Can White Elephants Kill?
The Unintended Consequences of Infrastructure Development

Antonella Bancalari

St. Andrews and IFS

NBER SI, Urban Economics, 2022
Investing in infrastructure

- Driver of productivity and economic development [Aschauer, 1989; Isham and Kaufmann, 1999].

- Total spending in infrastructure projects in LMICs $\approx 1.2$ trillion USD $\rightarrow 5\%$ of the global GDP [Fay et al., 2019].

- Ample evidence on the effectiveness of completed and in-use infrastructure to improve living standards:
  - Sewers, water-pipes, dams, and electricity and transportation networks [e.g. Watson, 2006; Duflo and Pande, 2007; Rud, 2012; Lipscomb et al., 2013; Donaldson, 2018; Alsan and Goldin, 2019; Bhalotra et al., 2021].

- Yet, we know little about the consequences of infrastructure projects that are being implemented.
Investing in infrastructure

• Driver of productivity and economic development [Aschauer, 1989; Isham and Kaufmann, 1999].

  • Total spending in infrastructure projects in LMICs $\approx 1.2$ trillion USD $\rightarrow 5\%$ of the global GDP [Fay et al., 2019].

• Ample evidence on the effectiveness of completed and in-use infrastructure to improve living standards:

  • Sewers, water-pipes, dams, and electricity and transportation networks [e.g. Watson, 2006; Duflo and Pande, 2007; Rud, 2012; Lipscomb et al., 2013; Donaldson, 2018; Alsan and Goldin, 2019; Bhalotra et al., 2021].

• Yet, we know little about the consequences of infrastructure projects that are being implemented.
Investing in infrastructure

- Driver of productivity and economic development [Aschauer, 1989; Isham and Kaufmann, 1999].
  - Total spending in infrastructure projects in LMICs $\approx 1.2$ trillion USD $\rightarrow$ 5% of the global GDP [Fay et al., 2019].

- Ample evidence on the effectiveness of completed and in-use infrastructure to improve living standards:
  - Sewers, water-pipes, dams, and electricity and transportation networks [e.g. Watson, 2006; Duflo and Pande, 2007; Rud, 2012; Lipscomb et al., 2013; Donaldson, 2018; Alsan and Goldin, 2019; Bhalotra et al., 2021].

- Yet, we know little about the consequences of infrastructure projects that are being implemented.
Implementing infrastructure

• Construction works expose the local population to **hazards** and **disruptions**.

• Implementation plagued by inefficiencies that increase private and **social costs** in LMICs and HICs alike.
  
  • Delays and cost over-runs in up to 45% of OECD projects [Flyvbjerg, Holm, and Buhl, 2004].
  
  • 1/3 of infrastructure projects are halted and even abandoned mid-construction [Rasul and Rogger, 2018; Williams, 2017].

• Overestimated welfare evaluations because social costs from the implementation phase are ignored.

\[
\frac{B^c}{C_{priv}^i + C_{soc}^i} \tag{1}
\]
Implementing infrastructure

- Construction works expose the local population to hazards and disruptions.

- Implementation plagued by inefficiencies that increase private and social costs in LMICs and HICs alike.
  
  - Delays and cost over-runs in up to 45% of OECD projects [Flyvbjerg, Holm, and Buhl, 2004].
  
  - 1/3 of infrastructure projects are halted and even abandoned mid-construction [Rasul and Rogger, 2018; Williams, 2017].

- Overestimated welfare evaluations because social costs from the implementation phase are ignored.

\[
\frac{B^c}{C^i_{priv} + C^i_{soc}}
\]  

(1)
Implementing infrastructure

- Construction works expose the local population to hazards and disruptions.

- Implementation plagued by inefficiencies that increase private and social costs in LMICs and HICs alike.
  
  - Delays and cost over-runs in up to 45% of OECD projects [Flyvbjerg, Holm, and Buhl, 2004].
  
  - 1/3 of infrastructure projects are halted and even abandoned mid-construction [Rasul and Rogger, 2018; Williams, 2017].

- Overestimated welfare evaluations because social costs from the implementation phase are ignored.

\[
\frac{B^c}{C^i_{priv} + C^i_{soc}}
\]  (1)
This paper

- **New focus:** study the social costs imposed by infrastructure projects that are being *implemented* (i.e., projects started but not yet completed).

- **Context:** nation-wide implementation of sewerage projects in urban areas of Peru and its effect on early-life mortality.

- **Data:** novel administrative data matched to spatial data to construct a panel of more than 1,400 districts spanning 2005-2015.

- **Strategy:** exploiting geographic features and nationwide availability of funds as an instrument.
This paper

- **New focus**: study the social costs imposed by infrastructure projects that are being **implemented** (i.e., projects started but not yet completed).

- **Context**: nation-wide implementation of sewerage projects in urban areas of Peru and its effect on early-life mortality.

- **Data**: novel administrative data matched to spatial data to construct a panel of more than 1,400 districts spanning 2005-2015.

- **Strategy**: exploiting geographic features and nationwide availability of funds as an instrument.
This paper

• **New focus**: study the social costs imposed by infrastructure projects that are being **implemented** (i.e., projects started but not yet completed).

• **Context**: nation-wide implementation of sewerage projects in urban areas of Peru and its effect on early-life mortality.

• **Data**: novel administrative data matched to spatial data to construct a panel of more than 1,400 districts spanning 2005-2015.

• **Strategy**: exploiting geographic features and nationwide availability of funds as an instrument.
This paper

• **New focus**: study the social costs imposed by infrastructure projects that are being **implemented** (i.e., projects started but not yet completed).

• **Context**: nation-wide implementation of sewerage projects in urban areas of Peru and its effect on early-life mortality.

• **Data**: novel administrative data matched to spatial data to construct a panel of more than 1,400 districts spanning 2005-2015.

• **Strategy**: exploiting geographic features and nationwide availability of funds as an instrument.
**Preview of findings**

- Projects increased mortality while being implemented, rather than not starting projects, due to hazards, infections and unsafe behavior.

- Most projects were halted mid-construction, exacerbating the effects.

*Source: Defensoria del Pueblo, Peru, 2016*
Preview of findings

- Projects increased mortality while being implemented, rather than not starting projects, due to hazards, infections and unsafe behavior.

- Most projects were halted mid-construction, exacerbating the effects.

Source: Defensoria del Pueblo, Peru, 2016
Data

- Budgetary reports for 6,000 projects in 1,400 districts:
  - Years of start and completion
  - **Number of implemented projects** per district–year

- Vital statistics and population forecasts:
  - Deaths by cause (ICD-10)
  - Infant (IMR) and under-5 mortality rate (U5MR) per 1,000 children per district–year

\[
\frac{\text{deaths of infants (children)}_{dt}}{\text{infant (child) population}_{dt}} \times 1,000
\]

- Grid-cell level spatial data: area, elevation, gradient and river density.

- Census, municipal records and sectoral public expenditure.
Data

- Budgetary reports for 6,000 projects in 1,400 districts:
  - Years of start and completion
  - **Number of implemented projects** per district–year

- Vital statistics and population forecasts:
  - Deaths by cause (ICD-10)
  - **Infant (IMR) and under-5 mortality rate (U5MR)** per 1,000 children per district–year
    \[
    \frac{\text{deaths of infants (children)}}{\text{infant (child) population}} \times 1,000
    \]

- Grid-cell level spatial data: area, elevation, gradient and river density.

- Census, municipal records and sectoral public expenditure.
Data

- Budgetary reports for 6,000 projects in 1,400 districts:
  - Years of start and completion
  - **Number of implemented projects** per district–year

- Vital statistics and population forecasts:
  - Deaths by cause (ICD-10)
  - **Infant (IMR) and under-5 mortality rate (U5MR)** per 1,000 children per district–year
    \[
    \frac{\text{deaths of infants (children)}}{\text{infant (child) population}} \times 1,000
    \]

- Grid-cell level spatial data: area, elevation, gradient and river density.

- Census, municipal records and sectoral public expenditure.
Data

- Budgetary reports for 6,000 projects in 1,400 districts:
  - Years of start and completion
  - **Number of implemented projects** per district–year

- Vital statistics and population forecasts:
  - Deaths by cause (ICD-10)
  - **Infant (IMR) and under-5 mortality rate (U5MR)** per 1,000 children per district–year
    \[
    \frac{\text{deaths of infants (children)}}{\text{infant (child) population}} \times 1,000
    \]

- Grid-cell level spatial data: area, elevation, gradient and river density.

- Census, municipal records and sectoral public expenditure.
Specification

\[ MR_{dt} = \beta S_{dt} + \gamma_d + \delta_t + \nu_{dt} \quad (2) \]

- \( MR_{dt} \) = mortality rate in district \( d \) and year \( t \)
- \( S_{dt} \) = number of implemented projects
- \( \gamma_d \) = district fixed effect
- \( \delta_t \) = year fixed effect
- \( \nu_{dt} \) = error term, standard error clustered at district level

- Sample: district–years before project completion.
- Counterfactual scenario: no project implemented.
**Specification**

\[ MR_{dt} = \beta S_{dt} + \gamma_d + \delta_t + \nu_{dt} \]  

- \( MR_{dt} \) = mortality rate in district \( d \) and year \( t \)
- \( S_{dt} \) = number of implemented projects
- \( \gamma_d \) = district fixed effect
- \( \delta_t \) = year fixed effect
- \( \nu_{dt} \) = error term, standard error clustered at district level

- Sample: district–years before project completion.
- Counterfactual scenario: no project implemented.
Empirical strategy

- Instrumental variable to deal with endogenous placement and timing of project implementation.

- Counterfactual implementation of projects, commonly used in the infrastructure literature [Duflo and Pande, 2007; Lipscomb et al., 2013; Burgess et al., 2015] based on:

  1. District-level geographical suitability to develop low-cost sewerage projects.
  2. Over time changes in national funds for sewerage projects.
Empirical strategy

• Instrumental variable to deal with endogenous placement and timing of project implementation.

• Counterfactual implementation of projects, commonly used in the infrastructure literature [Duflo and Pande, 2007; Lipscomb et al., 2013; Burgess et al., 2015] based on:

  1. District-level geographical suitability to develop low-cost sewerage projects.

  2. Over time changes in national funds for sewerage projects.
Instrumental variable

- Time-varying predicted projects:

\[ P_{dt} = \min(M, P_{dt-1} + I(fundrank_d \leq maxfund_t)) \]  (3)

- \( fundrank_d \rightarrow \) funding rank of district \( d \) based on its geographic suitability
- \( maxfund_t \rightarrow \) maximum fundable projects given the nationwide budget in year \( t \)
- \( M \rightarrow \) maximum prediction per district \( d \) (based on median)

- Identification assumption: no other demand-side factors, policies or infrastructure evolved over time following same spatial lines.
Instrumental variable

- Time-varying predicted projects:

\[
P_{dt} = \min(M, P_{dt-1} + I(fundrank_d \leq maxfund_t))
\] (3)

- \(fundrank_d\) → funding rank of district \(d\) based on its geographic suitability

- \(maxfund_t\) → maximum fundable projects given the nationwide budget in year \(t\)

- \(M\) → maximum prediction per district \(d\) (based on median)

- Identification assumption: no other demand-side factors, policies or infrastructure evolved over time following same spatial lines.
Sewerage implementation in Peru, snapshot of 2010

Antonella Bancalari (St.Andrews and IFS)

Can White Elephants Kill?

NERB SI, Urban Economics, 2022
Validity of the exclusion restriction

1. Lagged development indicators have no effect on the number of predicted projects.

2. While sewerage-intense districts experienced steeper mortality trends, low- and high-geographically suitable districts exhibit parallel trends.

<table>
<thead>
<tr>
<th>Dependent variable: Instrument for implemented sewerage projects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unit: Projects</td>
</tr>
<tr>
<td>(1) (2) (3) (4)</td>
</tr>
<tr>
<td>Population density ((t-5)) 0.000 (0.000) [0.937]</td>
</tr>
<tr>
<td>Revenues ((t-5)) 0.029 (0.037) [0.434]</td>
</tr>
<tr>
<td>Human resources ((t-5)) -0.000 (0.000) [0.245]</td>
</tr>
<tr>
<td>Internet access ((t-5)) 0.038 (0.030) [0.210]</td>
</tr>
<tr>
<td>District-year 8,448 6,889 8,345 8,414</td>
</tr>
<tr>
<td>Districts 1,408 1,408 1,408 1,408</td>
</tr>
</tbody>
</table>

Note. Estimates based on district-level panel data spanning the years 2005–2015. The dependent variable is the cumulative number of “predicted projects.” The regressors are the five-year-lagged values of: population density (population per km\(^2\)) in column 1; municipal revenues (hyperbolic syne transformation) in column 2; municipal total human resources in column 3; and an indicator equal to one if the municipality has Internet access, and zero otherwise, in column 4. All coefficients are estimated with ordinary least-squares (OLS), including district and year fixed effects. Standard errors clustered by district are reported in parentheses and \(p\)-values in brackets. See Online Appendix D for variable definitions and see the text for further details.
Validity of the exclusion restriction

1. Lagged development indicators have no effect on the number of predicted projects.

2. While sewerage-intense districts experienced steeper mortality trends, low- and high-geographically suitable districts exhibit parallel trends.
Main result: implementing infrastructure kills

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>IMR</th>
<th>U5MR</th>
<th>IMR</th>
<th>U5MR</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Unit:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IMR U5MR</td>
<td>Deaths per 1,000 infants or children</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) OLS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implemented projects</td>
<td>0.000</td>
<td>0.057</td>
<td>0.003</td>
<td>0.660</td>
</tr>
<tr>
<td>(0.000)</td>
<td>(0.016)</td>
<td>(0.001)</td>
<td>(0.312)</td>
<td></td>
</tr>
<tr>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[0.058]</td>
<td>[0.034]</td>
<td></td>
</tr>
<tr>
<td>Anderson–Rubin $p$-value</td>
<td>0.034</td>
<td>0.014</td>
<td>0.083</td>
<td>0.140</td>
</tr>
<tr>
<td>Mean (initial)</td>
<td>0.018</td>
<td>4.818</td>
<td>0.018</td>
<td>4.818</td>
</tr>
<tr>
<td>District-year</td>
<td>8,555</td>
<td>8,555</td>
<td>8,555</td>
<td>8,555</td>
</tr>
<tr>
<td>Districts</td>
<td>1,379</td>
<td>1,379</td>
<td>1,379</td>
<td>1,379</td>
</tr>
</tbody>
</table>

Notes. Clustered standard errors at the district level in parenthesis and $p$-values in brackets. District and year fixed effects included.
Robustness checks

Mechanisms

- Increase in mortality caused by waterborne diseases and accidents.  
  IMR  U5MR

- Increase in use of unsafe water for drinking purposes and sanitation practices.  WASH

- Alternative channels: no changes in fertility, migration and selective migration.  Demography
Mechanisms

- Increase in mortality caused by waterborne diseases and accidents.  
  - IMR  
  - U5MR

- Increase in use of unsafe water for drinking purposes and sanitation practices.  
  - WASH

- Alternative channels: no changes in fertility, migration and selective migration.  
  - Demography
Mechanisms

- Increase in mortality caused by waterborne diseases and accidents.  
  IMR  U5MR

- Increase in use of unsafe water for drinking purposes and sanitation practices.  WASH

- Alternative channels: no changes in fertility, migration and selective migration.  Demography
↑ unsafe water and sanitation behaviour

Many halted projects and for long

More than 70% of projects were halted, increasing average duration 3 to 5y
## Heterogenous effects by halting status

<table>
<thead>
<tr>
<th>Dependent variable: IMR</th>
<th>U5MR</th>
<th>IMR</th>
<th>U5MR</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unit:</strong></td>
<td>Deaths per 1,000 infants or children</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(1)</strong></td>
<td><strong>(2)</strong></td>
<td><strong>(3)</strong></td>
<td><strong>(4)</strong></td>
</tr>
<tr>
<td><strong>Panel A</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLS 2SLS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implemented projects</td>
<td>0.000</td>
<td>0.057</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.016)</td>
<td>(0.001)</td>
</tr>
<tr>
<td></td>
<td>[0.000]</td>
<td>[0.000]</td>
<td>[0.058]</td>
</tr>
<tr>
<td>Anderson–Rubin p-value</td>
<td>0.034</td>
<td>0.014</td>
<td></td>
</tr>
<tr>
<td>Mean (initial)</td>
<td>0.018</td>
<td>4.818</td>
<td>0.018</td>
</tr>
<tr>
<td>District-year</td>
<td>8,555</td>
<td>8,555</td>
<td>8,555</td>
</tr>
<tr>
<td>Districts</td>
<td>1,379</td>
<td>1,379</td>
<td>1,379</td>
</tr>
<tr>
<td><strong>Panel B</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2SLS 2SLS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No halting in district</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implemented projects</td>
<td>0.003</td>
<td>0.633</td>
<td>0.006</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.455)</td>
<td>(0.004)</td>
</tr>
<tr>
<td></td>
<td>[0.107]</td>
<td>[0.165]</td>
<td>[0.138]</td>
</tr>
<tr>
<td>Anderson–Rubin p-value</td>
<td>0.083</td>
<td>0.140</td>
<td>0.119</td>
</tr>
<tr>
<td>Mean (initial)</td>
<td>0.018</td>
<td>4.818</td>
<td>0.018</td>
</tr>
<tr>
<td>District-year</td>
<td>5,236</td>
<td>5,236</td>
<td>4,346</td>
</tr>
<tr>
<td>Districts</td>
<td>1,009</td>
<td>1,009</td>
<td>968</td>
</tr>
</tbody>
</table>

Note. Estimates based on district-level panel data spanning the years 2005–2015. The dependent variables are the infant mortality rate (IMR) per 1,000 infants in columns 1 and 3, and under-five mortality rate (U5MR) per 1,000 children under five years old in columns 2 and 4. In Panel A, columns 1 and 2 show OLS estimates following equation (1). Columns 3 and 4 in Panel A and columns 1–4 in Panel B show 2SLS estimates of the effect of “implemented projects” instrumented by “predicted projects” using equation (4). The sample of analysis is restricted to years prior to the completion of at least one sewerage project in a given district. The samples columns 1 and 2 in Panel B are further restricted to district-year observations in which no project is halted, and in columns 3 and 4 to district-year observations in which all projects are halted or no project has been started. All regressions include district and year fixed effects. Standard errors clustered by district are reported in parentheses and p-values in brackets. The table also reports the weak-instrument-robust Anderson–Rubin (AR) p-value, the Sanderson–Windmeijer (SW) F-statistic and the mean of each outcome in the initial year of the study (2005). See Online Appendix D for variable definitions and see the text for further details.
Conclusions

- An additional implemented project ↑ early-life mortality, with respect to not starting projects
  - ↑ IMR by 0.003 deaths (17%)
  - ↑ U5MR by 0.66 deaths (14%)

- Effect on mortality of an additional halted project is 3x as large as the effect of an additional project that is just underway.

- Ignoring social costs during project implementation overestimates welfare calculations.
Conclusions

• An additional implemented project ↑ early-life mortality, with respect to not starting projects
  • ↑ IMR by 0.003 deaths (17%)
  • ↑ U5MR by 0.66 deaths (14%)

• Effect on mortality of an additional halted project is 3x as large as the effect of an additional project that is just underway.

• Ignoring social costs during project implementation overestimates welfare calculations.
Conclusions

• An additional implemented project ↑ early-life mortality, with respect to not starting projects
  • ↑ IMR by 0.003 deaths (17%)
  • ↑ U5MR by 0.66 deaths (14%)

• Effect on mortality of an additional halted project is 3x as large as the effect of an additional project that is just underway.

• Ignoring social costs during project implementation overestimates welfare calculations.
Conclusions

- Benefit-cost ratio using estimated effects of completed projects [Alsan and Goldin, 2019; Galiani et al., 2005]
  - Social benefits of completed projects are $7 \times$ the estimated social cost of projects implemented without problems.
  - Social benefits just $7/10$ of the social costs associated with delayed and halted projects.
  - No social benefits if projects are abandoned.

- Abandoning projects: also wasteful use of public resources
  - By 2015, halted and unfinished projects had 40% of contractual sum disbursed.
  - Total waste $\approx 1/3$ of public expenditure on tertiary education in 2015 in Peru [World Bank, 2020].
Conclusions

- Benefit-cost ratio using estimated effects of completed projects [Alsan and Goldin, 2019; Galiani et al., 2005]
  - Social benefits of completed projects are 7x the estimated social cost of projects implemented without problems.
  - Social benefits just 7/10 of the social costs associated with delayed and halted projects.
  - No social benefits if projects are abandoned.

- Abandoning projects: also wasteful use of public resources
  - By 2015, halted and unfinished projects had 40% of contractual sum disbursed.
  - Total waste $\approx$ 1/3 of public expenditure on tertiary education in 2015 in Peru [World Bank, 2020].
Thank you!
Implementing agent

- District: 56.2%
- Province: 30.7%
- Region: 7.5%
- Ministry of Sanitation: 5.4%
- Other sectors: 0.2%

Can White Elephants Kill?
Antonella Bancalari (St.Andrews and IFS)
NBER SI, Urban Economics, 2022
Funding source

- Royalties: 38.9%
- Transfers: 29.2%
- Taxes: 23.8%
- Funds: 7.2%
- Loans: 1.0%

Nationwide funds for sewerage varies over time

Budget (million soles)

year


Can White Elephants Kill?

Antonella Bancalari (St.Andrews and IFS)
Driven by exogenous shock to royalties

Placebo test, reduced form and first stage

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Placebo test</th>
<th>Reduced-form</th>
<th>1st stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>IMR Deaths per 1,000 infants or children</td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>Predicted projects</td>
<td>0.000</td>
<td>0.071</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.064)</td>
<td>(0.000)</td>
</tr>
<tr>
<td></td>
<td>[0.156]</td>
<td>[0.266]</td>
<td>[0.034]</td>
</tr>
<tr>
<td>Mean (initial)</td>
<td>0.018</td>
<td>4.818</td>
<td>0.018</td>
</tr>
<tr>
<td>District-year</td>
<td>5,630</td>
<td>5,630</td>
<td>8,555</td>
</tr>
<tr>
<td>Districts</td>
<td>1,283</td>
<td>1,283</td>
<td>1,379</td>
</tr>
</tbody>
</table>

Notes. Clustered standard errors at the district level in parenthesis and p-values in brackets. District and year fixed effects included.
**Alternative mechanism: demographic changes**

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Fertility</th>
<th>Migration</th>
<th>Selective</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Infants (1)</td>
<td>Under-5 (2)</td>
<td>Population (3)</td>
</tr>
<tr>
<td>Implemented projects</td>
<td>0.173 (12.510)</td>
<td>0.864 (62.551)</td>
<td>3513.991 (1210.377)</td>
</tr>
<tr>
<td>Anderson-Rubin $p$-value</td>
<td>0.989 [0.989]</td>
<td>0.989 [0.989]</td>
<td>0.000 [0.004]</td>
</tr>
<tr>
<td>Mean (initial)</td>
<td>495.913</td>
<td>2479.565</td>
<td>23472.257</td>
</tr>
<tr>
<td>F-stat(SW)</td>
<td>15.861</td>
<td>15.861</td>
<td>15.861</td>
</tr>
<tr>
<td>District-year</td>
<td>8555</td>
<td>8555</td>
<td>8555</td>
</tr>
<tr>
<td>Districts</td>
<td>1379</td>
<td>1379</td>
<td>1379</td>
</tr>
</tbody>
</table>

Note. Estimates based on district-level panel data spanning the years 2005–2015. The dependent variables are the number of infants (‘Infants’, in column 1); under-five population (‘Under-5’, in column 2); total population (‘Population’, column 3); population density in km$^2$ (‘Density’, column 4); share of household heads with secondary education completed (‘Education’, column 5) and share of households connected to electricity grids (‘Electricity’, column 6) in a given district. Data for the outcome in column (5) is only available for the years 2005, 2010 and 2017, and for the outcome in column (6) only for the years 2005 and 2017 and for fewer districts out of the total sample. For these two outcomes, data from 2017 is used to impute the year 2015 and the missing years are imputed with the latest value available. Coefficients correspond to 2SLS estimates of the effect of ‘implemented projects’ instrumented by ‘predicted projects’ using Equation ?? . The sample of analysis is restricted to years prior to the completion of at least one sewerage project. All regressions include district and year fixed effects. Standard errors clustered by district are reported in parentheses and $p$-values in brackets. The table also reports the weak-instrument-robust Anderson-Rubin (AR) $p$-values, the Sanderson–Windmeijer (SW) $F$-statistic and the mean of each outcome in the initial year of the study (2005). See Appendix?? for variable definitions and the text for further details.
Project pathway

- Completed: 23%
- Not completed: 2.2%
- Halted: 20.6%
- Restarted and completed: 10.1%
- Restarted, but not completed: 44.1%