in-differences (DID) approach with community and birth year fixed effects to exploit the variation in early life exposure across cohorts within a given community.
- Model:

$$Y_{ict} = \phi \text{Exposure to tap IU-5}_{ct} + X_{ict}\beta + g_c + \gamma_t + \delta_{pt} + \theta_t W_c + \epsilon_{ict}, \qquad (1)$$

  - Y_{ict}, word recognition test score, math test score, or the average
  - $\text{Exposure to tap IU-5}_{ct}$, years of tap water exposure during early life
  - X_{ict}, individual and household characteristics
  - g_c, community fixed effects
  - γ_t, birth year fixed effects
  - δ_{pt}, province-year fixed effects
  - $\theta_t W_c$, interactions between birth year dummies and community characteristics
  - ϵ_{ict}, random error term

Main results

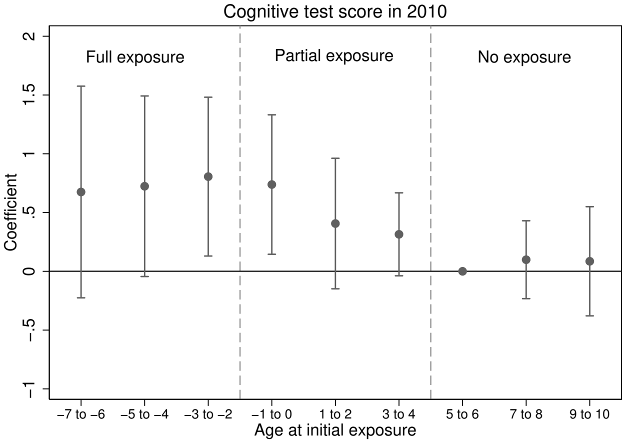
VARIABLES	Average cognitive test score (1)	(2)	(3)	Word (4)	Math (5)
Exposure to tap IU-5	0.108***	0.126***	0.132***	0.073	0.162***
	(0.034)	(0.045)	(0.046)	(0.049)	(0.050)
Observations	2,168	2,168	2,168	2,168	2,168
R^2	0.409	0.447	0.448	0.444	0.396
Number of clusters	382	382	382	382	382
Cohort FE	Y	Y	Y	Y	Y
Community FE	Y	Y	Y	Y	Y
Cohort-province FE, community controls × cohort	N	Y	Y	Y	Y
Early life exposure to other public facilities	N	N	Y	Y	Y

Notes: * means significant at 10%, ** significant at 5%, and *** significant at 1%.
Individual and household characteristics controlled for in all columns are gender, mother's age at birth, number of siblings, birth order, household size, father's years of schooling, and mother's years of schooling. In column (2), the community characteristics used to construct interactions with birth year dummies are suburban village (=1 if yes, =0 if not), distance to nearest town (city), log(population), log(area), average years of schooling of 25-55 year-old, and topographic characteristics (hills, mountains, plateaus, plains, and others). Columns (3)-(5) also include early life exposure to other public facilities, including roads, television, electricity, health facilities, and landline phones. Standard errors in parentheses are clustered at the community level.

Original

Event study

Event time: difference between the first year of connection and the year of birth



Notes: The figure plots coefficients (dot) and 95% confidence intervals (vertical line) from an event study analysis. Event time is defined as the difference between the first year of tap water connection at the community level and the year of birth. For example, a child born in 2000 in a community that first got tap water connection in 2002 would have an event time of 2. The child would have an event time of -2 if tap water was first connected to the birth community in 1998. A negative event time indicates full exposure in early life. The event time of 5-6 is taken as the base group so estimates for other age groups are relative to that point.

Heterogeneous effects

	Dep. var.: Average cognitive test score		
	(1)	(2)	(3)
Panel A. Socioeconomic status			
	Gender	Mother's schooling	Household income
Exposure to tap IU-5	0.129***	0.131***	0.124***
	(0.047)	(0.049)	(0.048)
Exposure to tap IU-5 × boy	0.005		
	(0.020)		
Exposure to tap IU-5 × low-educated mother		0.010	
		(0.019)	
Exposure to tap IU-5 × low-income household			0.003
			(0.015)
Observations	2,168	2,125	2,069
R^2	0.448	0.452	0.454
Panel B. Status of baseline water sources			
	Water source	Water amount	Water quality[a]
Exposure to tap IU-5	0.103**	0.097	0.115*
	(0.047)	(0.070)	(0.061)
Exposure to tap IU-5 × not using underground water	0.044		
	(0.326)		
Exposure to tap IU-5 × rainfall above median		0.058	
		(0.094)	
Exposure to tap IU-5 × best water quality			0.016
			(0.147)
Exposure to tap IU-5 × worst water quality			0.095
			(0.160)
Observations	1,798	2,164	1,540
R^2	0.458	0.448	0.483

Notes: [a] Data on water grades are obtained from 484 monitoring sites. Readings are averaged across water basins then matched to communities in our sample. Six water grades are categorized into three types: best (grade I-II), medium (grade III-IV), and worst (grade V-VI). The type of medium is taken as the base group. Each coefficient is estimated from a separate regression with the same specification applied as in column (3) of the Main results table. Standard errors in parentheses are clustered at the community level.

Possible mechanisms

- Health
  - Better access to clean water may improve childhood health by reducing the incidence of waterborne disease and improving nutritional status (Jalan and Ravallion, 2003; Zhang, 2012).
- Time use
  - The burden of water collection in developing countries is borne primarily by women and girls (Devoto et al., 2012).
  - If access to tap water increases maternal time investment, it might indirectly improve children's short- and long-term development (Ruhm, 2008; Miller and Urdinola, 2010).
- Educational attainment
  - Improvement in one's educational attainment may function as one of the channels via which cognitive test scores are increased (Carlsson et al., 2015).
- Household Income and education expenditure

Possible mechanism I

Health: 0-5 year-old children in 2010

	Last month		Last year		First 12 months after birth	
	# illness	# doctor visits	# doctor visits	Top 20% expenditure	# illness	# doctor visits
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Tap water in community	-0.378***	-0.217**	-1.005**	-0.068*	-1.397**	-1.328**
	(0.114)	(0.098)	(0.484)	(0.040)	(0.615)	(0.519)
Observations	2,036	2,049	1,952	1,996	1,641	1,635
R^2	0.425	0.435	0.429	0.436	0.423	0.449

Notes: * means significant at 10%, ** significant at 5%, and *** significant at 1%.
The sampled children are from communities that either received tap water before 2005 or after 2010 and were aged 0-5 in 2010. Tap water in community is equal to one if the community first got tap water connection before 2005, and zero if otherwise. Each coefficient is estimated from a separate regression. Individual-level control variables include the child's gender, age, age squared, number of siblings, mother's age at birth, parents' years of schooling, number of household members, an indicator of farm household, and log (family income). Community-level controls include an indicator of suburban village, log (distance to nearest town or city), log (population), log (area), average years of schooling among the 25-55 years old residents, topographic characteristics of the community, % of labor force working outside of community, log (annual income per capita), and community accessibility to other facilities (kindergarten, primary school, health facility, electricity, road, cable/satellite TV, landline phone, and mobile signal). County-cohort fixed-effects model is applied. Standard errors in parentheses are clustered at the community level.

Possible mechanism I

Health: children aged 6-15 in 2010

VARIABLES	Height-for-age z-score (1)	Stunting (2)
Exposure to tap IU-5	0.056	-0.027*
	(0.061)	(0.014)
Observations	3,253	3.253
R^2	0.424	0.383

Notes: * means significant at 10%, ** significant at 5%, and *** significant at 1%.
Each coefficient is estimated from a separate regression with the same specification applied as in column (3) of main results. Standard errors in parentheses are clustered at the community level.

Possible mechanism II

Time use: women aged 20-45 in 2010, hours per week

	Have 0-5 year-old child			Have no 0-5 year-old child		
	Family care	Work	Housework	Family care	Work	Housework
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
Tap water in community	6.240**	-2.878	0.713	-0.101	3.975**	1.480**
	(2.659)	(2.818)	(0.870)	(0.510)	(2.015)	(0.693)
Observations	1,233	1,238	1,232	3,127	3,138	3,124
R^2	0.269	0.293	0.282	0.199	0.215	0.251

Notes: * means significant at 10%, ** significant at 5%, and *** significant at 1%.

The sample women were 20-45 years old in 2010. Each coefficient is estimated from a separate regression. Individual-level control variables include age of the respondent and its quadratic form, years of schooling, an indicator of living with the spouse, number of household members, and an indicator of farm household. Community-level controls include an indicator of suburban village, log (distance to nearest town or city), log (population), log (area), average years of schooling among the 25-55 years old residents, topographic characteristics of the community, % of labor force working outside of community, log (annual income per capita), and community accessibility to other facilities (kindergarten, primary school, health facility, electricity, road, cable/satellite TV, landline phone, and mobile signal). County fixed-effects model is applied. Standard errors in parentheses are clustered at the community level.

Possible mechanism III

Educational attainment: 10-15 year-old children in 2010

VARIABLES	Years of schooling (1)	Grade-for-age (2)	School enrollment (3)
Exposure to tap IU-5	-0.016	0.003	0.001
	(0.054)	(0.028)	(0.010)
Observations	2,184	2,246	2,246
R^2	0.807	0.432	0.356

Notes: * means significant at 10%, ** significant at 5%, and *** significant at 1%.
Each coefficient is estimated from a separate regression with the same specification applied as in column (3) of main results. Grade-for-age status is measured by a dummy indicator that takes value one if the child was enrolled in the supposed grade for his/her age and zero otherwise. School enrollment status is measured by a dummy indicator which equals to one if the child was enrolled in school at the survey time and zero otherwise. Standard errors in parentheses are clustered at the community level.

Possible mechanism IV

Household income and education expenditure

	Logarithm of			
	Family income	Educational exp on the child	Household educational exp	Educational exp on members other than the child
VARIABLES	(1)	(2)	(3)	(4)
Exposure to tap IU-5	0.003	0.03	0.009	0.006
	-0.019	-0.056	-0.054	-0.08
Observations	3,380	3,511	2,685	2,685
R2	0.494	0.517	0.568	0.458
Mean of dep. var. before log	27810	804.2	2659	1998

Notes: Sample children were aged 6-15 in 2010. In columns (3) and (4), children with household expenditure less than individual expenditure are dropped. All regressions have cohort fixed-effects, community fixed-effects, cohort-province fixed-effects, the interactions of community controls and cohort fixed-effects, and early life exposure to other public facilities controlled for. The community characteristics used to construct interactions with cohort dummies are suburban village (=1 if yes, =0 if not), distance to nearest town (city), log(population), log(area), average years of schooling of 25-55-year-old, and topographic characteristics (hills, mountains, plateaus, plains, and others). Other public facilities including roads, television, electricity, health facilities, and landline phones. Standard errors in parentheses are clustered at the community level.

Implications

- Our study shows the essential role of early life exposure to tap water in the development of cognitive skills.
- It reveals the role of public infrastructure in human capital development and the importance of intervention timing.
    - More work on the relationship between infrastructure construction and human capital development would provide guidance for distributing fiscal resources across regions.
- Improving access to tap water in rural area may have the great potential to narrow down the large gap in human capital between urban and rural children in China.
    - Gap calculated using CFPS 2010 data: 0.5 standard deviations
    - Our estimated full-exposure effect: 0.8 standard deviations

Thank you!

Reference I

Adhvaryu, A., Fenske, J., and Nyshadham, A. (2017). Early life circumstance and adult mental health. *Working Paper*.

Attanasio, O. P. (2015). The determinants of human capital formation during the early years of life: Theory, measurement, and policies. *Journal of the European Economic Association*, 13(6):949–997.

Banerjee, A., Duflo, E., and Qian, N. (2012). On the road: Access to transportation infrastructure and economic growth in China. NBER Working Paper No. 17897.

Barham, T. (2012). Enhancing cognitive functioning: Medium-term effects of a health and family planning program in Matlab. *American Economic Journal: Applied Economics*, 4(1):245–273.

Bharadwaj, P., Løken, K. V., and Neilson, C. (2013). Early life health interventions and academic achievement. *American Economic Review*, 103(5):1862–91.

Bierman, K. L., Heinrichs, B. S., Welsh, J. A., Nix, R. L., and Gest, S. D. (2017). Enriching preschool classrooms and home visits with evidence-based programming: Sustained benefits for low-income children. *The Journal of Child Psychology and Psychiatry*, 582:129–137.

Bleakley, H. (2007). Disease and development: Evidence from hookworm eradication in the American South. *Quarterly Journal of Economics*, 122(1):73–117.

Reference II

Bloom, D. E., Canning, D., and Shenoy, E. S. (2012). The effect of vaccination on children's physical and cognitive development in the Philippines. *Applied Economics*, 44:2777–2783.

Carlsson, M., Dahl, G. B., Öckert, B., and Rooth, D.-O. (2015). The effect of schooling on cognitive skills. *Review of Economics and Statistics*, 97(3):533–547.

Choi, A. L., Zhang, Y., Sun, G., Bellinger, D. C., Wang, K., Yang, X. J., Li, J. S., Zheng, Q., Fu, Y., and Grandjean, P. (2015). Association of lifetime exposure to fluoride and cognitive functions in Chinese children: A pilot study. *Neurotoxicology and Teratology*, 47:96–101.

Cunha, F., Heckman, J. J., Lochner, L., and Masterov, D. V. (2006). Interpreting the evidence on life cycle skill formation. In Hanushek, E. A. and Welch, F., editors, *Handbook of the Economics of Education*, volume 1, pages 697–812. Elsevier.

Cunha, F., Heckman, J. J., and Schennach, S. M. (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931.

Currie, J. and Vogl, T. (2013). Early-life health and adult circumstance in developing countries. *Annual Review of Economics*, 5:1–36.

Cutler, D., Fung, W., Kremer, M., Singhal, M., and Vogl, T. (2010). Early-life malaria exposure and adult outcomes: Evidence from malaria eradication in India. *American Economic Journal: Applied Economics*, 2(2):72–94.

Reference III

Dahl, G. B. and Lochner, L. (2012). The impact of family income on child achievement: Evidence from the Earned Income Tax Credit. *American Economic Review*, 102(5):1927–1956.

Devoto, F., Duflo, E., Dupas, P., Parientó, W., and Pons, V. (2012). Happiness on tap: Piped water adoption in urban Morocco. *American Economic Journal: Economic Policy*, 4(4):68–99.

Dinkelman, T. (2011). The effects of rural electrification on employment: New evidence from South Africa. *American Economic Review*, 101:3078–3108.

Duflo, E. (2001). Schooling and labor market consequences of school construction in Indonesia: Evidence from an unusual policy experiment. *American Economic Review*, 91(4):795–813.

Duflo, E. (2004). The medium run effects of educational expansion: Evidence from a large school construction program in Indonesia. *Journal of Development Economics*, 74:163–197.

Galiani, S., Gertler, P., and Schargrodsky, E. (2005). Water for life: The impact of the privatization of water services on child mortality. *Journal of Political Economy*, 113(1):83–120.

Gamper-Rabindran, S., Khan, S., and Timmins, C. (2010). The impact of piped water provision on infant mortality in Brazil: A quantile panel data approach. *Journal of Development Economics*, 92(2):188–200.

Reference IV

Gertler, P., Heckman, J., Pinto, R., Zanolini, A., Vermeersch, C., Walker, S., Chang, S. M., and Grantham-McGregor, S. (2014). Labor market returns to an early childhood stimulation intervention in Jamaica. *Science*, 344(6187).

Glewwe, P. and King, E. M. (2001). The impact of early childhood nutritional status on cognitive development: Does the timing of malnutrition matter? *World Bank Economic Review*, 15(1):81–113.

Gruber, J., Hendren, N., and Townsend, R. M. (2014). The great equalizer: Health care access and infant mortality in Thailand. *American Economic Journal: Applied Economics*, 6:91–107.

Hanushek, E. A. and Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*, 46(3):607–668.

Heckman, J. J., Stixrud, J., and Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3):411–482.

Hoddinott, J. and Kinsey, B. (2001). Child growth in the time of drought. *Oxford Bulletin of Economics and Statistics*, 63(4):409–436.

Ilahi, N. and Grimard, F. (2000). Public infrastructure and private cost: Water supply and time allocation of women in rural pakistan. *Economic Development and Cultural Change*, 49(1):45–75.

Reference V

Jalan, J. and Ravallion, M. (2003). Does piped water reduce diarrhea for children in rural India? *Journal of Econometrics*, 112(1):153 – 173.

Koolwal, G. and van de Walle, D. (2013). Access to water, women's work, and child outcomes. *Economic Development and Cultural Change*, 61(2):369–405.

Kosec, K. (2014). The child health implications of privatizing Africa's urban water supply. *Journal of Health Economics*, 35:1–19.

Maluccio, J. A., Hoddinott, J., Behrman, J. R., Martorell, R., Quisumbing, A. R., and Stein, A. D. (2009). The impact of improving nutrition during early childhood on education among Guatemalan adults. *Economic Journal*, 119(537):734–763.

Mangyo, E. (2008). The effect of water accessibility on child health in China. *Journal of Health Economics*, 27(5):1343–1356.

Miller, G. and Urdinola, B. P. (2010). Cyclicality, mortality, and the value of time: The case of coffee price fluctuations and child survival in Colombia. *Journal of Political Economy*, 118:113–155.

Niehaus, M. D., Moore, S. R., Patrick, P. D., Derr, L. L., Lorntz, B., Lima, A. A., and Guerrant, R. L. (2002). Early childhood diarrhea is associated with diminished cognitive function 4 to 7 years later in children in a northeast brazilian shantytown. *American Journal of Tropical Medicine and Hygiene*, 66(5):590–593.

Reference VI

Ruhm, C. J. (2008). Maternal employment and adolescent development. *Labour Economics*, 15(5):958–983.

Spears, D. and Lamba, S. (2016). Effects of early-life exposure to sanitation on childhood cognitive skills: Evidence from India's Total Sanitation Campaign. *Journal of Human Resources*, 51(2):298–327.

Tyler, C. R. and Allan, A. M. (2014). The effects of arsenic exposure on neurological and cognitive dysfunction in human and rodent studies: A review. *Current Environmental Health Reports*, 1:132–147.

UNICEF and WHO (2015). Progress on sanitation and drinking water - 2015 update and MDG assessment. Technical report, UNICEF and World Health Organization, New York, United States; Geneva, Switzerland.

Zhang, J. (2012). The impact of water quality on health: Evidence from the drinking water infrastructure program in rural China. *Journal of Health Economics*, 31:122–134.

Zhang, J. and Xu, L. C. (2016). The long-run effects of treated water on education: The rural drinking water program in China. *Journal of Development Economics*, 122:1–15.

Zoni, S. and Lucchini, R. G. (2013). Manganese exposure: Cognitive, motor and behavioral effects on children: A review of recent findings. *Current Opinion in Pediatrics*, 25(2):255–60.

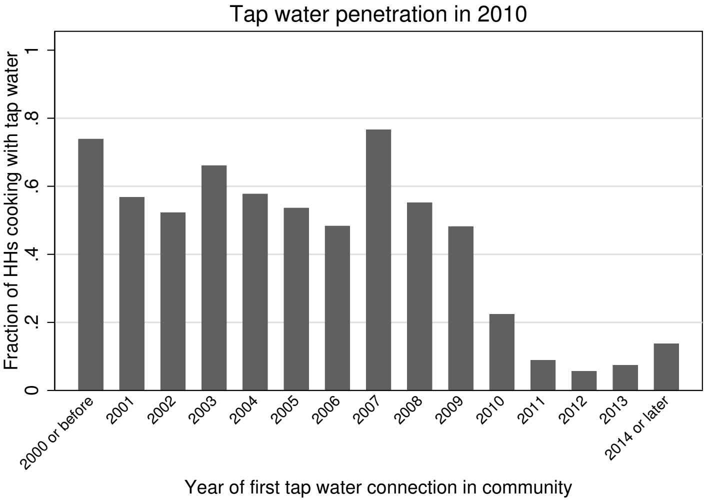
Determinants of tap water connection at the community level

VARIABLES	Connected to tap water in 2010 (1)	Year first getting connection (2)
Suburban village (0/1)	0.109*	-3.319***
	(0.064)	(1.105)
Log (distance to nearest town or city (hours))	-0.077**	1.732***
	(0.033)	(0.541)
Log (population)	-0.023	-0.909
	(0.041)	(0.632)
Log (area)	0.007	-0.018
	(0.012)	(0.197)
Average years of schooling of 25-55 year-old	0.018	-0.191
	(0.015)	(0.258)
Hills	-0.079	0.446
	(0.066)	(1.075)
Mountains	0.041	-0.478
	(0.097)	(1.460)
Plateaus	-0.054	-0.530
	(0.135)	(1.843)
Others	-0.012	-0.534
	(0.078)	(1.321)
Observations	404	404
R^2	0.198	

Notes: * means significant at 10%, ** significant at 5%, and *** significant at 1%.
We categorize the communities into five types by topographic characteristics: hills, mountains, plateaus, plains, and others, and generate dummy variables for each type. The type of plains is taken as the omitted group. Column (1) reports estimates from OLS regressions. Column (2) reports the marginal effects from Tobit regressions, conditional on community first getting tap water access no later than 2010. Robust standard errors are in parentheses.

Back

Tap water penetration



Back

Specification test - Test scores in original forms

	(1)	(2)	(3)
Panel A. Test scores in original forms			
	Average cognitive test score (original)	Word (original)	Math (original)
Exposure to tap IU-5	0.618***	0.510	0.512***
	(0.215)	(0.313)	(0.168)
Observations	2,168	2,168	2,168
R^2	0.650	0.566	0.646

Back

Conceptual Framework

- Skill formation theories suggest that cognitive skills can be produced at different stages of childhood into adulthood, taking parental background and other investments as inputs (Attanasio, 2015; Cunha et al., 2010; Heckman et al., 2006).
- Suppose at any time t, the human capital stock for each individual is a vector $H_t = (\theta_t^c, \theta_t^h)$, where θ_t^c and θ_t^h are cognitive skill and health.

Conceptual Framework- con'd

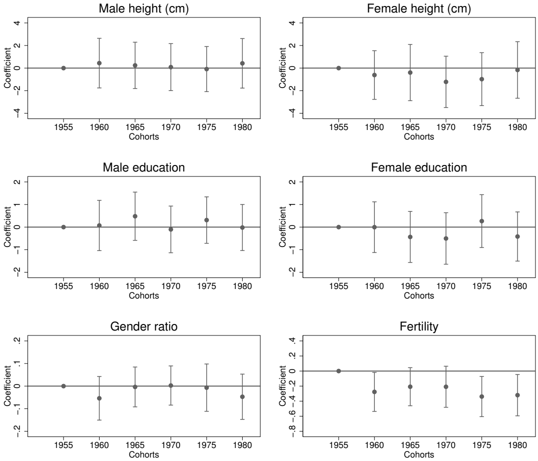
- The stock of human capital k at time $t+1$ is given by

$$\theta_{t+1}^{k} = g_t^k(H_t, Z_t, X_t, e_t^k), k \in \{c, h\}$$

- Exposure to tap water at time t can affect θ_{t+1}^{c} through the following channels
  - Z_t, such as environmental factors and household income(Zoni and Lucchini, 2013; Tyler and Allan, 2014; Choi et al., 2015; Ilahi and Grimard, 2000; Devoto et al., 2012; Koolwal and van de Walle, 2013)
  - X_t, particularly the time investment of main caregivers (Devoto et al., 2012; Ilahi and Grimard, 2000; Devoto et al., 2012; Koolwal and van de Walle, 2013).
  - Early childhood human capital θ_t^h and θ_t^c (Niehaus et al., 2002).

Validity of DID strategy I

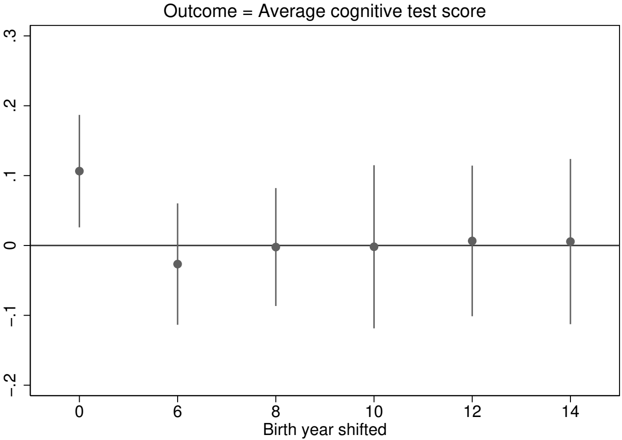
Comparing pre-trends between communities connecting earlier and later



Notes: The figure plots pre-intervention characteristics by every five-year period for males and females born between 1955 and 1984 from communities receiving tap water connection between 1995 and 2004 (earlier) and communities not receiving tap water connection before 2005 (later). Each small figure is from a separate regression. Each dot on the solid line is the estimated coefficient on the interactions between five-year-cohort dummies and a binary variable indicating whether the community receiving tap water connection earlier, and the 95% confidence interval is plotted by solid line. The cohort of 1955 is taken as the base group. Community fixed-effects model is used for estimation. Standard errors in parentheses are clustered at the community level.

Validity of DID strategy II

Time-shifted placebos: older cohorts in 2010



Notes: In the placebo test, we uniformly shift the birth year of five groups of individuals who were 16-21, 18-23, 20-25, 22-27, and 24-29 years old in 2010 forward by 6, 8, 10, 12, and 14 years, respectively, so that they were of the same "age", 10-15 years old in 2010, as the children in our main analysis. We generate pseudo Exposure to tap IU-5 for the individuals and estimate the effect for each of the five groups. Each dot on the solid line is the coefficient of interest and the 95% confidence interval is plotted by solid line. The estimates for the birth year shifted by "0" come from the baseline.

Sensitivity analysis

- Other potential confounding factors
- Reporting error
- Measurement error
- Sample selection

Sensitivity analysis- sanitation

	Dep. var.: Average cognitive test score		
	(1)	(2)	(3)
Panel A. Separately identifying the effects of tap water and sanitation [a]			
Average county tap coverage rate IU-5	0.760**	0.760**	0.336
	(0.381)	(0.381)	(0.442)
Average county toilet coverage rate IU-5		0.060	-0.030
		(0.346)	(0.383)
Average county tap coverage rate IU-5 × average county toilet coverage rate IU-5			0.657
			(0.655)
Observations	1,853	1,853	1,853
R^2	0.341	0.341	0.341
Panel B. Accounting for sanitation in the baseline specification [b]			
	+ County toilet coverage rate in 2000 × cohort	+ Change in county toilet coverage rate × cohort	+ County toilet coverage rate in 2000 × cohort + Change in county toilet coverage rate × cohort
Exposure to tap IU-5	0.097*	0.101*	0.098*
	(0.051)	(0.052)	(0.051)
Observations	1,884	1,884	1,884
R^2	0.446	0.447	0.447

Notes: [a] Data from 1% sample of China Census 2000 and 20% sample of China 1% Census 2005. [b] Sanitation data from CFPS 2010. A county fixed-effects model is applied for estimation in Panel A. We control for individual and household characteristics, province-specific cohort fixed effects, and interactions of cohort fixed effects and county characteristics. The county characteristics used include GDP per capita in 2000, ratio of urban population in 2000, and their growth rates from 2000 to 2010. In Panel B, the same specification as in column (3) of the Main results table is applied. Standard errors (in parentheses) clustered at the community level.

Sensitivity analysis - Other confounding factors

	(1)	(2)	(3)	(4)	(5)
Panel A. Other confounding factors					
A1. At community level (Dep. var.: Average cognitive test score, data from CFPS 2010)					
	Migrant workers	School supply		One child policy	
	+Prevalence of migrant workers	+Household reported distance to high school	+ Community average (own exclusive) × cohort	+ Number of births allowed × cohort	
Exposure to tap IU-5	0.133***	0.131***	0.138***	0.115**	
	(0.046)	(0.046)	(0.047)	(0.051)	
Observations	2,168	2,168	2,168	2,153	
R^2	0.449	0.449	0.450	0.452	
A2. At county level (Dep. var.: Average cognitive test score)					
	GDP[a]	Share of urban[b]	Emigration[c]	Immigration[b]	Left-behind[c]
	+ GDP pc in 2000 × cohort + change in GDP pc × cohort	+ Share of urban in 2000 × cohort + Change in share of urban × cohort	+ Emigration rate in 2000 × cohort + change in emigration rate × cohort	+ Immigration rate in 2000 × cohort + change in immigration rate × cohort	+ Left-behind rate in 2000 × cohort + Change in left-behind rate × cohort
Exposure to tap IU-5	0.145***	0.130***	0.134***	0.131***	0.102*
	(0.052)	(0.046)	(0.047)	(0.049)	(0.052)
Observations	2,133	2,164	2,098	2,164	1,946
R^2	0.452	0.450	0.453	0.450	0.456

Notes: [a] Data from China Data Online (CDO). [b] Data from county level assemblies of China Census 2000 and 2010. [c] Data from 1% sample of China Census 2000 and 20% sample of China 1% Census 2005. Only counties with more than 100 households in both samples are included. The same specification as in column (3) of the Main results table is applied for estimation. Standard errors (in parentheses) clustered at the community level.

Sensitivity analysis -Alternative measurement; measurement error;sample selection

	(1)	(2)	(3)	(4)	(5)
Panel B. Reporting error (Dep. var.: Average cognitive test score)					
	≥1994	≥1996	≥1998	Reliable answers	
Exposure to tap IU-5	0.134**	0.137***	0.147**	0.170***	
	(0.052)	(0.052)	(0.060)	(0.061)	
Observations	1,866	1,832	1,742	1,317	
R^2	0.443	0.443	0.448	0.448	
Panel C. Measurement error (Dep. var.: Average cognitive test score)					
Exposure to tap IU-5 (Mid-year connection)	0.127***				
	(0.044)				
Exposure to tap IU-5 (Year end connection)		0.110***			
		(0.040)			
Observations	2,168	2,168			
R^2	0.448	0.448			
Panel D. Sample selection (Dep. var. shown as column name)					
	Not in the final sample in 2010	Not living at home in 2010	Not surveyed in 2010	No test scores in 2010	Not surveyed in 2012
Exposure to tap IU-5	-0.002	0.006	-0.009	0.003	-0.002
	(0.020)	(0.015)	(0.015)	(0.010)	(0.007)
Observations	2,700	2,700	2,539	2,251	3,367
Observations with dep. var.=1	532	161	288	83	519
Observations with dep. var.=0	2,168	2,539	2,251	2,168	2,848
R^2	0.391	0.346	0.451	0.366	0.447

Notes:
The same specification as in column (3) of the Main results table is applied for estimation. Standard errors (in parentheses) clustered at the community level. In Panel D, each dependent variable is a dummy variable. Among the 2,700 children aged 10-15 in 2010, 161 children were not at home, 288 children were at home but did not participate in the survey, and 83 children did the child survey but did not do the cognitive test.