Multi-Section Article Estimate Sandbox

November 7, 2019

1 Estimation

1.1 Estimation One

Besides the age at school closure, the impact of school closure on educational attainment may also differ by the number of years of exposure to the policy: short-run effects of closure on a child's educational attainment progression could be dampened or amplified over the medium and long run.[1] In order to identify both age and duration effects with our cross-sectional data, we exploit the variation in the year of school closure. Under the assumption that the impact of the policy is not specific to the calendar year of closure, we can estimate Equation 1 to obtain the impact of the policy as a function of both starting age and the length of exposure.

In Equation 1, we use similar notations as in Equation **??**, the difference is that the policy's effects are now captured by $\hat{\lambda_{zl}}$ that varies by age-at-closure variable t_i and years-of-exposure variable τ_i:

$$\begin{aligned} E_{pvia} &= \phi + \beta_v + \rho_{pa} \\ &+\hat{\lambda_{zl}} \cdot \mathbb{1}\{(l_l \leqslant \tau_i \leqslant u_l) \cap (l_z \leqslant t_i \leqslant u_z)\} \cdot c_v \\ &+X_i \cdot \gamma + \epsilon_i \end{aligned} \tag{1}$$

where, as before, c_v is a binary variable indicating if individual i is from a village v with school consolidation (i.e. treatment village). As in Equation **??**, we group children in villages with school closure into Z groups based on their age at closure, with lower and upper bounds for each group, l_z and u_z. To capture duration effects, we further divide each of the Z groups of children into L groups based on the length of exposure τ_i, defined as the gap between individual i's age in 2011 and i's age at year of school closure, t_i.[2] Each l length of exposure group includes those with τ_i falling within lower and upper bounds, l_l and l_u. The exposure groups allow us to separately estimate the short, medium and long run effects of the consolidation policy on educational attainment. There are $Z \cdot L$ groups of interest for this regression.[3]

1. After an individual completes schooling, duration effects will become constant. In studies with cross-sectional data taken long after a policy has been implemented, **duflo_school_2001** for example, the duration effect is irrelevant because all educational attainment data is observed long after sample individuals have completed schooling. In our data, a significant proportion of individuals have not completed schooling, allowing us to have meaningful duration effects.

2. $\tau_i = \min(a_i, a_i - t_i)$: τ_i is the gap between age in 2011 and t_i if individual i was borne before the year of closure, and it is the age of the child in 2011 if the child was borne after school closure.

3. Ideally, we would estimate the policy effects for each t_i and τ_i combination separately, but we

1.2 Estimation Two

We include in the model measurement errors that allow us to estimate the parameters using maximum likelohood methods. As described previously, households observe $\Omega = \left(Y, p^N_{yv}, X\right)$, and the distributions of R_{yv}. In terms of choices and outcomes, the econometrician only observes F^* and N^*, which differ from the true optimal nutritional choice N by measurement error η and true height outcome h_{24} by ι:

$$\log(N^*) = \log(N(Y, X, \epsilon; p^N_{yv}, \mu_{R_{yv}}, \sigma_{R_{yv}})) + \eta \tag{2}$$

$$\log(h^*_{24}) = \log(h_{24}(N(Y, X, \epsilon; p^N_{yv}, \mu_{R_{yv}}, \sigma_{R_{yv}}), X, \epsilon)) + \iota \tag{3}$$

We assume that η and ι are normally distributed, and that ϵ, η and ι are independent. The standard deviation of η is σ_η and the mean is $\mu_\eta = -\frac{\sigma^2_\eta}{2}$. The standard deviation for ι is σ_ι with mean $\mu_\iota = -\frac{\sigma^2_\iota}{2}$. The log likelihood is based on the difference between model optimal nutritional choices and observed nutritional choices, as well as the model height outcome and observed heights at 24 months of age:

$$\max_{\theta\in\Theta} \sum_{y=1970}^{1975} \sum_{v} \left\{ \sum_{i=1}^{n_{yv}} \log\left(\left(\int_\epsilon \phi_\iota\left(\ln h^*_{24,i} - \ln h_{24}\left(\begin{smallmatrix} Y_i, X_i, \epsilon_i; \\ \theta, \mu_{R_{yv}}, \sigma_{R_{yv}} \end{smallmatrix}\right)\right) \cdot \phi_\eta\left(\ln N^*_i - \ln N\left(\begin{smallmatrix} Y_i, X_i, \epsilon_i; \\ \theta, \mu_{R_{yv}}, \sigma_{R_{yv}} \end{smallmatrix}\right)\right) dF(\epsilon_i)\right)\right)\right\} \tag{4}$$

where

$$\theta = \{\underbrace{\rho, \gamma, \lambda}_{\text{Preference}}, \overbrace{\delta}^{\substack{\text{Atole}\\\text{disc.}}}, \underbrace{A, \alpha, \beta, \sigma_\epsilon}_{\substack{\text{Production}\\\text{Function}}}, \sigma_\eta, \sigma_\iota\} \tag{5}$$

Equation 4 is determined by θ as well as a set of $\left(\mu_{R_{yv}}, \sigma_{R_{yv}}\right)$ that are village- and time-specific. This means that in estimating the model, we do not impose assumptions about where the current height distribution is with respect to the stationary height distribution. We solve for optimal choices given the observed individual specific Ω_i and the observed $\mu_{R_{yv}}, \sigma_{R_{yv}}$ for each year y in village v.

1.3 Editing

1. ☒ comment one
2. ☒ comment two

have constructed the Z and L groups due to limited sample size.

Table 1: Summary Statistics Various Variables

	Atole and Fresco Differences				
	All	Group Averages		p-Values Testing	
	mean (sd)	*Fresco Villages*	*Atole Villages*	*Gap*	*P-value*
Panel A: Gender Income Price (N=503, main sample)					
Male	0.52	0.52	0.52	-0.00	0.92
	(0.50)	(0.50)	(0.50)		
Income (peso)	515.57	503.68	526.00	22.32	0.59
	(460.9)	(464.4)	(458.4)		
Mth 15-24 Protein Price (peso/10k grams)	52.58	52.47	52.68	0.21	0.54
	(3.87)	(3.93)	(3.81)		
Panel B: Gender Income (N=1115, hgt observed once in first 24 months)					
Male	0.53	0.53	0.53	0.00	0.98
	(0.50)	(0.50)	(0.50)		
Income (peso)	449.49	444.63	454.06	9.43	0.72
	(432.3)	(446.4)	(419.0)		
Panel C: Height					
Month 0 (cm) N=503	49.64	49.79	49.52	-0.27	0.19
	(2.29)	(2.29)	(2.29)		
Month 6 (cm) N=463	62.72	62.49	62.93	0.44	0.05
	(2.46)	(2.50)	(2.42)		
Month 12 (cm) N=475	68.81	68.45	69.13	0.68	0.01
	(2.99)	(3.13)	(2.83)		
Month 18 (cm) N=482	73.37	72.88	73.80	0.92	0.00
	(3.23)	(3.26)	(3.15)		
Month 24 (cm) N=503	77.66	76.94	78.29	1.36	0.00
	(3.47)	(3.49)	(3.33)		
Panel D: Average Daily Nutritional Intake					
Month 15 (grams/day) N=464	17.43	14.29	20.07	5.78	0.00
	(10.5)	(9.14)	(10.9)		
Month 18 (grams/day) N=461	21.52	18.27	24.41	6.14	0.00
	(11.4)	(9.61)	(12.0)		
Month 21 (grams/day) N=475	24.45	20.17	27.99	7.82	0.00
	(11.4)	(9.03)	(11.9)		
Month 24 (grams/day) N=462	26.99	22.51	31.07	8.56	0.00
	(12.0)	(9.02)	(13.0)		
Avg Mth 15-24 (grams/day) N=503	22.54	18.81	25.80	6.99	0.00
	(8.98)	(6.43)	(9.62)		
Avg Mth 15-24 (kcal/day) N=503	691.78	681.78	700.55	18.77	0.38
	(236.5)	(236.2)	(236.9)		

Summary Statistics