Online Appendix

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

Yvonne Jie Chen*, Li Li†, Yun Xiao‡

*School of Entrepreneurship and Management, ShanghaiTech University.

†Faculty of Economics and Management, East China Normal University. Corresponding author, email: lli@fem.ecnu.edu.cn.

‡Amsterdam School of Economics, University of Amsterdam.

A Probability and Timing of Tap Water Connection

The rural drinking water program is not randomly placed across regions and time. We employ a province fixed-effects model to show the correlation between community characteristics in 2010 and the probability and timing of community-level tap water connection. Factors that may have affected the probability and timing of tap water access include a community's location, population, area, level of economic development, and topographic characteristics. Population and area are proxies for the scale of a community. The level of economic development is indicated by average years of schooling. Topographic characteristics capture the difficulty of constructing the tap water infrastructure in the community.

Table A1 presents the regression results. We find that whether a community is a suburban village and the location of a community are among the most important determinants for both the placement of tap water connections and the timing of construction: a suburban village or a community closer to a town or city is more likely to have access to tap water in 2010 and would receive a tap water connection earlier than other communities located in the same province.[1] Although we do not find any significant results for population, size, average years of schooling of the 25–55 year olds, or topographic characteristics, to account for the possible difference in cohort trends across communities, in our preferred specification (column (3) of Table 2) we control for the interactions of community characteristics and cohort dummies. Considering the endogeneity of a community's population size in 2010, we do not control for the interactions of population and cohort dummies. Our results remain unchanged whether or not the interactions are controlled for.

B Household Take-Up of Tap Water

To shed some light on the program take-up rate, we plot the fraction of households using tap water as the main source of water for cooking in 2010 by the first year of a tap water connection

[1] Among the 404 rural communities, 90 are located in suburban areas, and the rest are located in rural areas.

at the community level. The data on first year of connection in or before 2010 is extracted from the CFPS 2010 community survey, and that after 2010 is extracted from the CFPS 2014 community survey. As shown in Figure A1, among households residing in communities that had not obtained a tap water connection by 2010, less than 10% of them were using tap water for cooking in 2010. Among households residing in communities that newly received a tap water connection in 2009, around 50% of them reported using tap water for cooking in 2010, while, for communities that got a connection before 2009, the share of households using tap water for cooking falls in the range of 50% and 80%. This suggests that once tap water becomes available in the community, households take it up rapidly, and the penetration rate does not depend on the cumulative years of tap water availability at the community level.

C Evidence About the Reliability of the Reported Data on Tap Water Availability

We provide three pieces of empirical evidence to support the reliability of the reported data on tap water availability. First, we show that the time use pattern differs for women aged 20–45 by tap water availability. When having one or more children aged 0–5, women living in communities with tap water spent more hours taking care of family members than those in communities without tap water. In contrast, when having no children aged 0–5, women in communities with tap water spent the same amount of time on family care as those without tap water but more time on labor market activities and housework. A detailed discussion on the time use pattern is provided in Section VII.A.1.

Second, the household-level tap water penetration rate differs based on tap water availability in the community. As shown in Figure A1, in communities that had not obtained a tap water connection by 2010, less than 10% of the households were using tap water for cooking in 2010, while the share of households using tap water for cooking falls in the range of 50% to 80% in communities that had received a connection before 2010. The large gap in the tap water penetration rates between communities with and without tap water connection in 2010 supports the validity of the survey question on tap water availability in the CFPS. We discuss the household take-up of tap water in more detail in Online Appendix B.

Third, the tap water availability among CFPS communities is very close to the national level. According to the *China Urban-Rural Construction Statistical Yearbook (2010)*, 52.3% of the 294,749 rural communities nationwide had obtained a tap water connection by 2010, whereas 56% of the 404 rural communities in the CFPS survey had a tap water connection in 2010.

D Baseline Water Status

Next, we detail our approach for constructing the measures for the status of baseline water sources. In the CFPS 2010 family questionnaire, households are asked to report the source of water for cooking from a list: river/lake water, well/spring water, tap water, mineral/purified/filtered water, rainwater, cellar water, pond water, and other. We calculate the fraction of households using each water source within each community. The most frequently used water source other than tap water and mineral/purified/filtered water is the baseline water source. In each community, sources used by less than 10% of residents are excluded. Based on our definition, we find that well or spring water is the baseline water source in 320 of 345 communities (92.8%), water from a river, lake, or pond is used in five communities (1.5%), and rainwater or cellar water is used in 20 communities (5.8%). We generate a dummy indicator that takes on the value one if well or spring water is not the baseline water source (not using underground water) and zero otherwise.

We take precipitation amount as an indicator of the availability of water resources. A high precipitation amount suggests abundant water resources. We extract precipitation data from the National Centers for Environmental Information and map the data to the counties. We calculate the average annual precipitation between 1951 and 2010 using the monthly data of the 221 weather stations nationwide. Then we identify the centroid and five nearest weather stations of each county and use the Inverse Distance Weighting (IDW) interpolation method to calculate the annual precipitation amount at the county level. We match the imputed county-level precipitation data to our main sample in the CFPS 2010 and generate a dummy indicator equal to one if the precipitation amount is above the median (800mm) and zero otherwise.[2]

[2]As counties in the CFPS are not identifiable, we run regressions with county-level variables

To obtain water quality at the county level, we follow Ebenstein (2012) and merge the water grade readings with the river basin data. We first draw data from the surface water quality readings of 484 geographic points across China's nine river systems in 2004. Each reading contains information about a set of monitored chemicals. The overall water grade is measured on a six-point scale, from Grade I (best) to Grade VI (worst). Ebenstein (2012) provides data on China's river basins created by the Hydro1k project conducted by the U.S. Geological Survey. Based on the data set, China can be separated into 989 basins (level 6) and a smaller set of larger basins (level 5 and level 4). For each county, we assign the average water grade reading from the smallest basin in which a monitoring station is observed. Among the 2,876 counties, 1,340 are observed within a level-6 basin in which we also observe at least one water quality monitoring station. The remaining counties are assigned the average of water grade readings of larger basins. With this algorithm, we assign water grade readings to 2,197 out of the 2,876 counties. We categorize the six water grades into three types: best (Grade I–II), medium (Grade III–IV, base group), and worst (Grade V–VI). Among the 135 CFPS counties, 95 counties have water quality data matched: 13 counties have the best water quality, 41 counties have the worst water quality, and the rest of the 41 counties have medium water quality.

E Robustness Checks in Detail

E.1 The Effect of Sanitation

Evidence from other countries has pointed to the roles of sewerage and sanitation infrastructures in child outcomes. For example, Alsan and Goldin (2019) show that the combination of clean water and effective sewerage systems accounts for approximately one-third of the decline in log child mortality in Massachusetts between 1880 and 1920. Watson (2006) shows that improving sanitation infrastructure on Native American reservations leads to sharp reductions in both waterborne gastrointestinal disease and infectious respiratory disease among Native American infants. Hence, our estimated impact of early-life exposure to tap water may

matched to our main sample at the CFPS data center in Peking University, China. All county-level variables used for the analysis throughout this paper are matched in the same way unless otherwise specified.

be confounded by the improvement in sewerage and sanitation systems.

A sewerage system is a network of pipes, pumps, and force mains for the collection of wastewater or sewage from a community (Encyclopedia Britannica).[3] The sewerage system is underdeveloped in rural China. A survey conducted by the Ministry of Construction on 74 rural communities located in 43 counties within nine provinces in 2005 shows that the sewage system and wastewater treatment system were absent in 96% of the communities (Ministry of Construction, 2005). Consequently, wastewater from different sources was discharged directly into underground or surface water. According to the *China Urban-Rural Construction Statistical Yearbooks (2006–2010)*, the wastewater treatment rate in rural communities was as low as 1% nationwide in 2006. It increased to 6% by 2010 but is still very low compared to the tap water coverage rate in rural communities in 2010 of 52.3%. As we are exploring the variation in tap water connections between 1995 and 2004, the general absence of sewerage and wastewater treatment systems suggests that our estimated effect of tap water exposure is unlikely to be driven by these systems.

We then investigate whether our estimates of the impact of tap water exposure would be confounded by the construction of sanitation facilities. Specifically, we would like to know whether our results are driven by the construction and usage of toilets. In rural China, the construction of toilets is independent of the availability of tap water and a sewerage system. Unlike the tap water system and sewerage system, the decision to construct toilets is more likely to be decentralized to the household level. As a result, we find that, in communities without a tap water connection, the share of children having a toilet at home is as high as 81.8%, only slightly lower than the share in communities with a connection, which is 83.7%. As the information on household accessibility to a toilet is only available for the year 2010 in our data, we are not able to identify the impact of early-life exposure to a toilet at the household or community level. Instead, we estimate the separate effects of tap water and toilet coverage rates during early life at the county level.

We obtain the rural tap water and toilet coverage rates at the county level in 2000 and 2005 by aggregating the accessibility of tap water and toilets among rural households in the 1% sam-

[3]See online source at: https://www.britannica.com/technology/sewerage-system.

ple of the 2000 China Population Census and the 20% sample of the 2005 China 1% Population Sample Survey.[4] During 2000 and 2005, the average tap water coverage rate increases from 30% to 36%, and the average toilet coverage rate increases from 70% to 77%. The children in our sample were born between 1995 and 2000. By our definition, their early life spans from 1994 to 2004. For each child, we extrapolate the average tap water and toilet coverage rates at the county level during early life. Years in early life falling between 1994 and 2000 are assigned the coverage rate in 2000, and years falling between 2001 and 2004 are assigned the coverage rate in 2005. If a child's early life contains the year 2000, the average coverage rate during early life is then the weighted average of the coverage rates in 2000 and 2005.

In order to estimate the effects of tap water and toilet coverage at the county level during early life on cognitive development, we estimate a DID model similar to our preferred specification in column (3) of Table 2. Because the variation in exposure is now measured at the county level, we include county fixed effects, province-specific cohort fixed effects, as well as the interactions of cohort fixed effects and county characteristics (including GDP per capita in 2000, share of urban population in 2000, and their changes from 2000 to 2010). Column (1) of Panel A in Table 4 shows that full tap water coverage at the county level during early life significantly increases the average cognitive test score by 0.760 standard deviations. Our main results show that children receiving full community-level exposure could attain an average score 0.792 standard deviations higher than a child with no exposure. Using the county-level coverage rate, therefore, generates a comparable estimate to our baseline estimated effect of community-level exposure, providing support to the validity of our measure of early-life exposure at the community level. In column (2) of Table 4, we further add the county-level toilet coverage rate into the regression, and the estimate of the effect of tap water coverage does not change. In column (3), we control for the interaction of tap water and toilet coverage rates. The result indicates that the effect of the tap water coverage rate is higher when the toilet coverage rate is high, although the estimate is insignificant.

[4]Here we aggregate microdata to get county-level measures. To ensure precise county-level measures, we only include the counties having more than 100 observations in both samples of census data. We apply the same restrictions throughout the paper when aggregating data from censuses unless otherwise specified. Removing the restrictions does not affect our results.

In Panel B of Table 4, we address the concern about the confounding impact of sanitation facilities in our preferred specification in column (3) of Table 2. We add controls for the interactions of the county toilet coverage rate in 2000 and cohort fixed effects and the interactions of changes in the county toilet coverage rate from 2000 to 2005 and cohort fixed effects, both separately and jointly. The point estimates of tap water exposure remain significant and positive. We have confirmed that the differences between the point estimates here and the baseline estimate are due to the smaller sample size caused by missing values in tap water and toilet coverage rates.

In sum, we find strong evidence that our baseline estimate of the impact of early-life exposure to tap water is unlikely to be confounded by the development of sanitation facilities.

E.2 Other Potential Confounding Factors

In this section, we show that the observed effect of tap water on cognitive skills is not driven by other confounding factors, such as other governmental programs, migration, school supply, the one-child policy, real income, urbanization, and the prevalence of left-behind children. We construct measures for these potential confounding variables at the community level whenever data are available and only use county-level measures when community data are not available.

Community macro factors

Although the communities in the CFPS 2010 are not identifiable, we can use the recalled data on the first year of having infrastructure in the community survey of the CFPS 2010 to shed some light on the macro trends in the sampled communities. In Figure A2, we plot the rollout of six types of public infrastructure between 1980 and 2010. Panel 1) plots tap water (same as Figure 1); panels 2)–6) plot electricity, cable/satellite TV, road, health facility, and landline phone, respectively. Access to tap water, cable/satellite TV, and landline phone was underdeveloped in 1980. Less than 5% of the communities had a tap water connection or cable/satellite TV, and only 20% of the communities had landline phones. All three of these types of infrastructure experienced rapid grown during 1980 and 2010 with the share of communities having each type of infrastructure increasing by over 50%. Although electricity had been introduced to 63% of the communities by 1980, we still observe a 28% increase in the share in the

next 30 years. Both the share of communities having roads and having a health facility reached 38% in 1980. However, the expansion of roads is much faster than the expansion of health facility, with an increase of 47% compared to 19%. Note that in our study the variation in years of exposure to tap water comes from the children living in communities that first received a tap water connection between 1995 and 2004 (highlighted in red).

To address the concern that government programs other than tap water may have driven our estimates, we control for early-life exposure to all these public infrastructures in our preferred specification (column (3) of Table 2). Our estimates remain positive and significant after controlling for the exposure to other public infrastructure.

In the following, we further account for some other community macro factors, including migration workers, school supply, and the one-child policy.

Migration workers — The CFPS 2010 asks each child aged 6–16 "What is the longest time (in weeks) before age three that both parents are not around?" A longer absence period of parents indicates less time with parents for the children but may also suggest better nutrition supported by remittances from parents. Since this measure is endogenous, we calculate the community's average by every two adjacent cohorts exclusive of their own households, which captures the prevalence of migrant workers in early life at the community level. We then add this new control variable to our preferred specification in column (3) of Table 2. The results reported in Panel A1 column (1) of Table A2 show that adding this control does not change our baseline results.

School supply — In the absence of a direct measure of school supply, we use a household's distance to the nearest high school from the CFPS 2010 as a proxy. We augment our preferred specification in column (3) of Table 2 with 1) the household-reported distance to the nearest high school, 2) the interactions of community average distance to the nearest high school (own household inclusive) and cohort dummies, and 3) the interactions of the community average distance (own household exclusive) and cohort dummies. Panel A1 columns (2)–(4) of Table A2 show that the estimated effects seldom change.

The one-child policy — Existing literature (e.g., Ebenstein, 2010) uses the variation of the one-child policy intensity across provinces and years. Since we have controlled for province-

cohort fixed effects in our preferred specification, the possible confounding effect of the one-child policy has been captured if the variation of the policy comes from the province-cohort level. We further account for the possible difference in time trends across communities with different implementations of the one-child policy. The community survey of the CFPS 2010 collected the details on the implementations of the one-child policy at the community level. We add the interactions of the number of births allowed and the cohort fixed effects to our preferred specification in column (3) of Table 2. The estimate in Panel A1 column (5) of Table A2 remains positive and significant. Adding the interactions of the number of births allowed without a son and cohort fixed effects or the interactions of fines and cohort fixed effects generates similar results (Panel A of Table A3).

County macro factors

Even though the counties in the CFPS are not identifiable, we seek the assistance of the CFPS data center to merge the county-level data to our sample for analysis. The 2,168 children in our main sample come from 382 rural communities located in 135 counties. In the following, we discuss the trends of real income, urbanization, migration (immigration and emigration), and left-behind children in the 135 counties and present the results from regressions accounting for the county macro factors.

Real income — We use GDP per capita as a proxy of real income. GDP per capita at the county level (2000–2010) is extracted from China Data Online and the *China City Statistical Yearbooks (2001–2011)*. GDP deflators from the *China Statistical Yearbooks* are applied to obtain the 2000 constant price. The GDP per capita is 8,720 Chinese Yuan on average in the sampled counties in 2000, and it increases by 36% between 2000 and 2010.

Urbanization — The extent of urbanization is measured by the share of urban residents among the total population. In China, most urban residents reside in urban districts, while a minority of them live in the urban areas of the counties. Urbanization in the counties is lower and progresses slower than that in urban districts. Using the county-level assemblies of the 2000 and 2010 China Population Census, we find that the share of urban residents in the 135 sampled counties increases by 2% from 24% in 2000 to 26% 2010.

Migration — We consider immigration (migrants who come to a county from other coun-

ties) and emigration (migrants who have an *hukou* in a county but leave the county for other counties) separately. We obtain the immigration rate from the county-level assemblies of the 2000 and 2010 China Population Census. The share of immigrants among the total population decreases from 20% in 2000 to 18% in 2010. The share of emigrants among households with *hukou* in the county is calculated to be 9% using the county-level assemblies of the 2000 China Population Census. In the absence of emigration information in the county-level assemblies of the 2010 China Population Census, we are not able to determine the change between 2000 and 2010. Using the 1% sample of the 2000 China Population Census and the 20% sample of the 2005 China 1% Population Sample Survey, we find that the share of emigrants drops from 9% in 2000 to 4% in 2005.

Left-behind children — We use the 1% sample of the 2000 China Population Census and the 20% sample of the 2005 China 1% Population Sample Survey to calculate the share of left-behind children. Individuals aged below 18 years are considered as children. Children are considered to be left behind if both parents are not living at home at the time of census and half left behind if at least one parent is not living at home at the time of census. We exclude 22 counties with less than 100 children in the samples. Among the remaining 113 counties with over 100 children, the share of left-behind children increases from 2% in 2000 to 6% in 2005, and the share of half left-behind children increases from 4% in 2000 to 19% in 2005.

To account for the possible impact of macro trends on our estimation, we add the interactions of cohort fixed effects and the measures of county-level macro factors as well as their changes into our preferred specification in column (3) of Table 2. Panel A2 of Table A2 shows that the estimated effects remain positive and significant. Controlling for the interactions of cohort fixed effects and the above county measures alone or the interactions of cohort fixed effects and the changes in the county measures alone produces similar results (Panel B–E of Table A3). Our estimates remain positive and significant in all specifications, although mild declines are observed in Panel E of Table A3. Note that the samples used in these regressions are smaller than the main sample due to the missing values in left-behind rates. Applying our preferred specification to the smaller samples produces results similar to the estimates in Panel E (results available upon request).

To summarize, the exercise of accounting for the macro trends at the community and county levels provides supporting evidence that our estimates are not driven by these macro trends.

E.3 Reporting Error

As mentioned in Section IV.A, in the CFPS 2010, a knowledgeable individual who has access to statistical materials in each community was invited to complete the community questionnaire. As part of the survey, they were asked to provide the first year that the community received a tap water connection, the answer to which is used to construct the treatment variable in this study. Since receiving a tap water connection is a salient and socially desirable event in the community, and as the information is available in community statistical materials such as community chronicles of events, the knowledgeable respondents are likely to report accurate years (Beckett et al., 2001). However, reporting error, if it exists and is correlated with the introduction year of tap water in the community, would bias our estimates.

We address the issue of reporting error in two ways. First, Beckett et al. (2001) suggest that data quality likely deteriorates with the length of the recall period, so we exclude communities whose reported connection year is more than 16 years, 14 years, or 12 years ago (the reported year of first connection before 1994, 1996, or 1998). The results presented in Panel B columns (1)–(3) of Table A2 show that the exclusion of "vulnerable" observations results in similar estimates to the main results, suggesting that reporting error, if any, is not driving our results.

Second, we account for the correlation between interviewers' impressions of the quality of the respondents' answers and the quality of the data (Beckett et al., 2001). After each community survey of the CFPS 2010, interviewers evaluated the reliability of the respondent's answers on a scale of 1 to 7, with 1 being very unreliable and 7 very reliable. We construct a binary reliability measure and estimate our results using the subsample of children in the reliable communities only (with a score of at least 6). The estimated effect remains positive and significant, as indicated by column (4) in Panel B of Table A2.

To further rule out potential misreporting, we replace exposure to tap water at the community level with exposure at the county level. This results in an estimate comparable with the baseline estimate, providing support to the validity of our measure of early-life exposure at the

community level. These discussions are detailed in Online Appendix E.1.

E.4 Measurement Error

In our main analysis, we define early life as the period covering one year before birth through the first five years of life. Without information on the exact tap water connection time, we construct the treatment variable, namely, the early-life years of exposure to tap water, using children's birth year and the tap water connection year at the community level. We start counting early-life exposure in the first year of connection. By doing this, we assume that all the connections occur at the beginning of a year. The observed exposure, i.e., $\text{Exposure to tap IU–5}$ in Equation (3), is computed under this assumption.

In reality, however, the connection time may be spread out over a year. Suppose we know the exact date of connection, and hence the actual exposure is $\text{Exposure to tap IU–5}^*$. The measurement error $e = \text{Exposure to tap IU–5} - \text{Exposure to tap IU–5}^*$. If e is uncorrelated with the observed exposure, i.e., $\text{Cov}(e, \text{Exposure to tap IU–5}) = 0$, our estimate will remain consistent in the presence of the measurement error. Otherwise, our estimate will be subject to biases. If we know the distribution of the exact connection time, we can test whether $\text{Cov}(e, \text{Exposure to tap IU–5}) = 0$ holds. The magnitude of biases generated due to the measurement error will depend on the imposed distributional assumptions of the connection time.

To check if our estimates would be affected by a non-random connection time, we rerun the regressions considering two extreme cases. In the first case, we assume all the connections happen in the middle of a year, and the early-life exposure in the first year of connection is counted as 0.5 year. In the second case, we assume that all the connections happen at the end of a year, and we start counting early-life exposure in the second year of connection. The results are reported in Panel C of Table A2. In both extreme cases, we get results close to our baseline estimate. We also check the scenario in which the connection timing is uniformly distributed throughout the year, and the results do not change (available upon request).

E.5 Sample Selection

Our sampled children may not be randomly selected. Children may be absent in the CFPS 2010 due to leaving home for study or work, household emigration, or other reasons. If the absence of the children in the survey is affected by the introduction of tap water in the community, our estimates may be biased. For instance, having access to tap water may improve children's school performance, which in turn may increase their probability of being admitted to prestigious middle schools outside their communities and consequently their probability of leaving home. Since the CFPS does not track children who have moved out of the community, we cannot test the above possibility directly. Instead, we rely on the existing data to shed some light.

Of the 2,700 children aged 10–15 in the CFPS 2010, there are 449 children whose families were surveyed while they themselves were not: 161 of them had left home, and the remaining 288 children were living at home but were not surveyed. There are another 83 children who completed the child survey but did not take the cognitive test. For these 532 children, we are able to calculate their actual years of exposure to tap water based on their parents' reporting of their basic information, such as gender and birth date. We estimate the effect of early-life exposure to tap water on the likelihood of the children not being included in the final sample, the likelihood of leaving home, the likelihood of not completing the child survey, and the likelihood of not taking cognitive tests in 2010. Panel D columns (1)–(4) of Table A2 show that early-life exposure to tap water does not affect any of these likelihoods.

Children could also be absent in the survey due to household emigration and other reasons. To check this, we examine the sample attrition between two adjacent CFPS waves, 2010 and 2012. We choose children who were aged 0–13 and participated in the survey in 2010. The starting point is chosen to be zero, as the information on fetuses is absent in 2010. The oldest child included was 13 in 2010 and turned 15 in 2012. Of the 3,367 children who were aged 0–13 and surveyed in 2010, 519 (15%) were not surveyed in 2012. Panel D column (5) of Table A2 shows that, for the children surveyed in the CFPS 2010, the possibility of absence from the CFPS 2012 is not affected by years of exposure to tap water in early life. Overall, we find no evidence of sample selection due to early-life tap water exposure. Therefore, our

findings are unlikely to be affected by sample selection.

E.6 Test Scores in Original Form

In our main analysis, the test scores are standardized to facilitate comparison with other child development studies. In this section, we investigate whether the results are sensitive to the transformation of test scores. As there are 34 questions in the word recognition test and 24 questions in the math test, we get the average score by first multiplying the math test score by 34/24 and then taking the average. The highest possible average score is 34. The same specification as in column (3) of Table 2 is applied for estimation. The results are reported in Panel A of Table A4. One additional year of tap water exposure during early life increases the average test score by 0.618 (out of 34). Similar to the main results, only the math test score is significantly increased (by 0.512 out of 24 and significant at the 1% level), while the effect on the word recognition test score is insignificant. Taking the math test for example, a child receiving full exposure during early life (six years) would answer three more questions correctly in the math test than a child without any exposure, comparable to the estimate of four when the scores are standardized. In general, we find that our results are robust to the transformation of the cognitive test scores.

F Tap Water Coverage Rate at the County Level

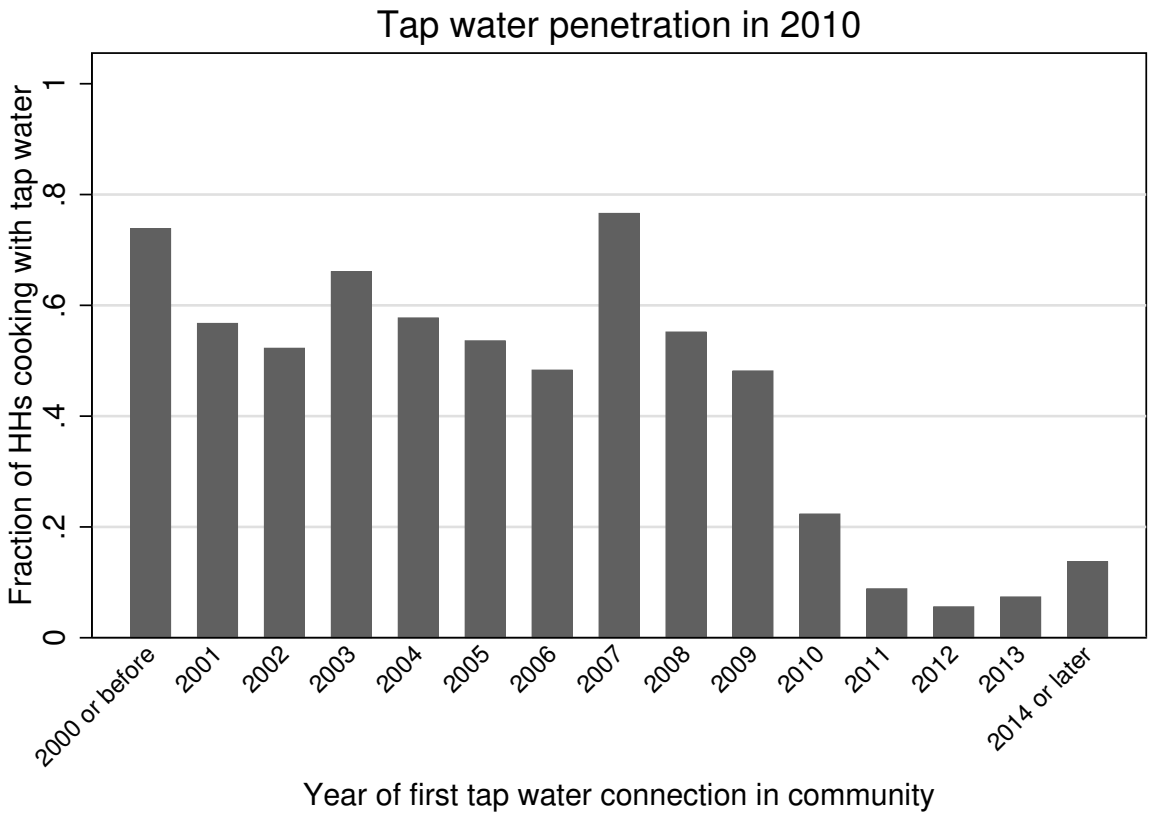
In this section, we present the change in tap water coverage at the county level over time and discuss reasons why the community level measure is used in the main analysis.

Using the county-level assemblies of the 2000 and 2010 China Population Census, we plot the nationwide county-level tap water coverage rates in 2000 and 2010, respectively (see Figure A3). We observe a rapid expansion of tap water connection nationwide during 2000 and 2010, and more areas having a high coverage rate in 2010 than in 2000.

We use the rural population in the 1% sample of the 2000 China Population Census and the 20% sample of the 2005 China 1% Population Sample Survey to calculate the average county tap water coverage rate during early life for the sampled children. Although we find that using the county tap water coverage rate generates comparable results to the baseline estimate (see

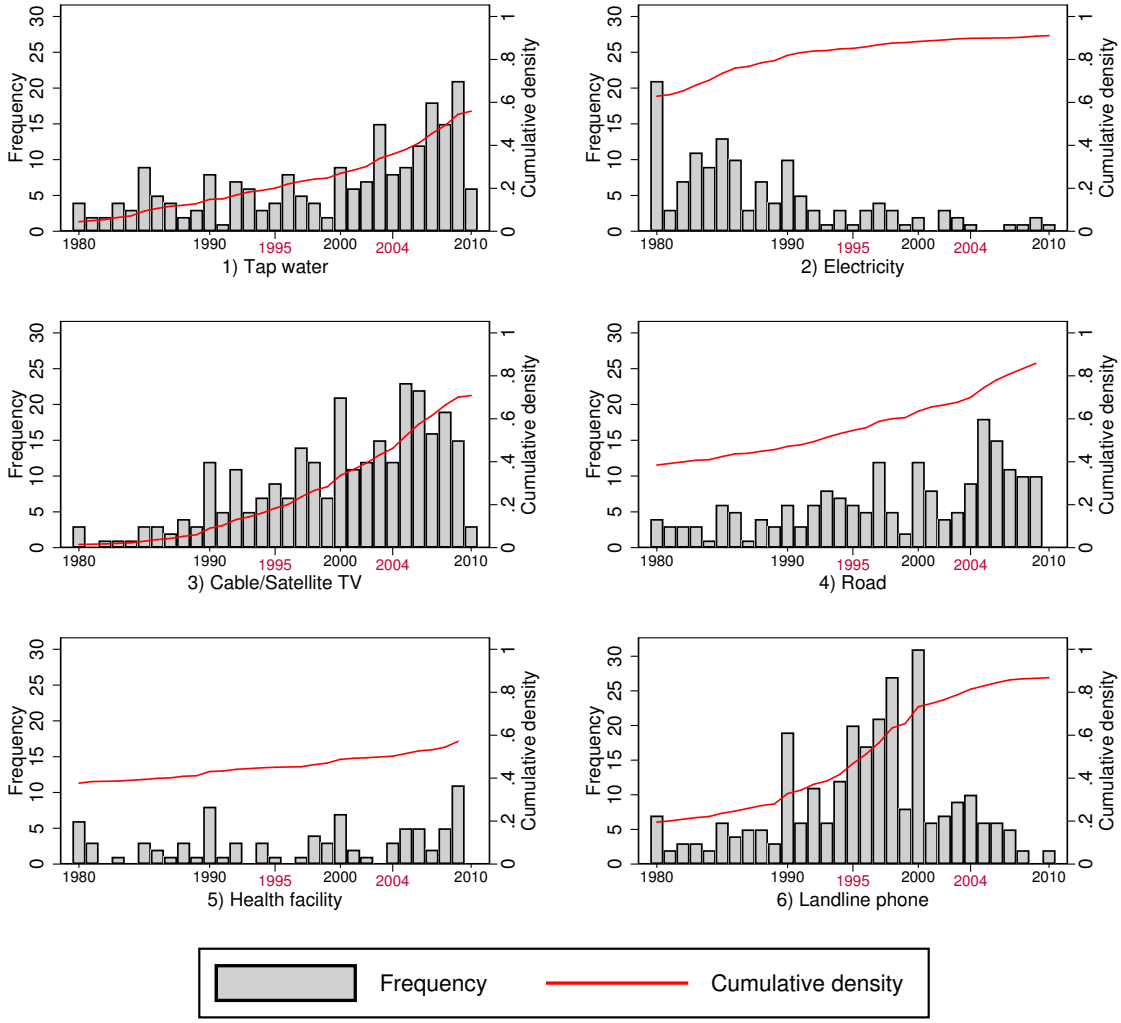
Section VI.A), we prefer to use early-life exposure to tap water at the community level for two reasons. First, it is more relevant to policy to understand the effect of exposure at the community level, as the construction of tap water infrastructure is decentralized, and the timing of connection could vary across communities within a county. To get a sense of the size of the variation in tap water connection year across communities in the same county, we examine the difference between the earliest and latest connection years for communities within counties. The children in the baseline analysis come from 382 rural communities in 135 counties. We exclude 18 counties that have only one rural community surveyed by the CFPS 2010. Each of the remaining 117 counties has 2–5 rural communities covered in the survey. The difference in connection years can be categorized into eight types: no connection within the county by 2010, 0–4 years, 5–9 years, 10–14 years, 15–19 years, 20–24 years, 25–29 years, and 30+ years. Note that the difference here is underestimated for counties in which not all communities had received a connection by 2010. Figure A4 plots the distribution of the difference. Around 19% of the counties had no connection for all sampled rural communities. For 21% of the counties, the difference ranges from zero to four years. In over 60% of the counties, the difference is at least five years. The plotting suggests that using the county-level coverage rate would smooth out the variation across communities.

Second, the measure of county-level tap water coverage rate has a few drawbacks. Censuses are available for every five years, but tap water coverage rate of rural population can only be calculated using the 1% sample of the 2000 China Population Census and the 20% sample of the 2005 China 1% Population Sample Survey. If the expansion of tap water is not linear, the weighted average of the county tap water coverage rate during early life calculated using two waves of censuses may be a poor measure of the actual exposure. Without information on the tap water coverage rate in other periods of life for the children aged 10–15 in 2010, we are unable to prove that early life is the most critical time window in the process of cognitive development. Moreover, the county-level aggregated data from the samples of census data might be imprecise measures of the true coverage rates. To obtain precise measures, we have to drop counties with too few observations and reduce the sample size.



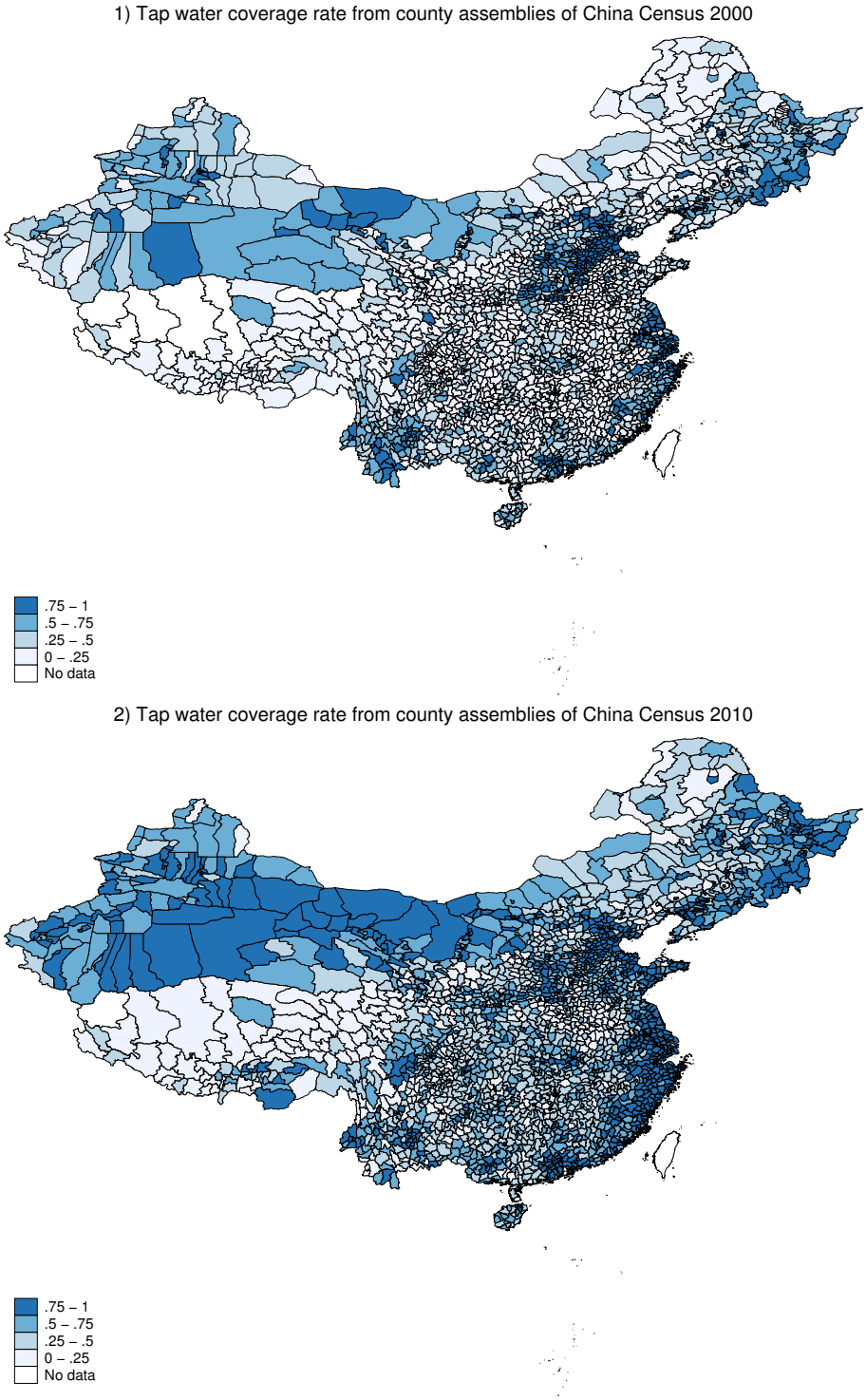
Data source: China Family Panel Studies (CFPS) 2010 and 2014.

Figure A1: Tap water penetration in rural China by first year of connection



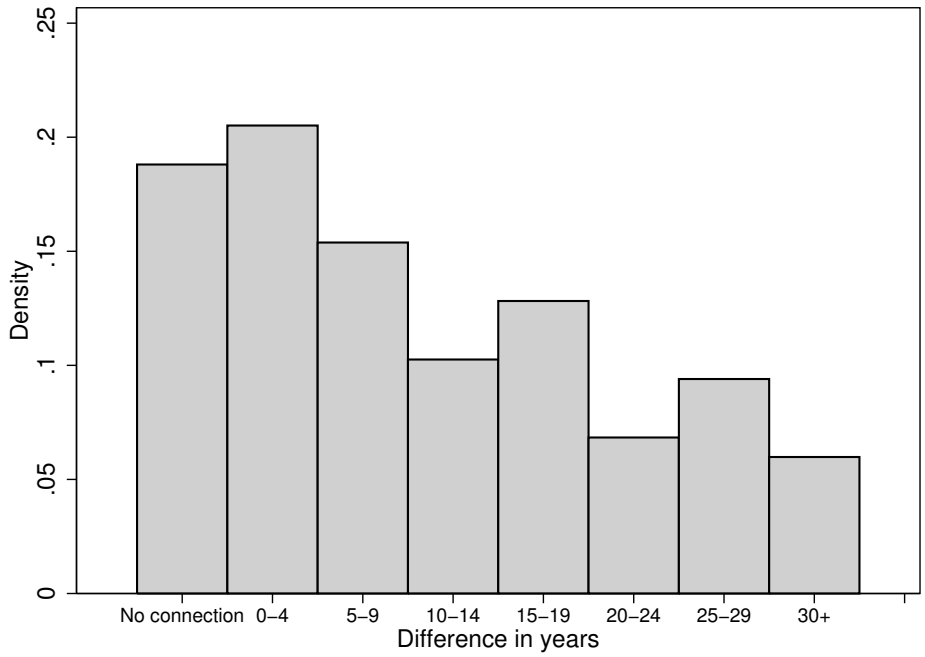
Notes: Data from the China Family Panel Studies (CFPS) 2010. X-axis indicates the first year of connection at the community level. The histogram plots the number of new connections each year, and the red curve plots the cumulative distribution of the community-level rollout during 1980 and 2010 among 404 rural communities. The variation in the treatment variable in this study comes from the children living in communities that first obtained a tap water connection between 1995 and 2004 (highlighted in red).

Figure A2: Trends of infrastructure construction in rural communities in China



Data source: County-level assemblies of the 2000 and 2010 China Population Census.

Figure A3: County-level coverage rate of tap water in 2000 and 2010



Notes: Data from the China Family Panel Studies (CFPS) 2010.
"Difference in years" indicates the gap between the earliest tap water connection year and the latest tap water connection year (or the year of 2010 if having no connection by 2010) among rural communities in a county. "No connection" indicates no tap water connection by 2010 for all sampled rural communities within a county.

Figure A4: Difference between earliest and latest connection years within counties

Table A1: Determinants of tap water connection at the community level

Dep. var.:	Connected to tap water in 2010	Year first getting connection
	(1)	(2)
Suburban village	0.109*	-3.319***
	(0.064)	(1.105)
Log of distance to nearest town or city (hours)	-0.077**	1.732***
	(0.033)	(0.541)
Log of population	-0.023	-0.909
	(0.041)	(0.632)
Log of area (km^2)	0.007	-0.018
	(0.012)	(0.197)
Average years of schooling of 25–55-year-olds	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%.
Column (1) reports OLS estimates. Column (2) reports the marginal effects of the Tobit model, conditional on community first getting tap water access no later than 2010. The communities are categorized into five types by topographic characteristics: hills, mountains, plateaus, plains, and others, and dummy variables are generated for each type. The type of plains is taken as the omitted group. Robust standard errors are in parentheses.

Table A2: Other robustness checks

	(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 inclusive)× cohort	+ Community average (own exclusive) × cohort	+ Number of births allowed × cohort
Exposure to tap IU–5	0.133***	0.131***	0.138***	0.138***	0.115**
	(0.046)	(0.046)	(0.047)	(0.047)	(0.051)
Observations	2,168	2,168	2,168	2,168	2,153
R^2	0.449	0.449	0.450	0.450	0.452
A2. At county level (Dep. var.: Average cognitive test score)					
	GDP per capita[a]	Share of urban[b]	Emigration[c]	Immigration[b]	Left-behind[c]
	+ GDP pc in 2000 × cohort + GDP pc growth rate × 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
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: * means significant at 10%, ** significant at 5%, and *** significant at 1%.
[a] Data from China Data Online (CDO) and the *China City Statistical Yearbooks*.
[b] Data from the county-level assemblies of the 2000 and 2010 China Population Census.
[c] Data from the 1% sample of the 2000 China Population Census and the 20% sample of the 2005 China 1% Population Sample Survey. Only counties with more than 100 observations in both samples are included.
Each column in each panel contains estimates from a separate regression with the same specification as in column (3) of Table 2. Standard errors in parentheses are 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 take the cognitive tests.

Table A3: Accounting for the macro trends at community and county level

	Dep. var.: Average cognitive test score					
	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A. One-child policy					
	+ Number of births allowed × cohort	+ Number of births allowed without son × cohort	+ Fines × cohort			
Exposure to tap IU–5	0.115**	0.127***	0.123**			
	(0.051)	(0.046)	(0.054)			
Observations	2,153	2,139	1,967			
R^2	0.452	0.453	0.456			
	Panel B. GDP per capita[a]			***Panel C. Share of urban population***[b]		
	+ GDP pc in 2000 × cohort	+ GDP pc growth rate × cohort	+ GDP pc in 2000 × cohort + GDP pc growth rate × cohort	+ Share of urban in 2000 × cohort	+ Change in share of urban × cohort	+ Share of urban in 2000 × cohort + Change in share of urban × cohort
Exposure to tap IU–5	0.143***	0.149***	0.145***	0.127***	0.132***	0.130***
	(0.051)	(0.049)	(0.052)	(0.046)	(0.045)	(0.046)
Observations	2,133	2,133	2,133	2,164	2,164	2,164
R^2	0.450	0.451	0.452	0.449	0.450	0.450
	Panel D. Migration					
	a. Emigration[c]			***b. Immigration***[b]		
	+ Emigration rate in 2000 × cohort	+ Change in emigration rate × cohort	+ Emigration rate in 2000 × cohort + Change in emigration rate × cohort	+ Immigration rate in 2000 × cohort	+ Change in immigration rate × cohort	+ Immigration rate in 2000 × cohort + Change in immigration rate × cohort
Exposure to tap IU–5	0.139***	0.128***	0.134***	0.129***	0.135***	0.131***
	(0.047)	(0.046)	(0.047)	(0.047)	(0.048)	(0.049)
Observations	2,098	2,098	2,098	2,164	2,164	2,164
R^2	0.452	0.451	0.453	0.450	0.449	0.450
	Panel E. Left behind children[c]					
	+ Left-behind rate in 2000 × cohort	+ Change in left-behind rate × cohort	+ Left-behind rate in 2000 × cohort + Change in left-behind rate × cohort	+ Half-left-behind rate in 2000 × cohort	+ Change in half-left-behind rate × cohort	+ Half-left-behind rate in 2000 × cohort + Change in half-left-behind rate × cohort
Exposure to tap IU–5	0.101**	0.098*	0.102*	0.101*	0.107**	0.110**
	(0.051)	(0.051)	(0.052)	(0.052)	(0.050)	(0.051)
Observations	1,946	1,946	1,946	1,946	1,946	1,946
R^2	0.455	0.455	0.456	0.455	0.457	0.458

Notes: * means significant at 10%, ** significant at 5%, and *** significant at 1%.
[a] Data from China Data Online (CDO) and the *China City Statistical Yearbooks*.
[b] Data from the county-level assemblies of the 2000 and 2010 China Population Census.
[c] Data from the 1% sample of the 2000 China Population Census and the 20% sample of the 2005 China 1% Population Sample Survey. Only counties with more than 100 observations in both samples are included.
Each column in each panel contains estimates from a separate regression with the same specification as in column (3) of Table 2. Standard errors in parentheses are clustered at the community level.

Table A4: Specification tests

	(1)	(2)	(3)
Panel A. Test scores in original form			
Dep. var.:	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
Panel B. Exposure in other forms			
	Dep. var.: Average cognitive test score		
Full exposure to tap IU–5	0.576**		
	(0.244)		
Partial exposure to tap IU–5	0.291**		
	(0.145)		
One year tap IU–5		0.100	
		(0.203)	
Two years tap IU–5		0.412*	
		(0.210)	
Three years tap IU–5		0.559**	
		(0.238)	
Four years tap IU–5		0.210	
		(0.310)	
Five years tap IU–5		0.612*	
		(0.319)	
Six years tap IU–5		0.774***	
		(0.271)	
Exposure to tap IU–5			0.154
			(0.095)
(Exposure to tap IU–5)^2			-0.004
			(0.013)
Observations	2,168	2,168	2,168
R^2	0.448	0.450	0.448

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

In Panel A, each column is estimated from a separate regression with the same specification applied as in column (3) of Table 2. All test scores are in the original form. "Average cognitive test score (original)" is obtained by first multiplying the math test score by 34/24 and then taking the average.

In Panel B, each column is estimated from a separate regression with Exposure to tap IU–5 in column (3) of Table 2 replaced by exposure in other forms. "No exposure to tap IU–5" is omitted in both columns (1) and (2).

In both panels, standard errors in parentheses are clustered at the community level.

Table A5: The impact of tap water exposure on cognitive skills for children aged 10–17

Dep. var.:	Average cognitive test score			Word	Math
	(1)	(2)	(3)	(4)	(5)
Exposure to tap IU–5	0.050**	0.056*	0.056*	0.016	0.085**
	(0.025)	(0.031)	(0.031)	(0.033)	(0.034)
Boy	-0.101**	-0.103**	-0.107**	-0.139***	-0.030
	(0.039)	(0.043)	(0.043)	(0.044)	(0.040)
Mother's age at birth	-0.004	-0.004	-0.004	-0.002	-0.004
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Number of siblings	-0.049	-0.054	-0.054	-0.022	-0.078**
	(0.034)	(0.035)	(0.036)	(0.035)	(0.038)
Number of household members	0.012	0.016	0.015	0.006	0.023
	(0.017)	(0.017)	(0.018)	(0.017)	(0.019)
Father's years of schooling	0.033***	0.032***	0.032***	0.031***	0.025***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Mother's years of schooling	0.016***	0.015**	0.015**	0.010*	0.018***
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)
Observations	2,782	2,782	2,782	2,782	2,782
R^2	0.380	0.430	0.431	0.427	0.370
Number of clusters	388	388	388	388	388
Cohort FE	Yes	No	No	No	No
Community FE	Yes	Yes	Yes	Yes	Yes
Cohort-province FE, community controls × cohort	No	Yes	Yes	Yes	Yes
Early-life exposure to other public facilities	No	No	Yes	Yes	Yes

Notes: * means significant at 10%, ** significant at 5%, and *** significant at 1%.
In addition to the 2,168 children aged 10–15 in the main sample, we include 615 children aged 16–17 in 2010 for analysis. In column (2), the community characteristics used to construct interactions with cohort dummies are an indicator of suburban village, log of distance to nearest town or city, log of area, average years of schooling of 25–55-year-olds, and topographic characteristics (hills, mountains, plateaus, plains, and others). Columns (3)–(5) also include early-life exposure to other public facilities, including electricity, cable/satellite TV, roads, health facilities, and landline phones. Standard errors in parentheses are clustered at the community level.

Table A6: The impact of tap water exposure on cognitive skills (adding household income and household doing non-farm work)

	Dep. var.: Average cognitive test score		
	(1)	(2)	(3)
Exposure to tap IU–5	0.125***	0.121***	0.121***
	(0.047)	(0.047)	(0.046)
Log of household gross income last year (yuan)	0.007		0.011
	(0.041)		(0.041)
Household doing non-farm work		0.166**	0.167**
		(0.084)	(0.084)
Boy	-0.113**	-0.111**	-0.112**
	(0.056)	(0.056)	(0.056)
Mother's age at birth	0.010	0.010	0.010
	(0.007)	(0.007)	(0.007)
Number of siblings	-0.015	-0.011	-0.011
	(0.059)	(0.060)	(0.060)
Birth order	-0.089*	-0.089*	-0.089*
	(0.048)	(0.048)	(0.048)
Number of household members	0.009	0.011	0.010
	(0.021)	(0.021)	(0.021)
Father's years of schooling	0.032***	0.031***	0.031***
	(0.008)	(0.008)	(0.008)
Mother's years of schooling	0.016*	0.014*	0.014*
	(0.008)	(0.008)	(0.008)
Observations	2,069	2,069	2,069
R^2	0.454	0.456	0.456

Notes: * means significant at 10%, ** significant at 5%, and *** significant at 1%.
The same specification as in column (3) of Table 2 is applied for estimation. Standard errors in parentheses are clustered at the community level.

References for Online Appendix

Alsan, Marcella, and Claudia Goldin. 2019. "Watersheds in Child Mortality: The Role of Effective Water and Sewerage Infrastructure, 1880–1920." Journal of Political Economy, 127(2): 586–638.

Beckett, Megan, Julie Da Vanzo, Narayan Sastry, Constantijn Panis, and Christine Peterson. 2001. "The Quality of Retrospective Data: An Examination of Long-Term Recall in a Developing Country." Journal of Human Resources, 36(3): 593–625.

Ebenstein, Avraham. 2010. "The "Missing Girls" of China and the Unintended Consequences of the One Child Policy." Journal of Human Resources, 45(1): 87–115.

_____. 2012. "The Consequences of Industrialization: Evidence from Water Pollution and Digestive Cancers in China." Review of Economics and Statistics, 94(1): 186–201.

Ministry of Construction Village and Town Construction Office. 2005. "The Current Situation and Problems of Village Living Environment." Beijing, China.

Ministry of Construction Department of Integrated Finance. 2006. China Urban-Rural Construction Statistical Yearbook. China Architecture & Building Press, Beijing, China.

Ministry of Housing and Urban-Rural Development. 2010. China Urban-Rural Construction Statistical Yearbook. China Planning Press, Beijing, China.

National Bureau of Statistics Urban Socioeconomic Survey Team. 2001. China City Statistical Yearbook. Beijing, China.

_____. 2011. China City Statistical Yearbook. Beijing, China.

Watson, Tara. 2006. "Public Health Investments and the Infant Mortality Gap: Evidence from Federal Sanitation Interventions on U.S. Indian Reservations." Journal of Public Economics, 90(8–9): 1537–1560.