A Appendix

A.1 Data

We use India's repeated cross-sectional data from the Demographic Health Survey (DHS) to empirically motivate the key features of our theoretical framework. All four waves of this qualitatively rich survey for the years 1992–93, 1998–2000, 2005–06, and 2015–16 have been utilized for the background checks and trend analysis. The full sample of four rounds of the DHS data contains 720,470 households with a total of 1,000,228 respondents or ever-married women under 49 years of age by the survey date from 38,436 villages/towns (or primary sampling units – PSUs) in all 36 states and union territories (UTs). The background checks based on trend analysis generally use a full sample, i.e., as many observations as possible. However, the K-modes clustering was performed based on a randomly drawn subsample, 0.1% of India's DHS-4 (2015–16) data, only in which indoor smoking behavior is recorded. We use a small portion of the data (84 observations) but still nationally representative for a clear illustration purpose.

Table A.1 provides detailed descriptions of the variables used in our analysis. It is important to note that the fertility measure is defined in two different ways. First, the fertility rate used in background checks and trend analysis is the total fertility rate (TFR). Second, we employ a dummy variable of fertility decisions for our clustering analysis, that is defined as whether a woman had a child over the period of 30 months before the date of the survey.

Table A.1: Variable Descriptions

Variable	Definition
(a) Instances of Household Public Goods, mother (j) × year of survey (t)	
Indoor air quality	Proxied by a dummy variable indicating whether household fuel choice or clean energy access. A total of ten different types of fuel are reported in India's DHS datasets as the main cooking fuels, and we reclassify these self-reported fuels used in the household into a dummy. 1 if a household in which the mother lives uses one of the clean fuels for cooking, 0 otherwise. Electricity, liquid petroleum gas (LPG) or natural gas, and biogas are counted as clean fuels, whereas kerosene, coal or lignite, charcoal, straw/shrubs/grass, wood, crop waste, and dung are considered as dirty fuels.
Indoor smoking behavior	A dummy variable indicating whether a household in which the mother lives has someone smoking inside the house. 1 if family members never smoked indoor, 0 otherwise.
Child vaccination	A dummy variable indicating whether a mother ever treated at least one of her under-five children with vaccination. 1 if a child ever had vaccination, 0 otherwise.
(b) Fertility Measures, mother (j) × year of survey (t)	
Fertility decision	A dummy variable of whether a woman had a child over the period of 30 months before the survey. 1 if a woman had a child within 30 months before the survey date, 0 otherwise.
Fertility rate, country (i) × year of survey (t)	The total fertility rate (TFR) – the sum of the age-specific fertility rate (or fertility rate for the seven five-year age groups from 15-19 to 45-49) for all eligible women in the household multiplied by five. When calculating the weighted estimates of the total fertility rates, we use a national women's sample weight to adjust the cluster sampling survey design.
(c) Child Mortality Measure, mother (j) × year of survey (t)	
Child mortality incidence	A dummy variable indicating whether a mother experienced mortality of a child born over the period 5 years (or 60 months) before the survey. 1 if a child died in the first 30 months of life, 0 otherwise. Twins are excluded.

Notes: The unit of observation is mother (j) × year of survey (t) unless otherwise specified.

A.2 *K*-Modes Clustering

To define the number of equilibria and explore the relationships between variables of interest, we use a simple machine learning technique that is an unsupervised clustering algorithm. Particularly, we use the K-modes clustering algorithm, and its paradigm is described here. The K-modes clustering algorithm is an extension of the K-means clustering algorithm (MacQueen et al., 1967) and was developed by Huang (1997, 1998) to cluster categorical data by dealing with modes instead of means.

The K-modes clustering method classifies the objects given by the data into K groups such that the distance from objects to the assigned cluster modes is minimized. The dissimilarity of two objects is determined by simple matching distance measure[14] for

[14]The simple matching distance also can be seen as the Hamming distance, which determines the dissimilarity between two string objects of equal length by counting the distinctive characteristics of the two

categorical objects, which counts the number of mismatches in all variables, and it is the first difference from the K-means clustering method that uses Euclidean distance function for continuous numeric values. Secondly, the K-modes method replaces mean of a cluster with the mode to compute the center of a cluster.

The steps of K-modes algorithm are then the same as those of K-means algorithm, except for the cost function it minimizes. Suppose $\boldsymbol{X} = [X_1 \cdots X_m \cdots X_M]_{(N \times M)}$ be a vector of categorical objects of interest or data vectors, where M is the number of objects or variables. The m-th and j-th objects can be written as $X_m = [x_{1m} \cdots x_{im} \cdots x_{Nm}]$ and $X_j = [x_{1j} \cdots x_{ij} \cdots x_{Nj}]$, where N is the total number of observations. It is considered that $X_m = X_j$ if and only if $x_{im} = x_{ij}$ for all $1 \le m, j \le M$ and $1 \le i \le N$, where m and j are indexes of objects and i is an index of observations.

The simple matching distance function of X_m and X_j represented above is defined as

$$d(X_m, X_j) = \sum_{i=1}^{N} \gamma(x_{im}, x_{ij}), \tag{B.1}$$

where

$$\gamma(x_{im}, x_{ij}) = \begin{cases} 0, & \text{if } x_{im} = x_{ij} \\ 1, & \text{if } x_{im} \neq x_{ij}. \end{cases} \tag{B.2}$$

The objective of clustering a vector of M categorical objects into K clusters is to find Y and S such that minimize the cost function $F(\cdot)$

$$F(S, Y) = \sum_{k=1}^{K} \sum_{m=1}^{M} s_{km} d(Y_k, X_m), \tag{B.3}$$

subject to

$$s_{km} \in \{0, 1\}, \quad \sum_{k=1}^{K} s_{km} = 1, \quad \text{and} \quad 0 < \sum_{m=1}^{M} s_{km} < M, \tag{B.4}$$

for all $1 \le k \le K$ and $1 \le m \le M$, where K is a predetermined number of clusters ($K \le M$), S is a $K \times M$ matrix of $s_{km} \in \{0, 1\}$, $Y = [Y_1 \cdots Y_k \cdots Y_K]$, and Y_k is center (e.g., mode) of the k-th cluster. In terms of optimization procedures, we first minimize the objective function $F(\cdot)$ with respect to S given Y, and then in the second step minimize $F(\cdot)$ with respect to Y by fixing the S at the level determined in the first step. We perform the K-modes clustering using `Rstudio` package `kmodes`.

---

categorical objects.

A.3 Detailed Solution of the Model

In this section, we provide a solution to our model in greater detail. The consumer's utility maximization problem is defined as

$$\begin{aligned}
\max_{\{k_t,n_t\}_{t=0}^{\infty}} \quad & U_t = \log(c_t - \bar{c}) + \beta \log(c_{t+1} - \bar{c}) \\
\text{s.t.} \quad & y_0(1 + k_{t-1}) = c_t + vn_t + k_t + \sigma y_0(1 + k_{t-1}), \\
& \pi_{t+1} n_t \sigma y_{t+1} = c_{t+1}, \\
& 0 \le k_t \le \bar{k} \text{ and } c_t, n_t, k_{t-1} \ge 0 \text{ for all } t, \\
& y_0, \pi_0 \text{ given.}
\end{aligned}$$

To solve this maximization problem, we first find c_t and c_{t+1} from the two budget constraints and then plug them into the utility function. For the moment, we intentionally ignore non-negativity constraints and upper-bound constraint on investment in household public goods. Then we have the following unconstrained utility maximization problem:

$$\max_{\{k_t,n_t\}_{t=0}^{\infty}} U_t = \log\left[y_0(1-\sigma)(1+k_{t-1}) - vn_t - k_t - \bar{c}\right] + \beta \log\left(\pi_{t+1} n_t \sigma y_{t+1} - \bar{c}\right), \tag{B.5}$$

where the first term corresponds to utility during adulthood, while the utility during old age is captured by the second term. The first order condition with respect to fertility level, n_t, is given by

$$\frac{v}{y_0(1-\sigma)(1+k_{t-1}) - vn_t - k_t - \bar{c}} = \frac{\beta \pi_{t+1} \sigma y_{t+1}}{\pi_{t+1} n_t \sigma y_{t+1} - \bar{c}}.$$

After few manipulations, we can define the utility maximizing fertility level, n_t^*, as

$$n_t^* = \frac{\beta}{(1+\beta)v}\left[y_0(1-\sigma)(1+k_{t-1}) - k_t - \bar{c}\left(1 - \frac{v}{\beta \pi_{t+1}\sigma y_{t+1}}\right)\right].$$

To substitute the utility-maximizing fertility level, n_t^*, back into the utility function in (B.5), let us do this for adulthood and old age separately for notational economy and then we combine those two expressions together. First, utility during the adulthood is

$$\log\left[y_0(1-\sigma)(1+k_{t-1}) - vn_t^* - k_t - \bar{c}\right] = \log\left(\frac{1}{1+\beta}\right) + \\ + \log\left[y_0(1-\sigma)(1+k_{t-1}) - k_t - \bar{c}\left(1 + \frac{v}{\pi_{t+1}\sigma y_{t+1}}\right)\right].$$

Second, utility during the old age period is

$$\beta \log\left(\pi_{t+1} n_t^* \sigma y_{t+1} - \bar{c}\right) = \beta \log\left[\frac{\beta}{(1+\beta)v}\right] + \beta \log(\pi_{t+1}\sigma y_{t+1}) +$$
$$+ \beta \log\left[y_0(1-\sigma)(1+k_{t-1}) - k_t - \bar{c}\left(1 + \frac{v}{\pi_{t+1}\sigma y_{t+1}}\right)\right].$$

Combining these two expressions, we can write the maximal discounted utility as

$$U_t(k_t, k_{t-1}) = (1+\beta)\log\left[y_0(1-\sigma)(1+k_{t-1}) - k_t - \bar{c}\left(1 + \frac{v}{\pi_{t+1}\sigma y_{t+1}}\right)\right] +$$
$$+ \beta \log\left(\pi_{t+1}\sigma y_{t+1}\right) + \log\left(\frac{1}{1+\beta}\right) + \beta \log\left[\frac{\beta}{(1+\beta)v}\right].$$

Now we maximize this maximum $U_t(k_t, k_{t-1})$ by choice of k_t, and the first order condition (FOC) with respect to k_t is given by

$$\frac{(1+\beta)\left[1 - \frac{2v\bar{c}}{\pi_0 y_0 \sigma (1+k_t)^3}\right]}{y_0(1-\sigma)(1+k_{t-1}) - k_t - \bar{c}\left[1 + \frac{v}{\pi_0 y_0 \sigma (1+k_t)^2}\right]} = \frac{2\beta}{1+k_t}.$$

By rearranging and collecting the terms, we obtain

$$2\beta y_0(1-\sigma)k_{t-1} = (1+3\beta)k_t - \left(\frac{2v\bar{c}}{\pi_0 y_0 \sigma}\right)\frac{1}{(1+k_t)^2} - 2\beta y_0(1-\sigma) + 2\beta\bar{c} + \beta + 1,$$

and if we divide the both sides of above expression by $2\beta y_0(1-\sigma)$ where we assume that $\sigma \neq 1$, we define the following law of motion equation which demonstrates the dynamic relationship between k_t and k_{t-1}

$$\begin{aligned} k_{t-1} &= \left[\frac{(1+3\beta)}{2\beta(1-\sigma)y_0}\right]k_t - \left[\frac{1}{\beta(1-\sigma)\pi_0\sigma}\right]\left(\frac{v}{y_0}\right)\left(\frac{\bar{c}}{y_0}\right)\frac{1}{(1+k_t)^2} + \\ &\quad + \frac{1}{1-\sigma}\left(\frac{\bar{c}}{y_0}\right) + \left[\frac{1+\beta}{2\beta(1-\sigma)}\right]\frac{1}{y_0} - 1, \\ &= \alpha k_t - \frac{\hat{k}_0}{(1+k_t)^2} + \tilde{k}_l, \\ &= g(k_t). \end{aligned}$$