

You are What Your Parents Expect: Height and Local Reference Points*

Fan Wang

March 17, 2022

Abstract

Recent estimates are that about 150 million children under five years of age are stunted, with substantial negative consequences for their schooling, cognitive skills, health, and economic productivity. Therefore, understanding what determines such growth retardation is significant for designing public policies that aim to address this issue. We build a model for nutritional choices and health with reference-dependent preferences. Parents care about the health of their children relative to some reference population. In our empirical model, we use height as the health outcome that parents target. Reference height is an equilibrium object determined by earlier cohorts' parents' nutritional choices in the same village. We explore the exogenous variation in reference height produced by a protein-supplementation experiment in Guatemala to estimate our model's parameters. We use our model to decompose the impact of the protein intervention on height into price and reference-point effects. We find that the changes in reference points account for 65% of the height difference between two-year-old children in experimental and control villages in the sixth annual cohort born after the initiation of the intervention.

JEL: I15, D8, D9, O15

Keywords: Early Childhood, Height, Reference Points, Nutrition, Anthropometrics

*Fan Wang: Department of Economics, University of Houston, Houston, Texas, USA (email: fwang26@uh.edu)
We thank Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like “Huardest gefburn”? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language.

1 Introduction

Insufficient height and weight growth still affect many children around the globe. Estimates are that about 150 million (22 percent) of children under five years of age are stunted (FAO 2019).¹ Studies suggest that these children are at risk of not developing their full human capital potential (Behrman et al. 2009; Black et al. 2017; Hoddinott, Alderman, et al. 2013; Hoddinott, Behrman, et al. 2013; Hoddinott et al. 2008; Maluccio et al. 2009; Richter et al. 2017; Victora et al. 2008). Our paper contributes to a growing literature that investigates how interventions in early childhood can contribute to foster human-capital formation for at-risk children (e.g., Campbell et al. 2014; Carneiro and Heckman 2004; Cunha and Heckman 2007, 2009; Cunha et al. 2010; Gertler et al. 2014; Heckman et al. 2010a, 2010b; Heckman et al. 2013).

2 The Model

Following our discussions on parental beliefs, given $\Omega_{y,v} = (Y, p_{yv}^N, \delta, X, \varepsilon, \mu_{R_{y,v}}, \sigma_{R_{y,v}}^2)$, each household solves the following maximization problem:

$$\max_{C,N} \left\{ C + \rho \cdot C^2 + \left\{ \gamma \cdot H^{24} + \lambda \cdot \int_{R_{yv}} (H^{24} - R_{yv}) \mathbb{1} \{ H^{24} \geq R_{yv} \} \phi(R_{yv}; \mu_{R_{yv}}, \sigma_{R_{yv}}^2) dR_{yv} \right\} \right\} \quad (1)$$

subject to the budget constraint (??), the production function (??), the equation that determines mean (??), and some measure of variance beliefs $\sigma_{R_{y,v}}^2$. Let $N(\Omega_{y,v})$ denote the policy function for nutrition. If $\sigma_{R_{y,v}}^2 = 0$, preferences for health would be piecewise linear with a kink at $\mu_{R_{y,v}}$. If $\sigma_{R_{y,v}}^2 > 0$, preferences for health are continuously differentiable; if additionally $\gamma > 0$ and $\lambda < 0$, preferences for health are concave.

The optimization problem in Equation (1) does not permit analytical solutions.² To illustrate how changes in key parameters impact optimal choices, we present in Figure 1 the

1. Stunted children have height-for-age more than two standard deviations below the median for a well-nourished population. In 2018, 59 million (30 percent) and 82 million (23 percent) of African and (non-Japanese) Asian children under five years of age were stunted, respectively (FAO 2019).

2. In Appendix Section A.1, we discuss the first-order condition for the optimal nutritional choice problem. In Appendix Section ??, we describe the procedures for numerical solutions.

consumption-health possibility frontier along with several indifference curves for an individual, using estimated parameters from our model.³ In each panel, solid dots indicate optimal choices (i.e., the combination of height and household consumption where indifference curves are tangent to the consumption-health possibility frontier).

Panel a shows several indifference curves and the consumption-health possibility frontier. The figure clearly shows the asymmetry in indifference curves. To the left of the mean reference height, parents are willing to sacrifice a large amount of consumption for a small increase in the child's height. In contrast, to the right of the mean reference height, parents are willing to forego only a small amount of consumption for a large increment in height.

In panel b, we vary the λ parameter. Marginal benefits of additional nutritional intakes on expected utility are increasing in λ . Additionally, reference points matter more when λ deviates more from zero. Visually, at more negative values of λ , the curvature of the indifference curve increases. The greater the curvature of the indifference curve, the more critical the role of reference height in determining parental nutritional choices and, consequently, the child's height at age 24 months.

In panel b, we fix λ and vary $\mu_{R_{y,v}}$. When $\lambda < 0$, the marginal benefits of additional nutritional intakes are increasing in $\mu_{R_{y,v}}$. The greater the value of $\mu_{R_{y,v}}$, the lower the household consumption and the taller the child at age 24 months.

Finally, panel d of Figure 1 presents the effect of increasing uncertainty about the parental estimates of mean height, $\sigma_{R_{y,v}}^2$. Given the estimated parameters from the empirical model, higher values of $\sigma_{R_{y,v}}^2$ increase the marginal benefits of additional nutrition when expected height exceeds $\mu_{R_{y,v}}$ and reduces the curvature of the indifference curves. These lead to an increase in nutritional choices. In the limit, as $\sigma_{R_{y,v}}^2$ approaches infinity, preferences become linear in height, and predictions from models with and without reference points become empirically indistinguishable.

We use the estimated parameters to produce Figure 1. However, in general, the household response to an increase in uncertainty depends on the values of γ and λ . When marginal utility gains from additional investment in health are positive after the mean reference point μ_R – which

3. The consumption-health possibility frontier is jointly determined by the household budget and the child height production function.

means $-\gamma < \lambda < 0$ – an increase in σ_R will lead to an increase in investments in health. In contrast, if marginal utility from additional investment in health is negative after the mean reference point μ_R – which generates backward bending indifference curves – a similar increase of σ_R could reduce investment in health. *Ceteris paribus*, there is a threshold level of λ where increases in σ_R have no impact on health investments. In Appendix Section A.1, we discuss the intuition behind these results and provide graphical illustrations.

3 Data

The data we use in this paper comes from an experimental intervention conducted by The Institute of Nutrition of Central America and Panama (INCAP), which started a nutritional-supplementation trial in 1969. Four villages from eastern Guatemala were selected, one pair of villages that was relatively populous (~900 residents each) and one pair that was less populous (~500 residents each). The villages were similar in child nutritional status, measured as the height at three years of age (Habicht et al. 1995). Over 50% of children lacked proper nutrition and were severely stunted, measured as height-for-age z-scores less than -3.⁴ The intervention consisted of randomly assigning nutritional supplements. One large and one small village were selected to receive a high-protein drink called Atole, and the other two were selected to receive an alternative supplement called Fresco. Each serving of Atole (180 ml) contained 11.5 grams of protein and 163 kcal. Fresco had no proteins, and each serving (180 ml) had 59 kcal. The central hypothesis was that the protein supplementation would accelerate physical and mental development (Habicht et al. 1995). The intervention started in February 1969 in the larger villages and in May 1969 in the smaller villages and lasted until the end of February 1977, with data collection taking place until September 1977 (Islam and Hoddinott 2009; Maluccio et al. 2009). The nutritional supplements were distributed in feeding centers located centrally in each village. The centers were open twice a day, two to three hours in the mid-morning and two to three hours in the mid-afternoon. All village members had access to the supplements at the

4. Guatemalan children continue to suffer from severe malnutrition. In 2015, among Guatemalan households in the lowest quintile of wealth, approximately 70 percent of children younger than five were stunted. In middle-quintile Guatemalan households, 45 percent of children younger than five were stunted (FAO 2019).

feeding centers.

Table 1 presents summary statistics for the variables that we use in our analysis. In Panels a and b, we show statistics on gender, income, and prices for our main sample of 503 individuals (panel a) and gender and income for the fuller sample of 1155 individuals (panel b). In Panels c and d, we show statistics on heights and nutritional intakes, respectively. Table 1 has five columns. The first column presents the overall means and standard deviations in Atole and Fresco villages combined. The second and third columns present Atole and Fresco village-specific means and standard deviations. Column four shows the gaps in means between Atole and Fresco villages for each variable, and column five presents the p-values for the statistical significance of these gaps.

As mentioned before, the intervention took place in four villages, two Atole or treatment villages and two Fresco or control villages. In the rest of the paper, when we refer to Atole and Fresco villages, we merge the information of the two villages that received the same supplement. The limited number of villages might impact the descriptive statistics' standard errors since villages can share common unobserved shocks. We follow the methods developed by Donald and Lang (2007) and Cameron and Miller (2015) to study how robust our results are to this clustering. The method proposed by Donald and Lang (2007) consists of estimating averages by clusters, controlling for individual variables, and using those averages in the regressions. This method greatly reduces the number of observations. We define cluster year-village pairs and half-year-village pairs to implement this procedure. Following Cameron and Miller (2015) we also implement a pair-cluster bootstrap, using the same cluster definition as in the Donald and Lang (2007) method. Table 1 and Figure 2 report the results without using the clusters corrections, but the results are robust to those methods.

4 Results

We present estimated parameters in Table 2, with standard errors shown in parentheses. For preferences, ρ is -0.047, indicating that preferences are concave in c . γ and λ are 0.033 and -0.026, respectively. Therefore, preferences are also concave in height, and the curvature of

the indifference curves indicates the asymmetry in responses. Parents prefer taller to shorter children, but the marginal benefit from an additional centimeter of height beyond the reference comparison height is close to zero (albeit positive).

The price discount parameter δ is 0.376, representing a 38% discount in protein prices in Atole villages. The production-function parameters are $\beta = 0.073$, $\alpha_{H_0} = 0.022$, $\alpha_{male} = 0.009$, $A = 4.144$, and $\sigma_\varepsilon = 0.010$. Given these parameters, we show the consumption- and height-possibility frontier along with indifference curves for an individual in panel a of Figure 1.

The measurement-error estimates are $\sigma_\eta = 0.382$ and $\sigma_l = 0.043$. These findings indicate that measurement error is more severe for nutritional intake than height, which is intuitive.

We use the estimated model to decompose the relative contributions of prices and reference points to the height dynamics in Atole and Fresco villages. In Appendix ??, we discuss additional counterfactuals in which we compare the relative impacts of poor-targeted and universal policy experiments given reference points.

5 Conclusion

In this paper, we build and estimate a child-nutritional investment model. The model considers reference-dependent preferences, where the reference is with respect to the heights of the previous cohort of children who live in the same village. Reference points shift endogenously as households observe children’s heights from earlier birth cohorts changing.

For researchers interested in the impact of price subsidies and income transfers, we have introduced a long-term secondary channel – endogenous changes in reference points – that might affect the impacts of these policies. For the protein-supplement experiment implemented in Guatemala, which we interpret as a price-discount policy, by 1975 – six years after the start of the policy – 57% to 65% of the policy’s impact are due to its impact on shifting reference points.

Our paper also shows significant height increases might be realized from shifting reference points for highly-stunted populations. It is an open question how to exogenously shift these reference points in the short run, although the Peruvian experience mentioned in the introduction indicates that educational campaigns could be effective on a large scale over time. The cost of

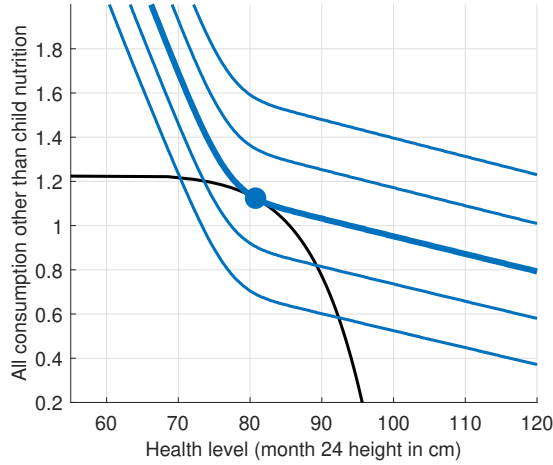
an educational campaign to inform households about alternative reference heights well might be lower compared to income transfers and price subsidies with similar effects on heights.

We recognize that our analysis requires a crucial assumption: parents use the data on a selected group of children to estimate mean and variance beliefs about the reference height. We argued that our assumption is consistent with research in economics, medicine, and anthropology, but we recognize that there are many other alternatives, such as rational expectations, Bayesian updating, or social learning models. Unfortunately, the literature in economics knows very little about such a critical element of the model we propose in this paper. Therefore, it is necessary to encourage research that elicits parental subjective distributions of reference points as part of parent-directed interventions to reduce stunting. Within a randomized controlled trial, the availability of such data would shed light on the process parents use to update critical moments of the subjective distribution. With such evidence, one could have a more robust representation of our model that would better inform the design of public policies to improve children's health.

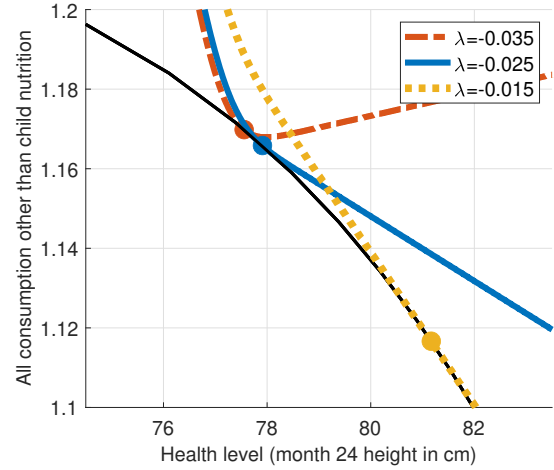
References

- Behrman, J. R., M. C. Calderon, S. H. Preston, J. Hoddinott, R. Martorell, and A. D. Stein. 2009. "Nutritional Supplementation in Girls Influences the Growth of Their Children: Prospective Study in Guatemala." *The American Journal of Clinical Nutrition* 90, no. 5 (November): 1372–1379. <https://doi.org/10.3945/ajcn.2009.27524>.
- Black, M. M., S. P. Walker, L. C. Fernald, C. T. Andersen, A. M. DiGirolamo, C. Lu, D. C. McCoy, et al. 2017. "Early Childhood Development Coming of Age: Science through the Life Course." *The Lancet* 389, no. 10064 (January): 77–90. [https://doi.org/10.1016/S0140-6736\(16\)31389-7](https://doi.org/10.1016/S0140-6736(16)31389-7).
- Cameron, A. C., and D. L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *Journal of Human Resources* 50, no. 2 (March): 317–372. <https://doi.org/10.3368/jhr.50.2.317>.
- Campbell, F., G. Conti, J. J. Heckman, S. H. Moon, R. Pinto, E. Pungello, and Y. Pan. 2014. "Early Childhood Investments Substantially Boost Adult Health." *Science* 343, no. 6178 (March): 1478–1485. <https://doi.org/10.1126/science.1248429>.
- Carneiro, P., and J. J. Heckman. 2004. "Human Capital Policy." In *Inequality in America: What Role for Human Capital Policies?*, edited by J. J. Heckman and A. Krueger, 77–240. Cambridge, MA: MIT Press. <https://mitpress.mit.edu/books/inequality-america>.
- Cunha, F., and J. J. Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97, no. 2 (May): 31–47. <https://doi.org/10.1257/aer.97.2.31>.
- . 2009. "The Economics and Psychology of Inequality and Human Development." *Journal of the European Economic Association* 7, nos. 2-3 (May): 320–364. <https://doi.org/10.1162/jeea.2009.7.2-3.320>.
- Cunha, F., J. J. Heckman, and S. M. Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78, no. 3 (May): 883–931. <https://doi.org/10.3982/ECTA6551>.
- Donald, S. G., and K. Lang. 2007. "Inference with Difference-in-Differences and Other Panel Data." *The Review of Economics and Statistics* 89, no. 2 (May): 221–233. <https://doi.org/10.1162/rest.89.2.221>.
- FAO. 2019. *The State of Food Security and Nutrition in the World 2019*. Rome, Italy: Food and Agriculture Organization of the United Nations. <https://doi.org/10.4060/CA5162EN>.
- Gertler, P., J. J. Heckman, R. Pinto, A. Zanolini, C. Vermeersch, S. Walker, S. M. Chang, et al. 2014. "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica." *Science* 344, no. 6187 (May): 998–1001. <https://doi.org/10.1126/science.1251178>.
- Habicht, J.-P., R. Martorell, and J. A. Rivera. 1995. "Nutritional Impact of Supplementation in the INCAP Longitudinal Study: Analytical Strategies and Inferences." *The Journal of Nutrition* 125, no. 4 (April): 1042–50. https://academic.oup.com/jn/article-abstract/125/suppl_4/1027S/4730854.
- Heckman, J. J., S. H. Moon, R. Pinto, P. A. Savelyev, and A. Yavitz. 2010a. "Analyzing Social Experiments as Implemented: A Reexamination of the Evidence from the Highscope Perry Preschool Program." *Quantitative Economics* 1, no. 1 (July): 1–46. <https://doi.org/10.3982/QE8>.
- . 2010b. "The Rate of Return to the Highscope Perry Preschool Program." *Journal of Public Economics* 94, nos. 1-2 (February): 114–128. <https://doi.org/10.1016/j.jpubeco.2009.11.001>.

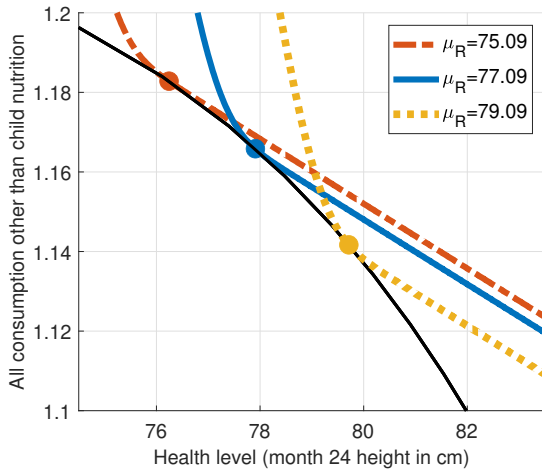
- Heckman, J. J., R. Pinto, and P. Savelyev. 2013. "Understanding the Mechanisms through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103, no. 6 (October): 2052–86. <https://doi.org/10.1257/aer.103.6.2052>.
- Hoddinott, J., H. Alderman, J. R. Behrman, L. Haddad, and S. Horton. 2013. "The Economic Rationale for Investing in Stunting Reduction." *Maternal & Child Nutrition* 9 (September): 69–82. <https://doi.org/10.1111/mcn.12080>.
- Hoddinott, J., J. R. Behrman, J. A. Maluccio, P. Melgar, A. R. Quisumbing, M. Ramirez-Zea, A. D. Stein, et al. 2013. "Adult Consequences of Growth Failure in Early Childhood." *The American Journal of Clinical Nutrition* 98, no. 5 (November): 1170–78. <https://doi.org/10.3945/ajcn.113.064584>.
- Hoddinott, J., J. A. Maluccio, J. R. Behrman, R. Flores, and R. Martorell. 2008. "Effect of a Nutrition Intervention during Early Childhood on Economic Productivity in Guatemalan Adults." *The Lancet* 371, no. 9610 (February): 411–416. [https://doi.org/10.1016/S0140-6736\(08\)60205-6](https://doi.org/10.1016/S0140-6736(08)60205-6).
- Islam, M., and J. Hoddinott. 2009. "Evidence of Intrahousehold Flypaper Effects from a Nutrition Intervention in Rural Guatemala." *Economic Development and Cultural Change* 57, no. 2 (January): 215–238. <https://doi.org/10.1086/592876>.
- Maluccio, J. A., J. Hoddinott, J. R. Behrman, R. Martorell, A. R. Quisumbing, and A. D. Stein. 2009. "The Impact of Improving Nutrition during Early Childhood on Education among Guatemalan Adults." *The Economic Journal* 119, no. 537 (March): 734–763. <https://doi.org/10.1111/j.1468-0297.2009.02220.x>.
- Puentes, E., F. Wang, J. R. Behrman, F. Cunha, J. Hoddinott, J. A. Maluccio, L. S. Adair, et al. 2016. "Early Life Height and Weight Production Functions with Endogenous Energy and Protein Inputs." *Economics & Human Biology* 22 (September): 65–81. <https://doi.org/10.1016/j.ehb.2016.03.002>.
- Richter, L. M., B. Daelmans, J. Lombardi, J. Heymann, F. L. Boo, J. R. Behrman, C. Lu, et al. 2017. "Investing in the Foundation of Sustainable Development: Pathways to Scale up for Early Childhood Development." *The Lancet* 389, no. 10064 (January): 103–118. [https://doi.org/10.1016/S0140-6736\(16\)31698-1](https://doi.org/10.1016/S0140-6736(16)31698-1).
- Victora, C. G., L. Adair, C. Fall, P. C. Hallal, R. Martorell, L. M. Richter, H. S. Sachdev, et al. 2008. "Maternal and Child Undernutrition: Consequences for Adult Health and Human Capital." *The Lancet* 371, no. 9609 (January): 340–357. [https://doi.org/10.1016/S0140-6736\(07\)61692-4](https://doi.org/10.1016/S0140-6736(07)61692-4).



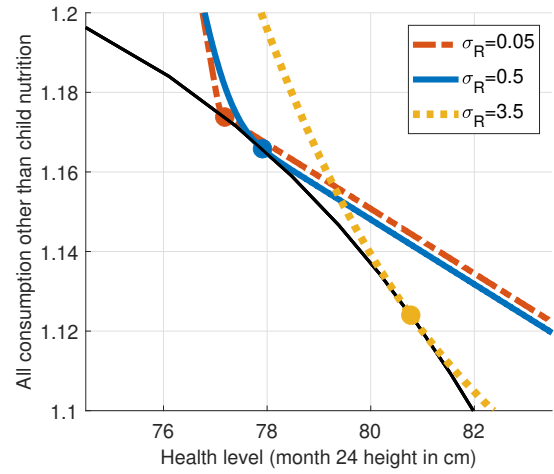
(a) Contours of indifference curves at given estimates



(b) Vary λ



(c) Vary μ_R



(d) Vary σ_R

Fig. 1. Consumption and height outcome indifference curves and choice frontiers. *Note:* Health possibility frontier and indifference curves are visualized in blue (solid-line) given estimated parameters and characteristics for one particular child. The consumption and expected height outcome frontier is determined by the budget constraint as well as the production function. Panels b, c and d show results from varying one parameter at a time in red (dashed-line) and orange (dotted-line).

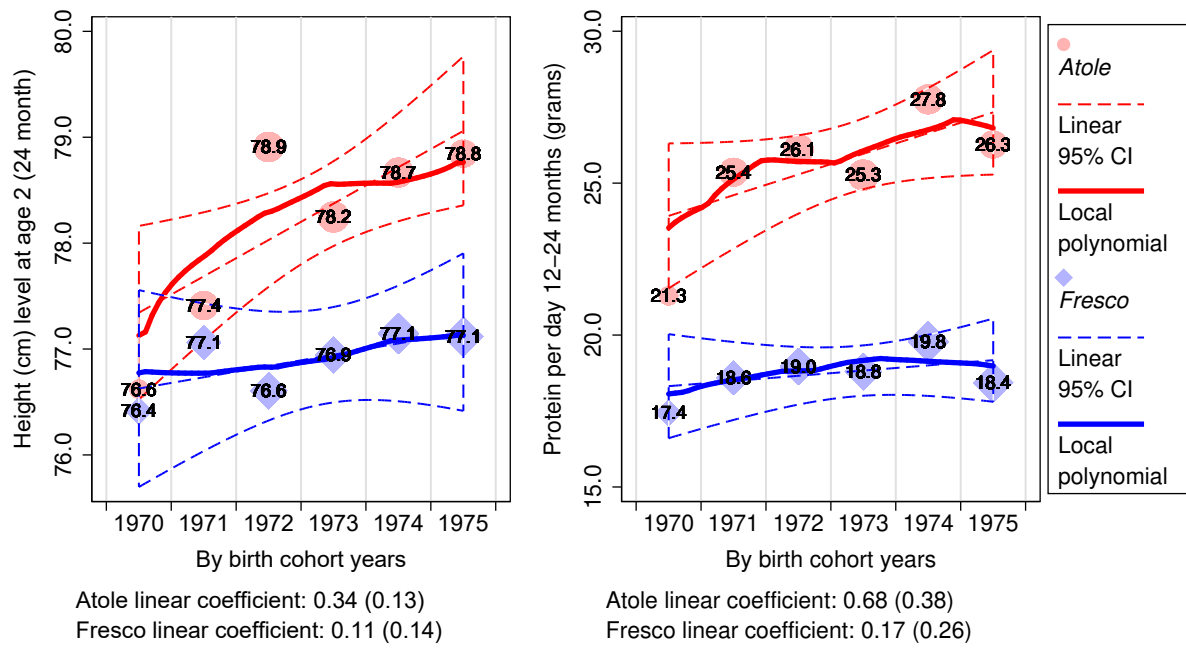


Fig. 2. Increasing protein and height gaps across cohorts.

Table 1: Summary statistics for various variables

	Atole and Fresco villages differences				
	All	Group averages		p-values testing	
	<i>mean (sd)</i>	<i>Fresco</i>	<i>Atole</i>	<i>gap</i>	<i>p-value</i>
Panel a: Gender income price (N=503, main sample)					
Male	0.52 (0.50)	0.52 (0.50)	0.52 (0.50)	-0.00	0.92
Income (quetzale)	515.57 (460.9)	503.68 (464.4)	526.00 (458.4)	22.32	0.59
Mth 15-24 protein price (quetzale/10k grams)	52.58 (3.87)	52.47 (3.93)	52.68 (3.81)	0.21	0.54
Panel b: Gender income (N=1115, height observed once in first 24 months)					
Male	0.53 (0.50)	0.53 (0.50)	0.53 (0.50)	0.00	0.98
Income (quetzale)	449.49 (432.3)	444.63 (446.4)	454.06 (419.0)	9.43	0.72
Panel c: Height					
Month 0 (cm) N=503	49.64 (2.29)	49.79 (2.29)	49.52 (2.29)	-0.27	0.19
Month 6 (cm) N=463	62.72 (2.46)	62.49 (2.50)	62.93 (2.42)	0.44	0.05
Month 12 (cm) N=475	68.81 (2.99)	68.45 (3.13)	69.13 (2.83)	0.68	0.01
Month 18 (cm) N=482	73.37 (3.23)	72.88 (3.26)	73.80 (3.15)	0.92	0.00
Month 24 (cm) N=503	77.66 (3.47)	76.97 (3.49)	78.30 (3.33)	1.33	0.00
Panel d: Average daily nutritional intake					
Month 15 (grams/day) N=464	17.43 (10.5)	14.29 (9.14)	20.07 (10.9)	5.78	0.00
Month 18 (grams/day) N=461	21.52 (11.4)	18.27 (9.61)	24.41 (12.0)	6.14	0.00
Month 21 (grams/day) N=475	24.45 (11.4)	20.17 (9.03)	27.99 (11.9)	7.82	0.00
Month 24 (grams/day) N=462	26.99 (12.0)	22.51 (9.02)	31.07 (13.0)	8.56	0.00
Avg mth 15-24 (grams/day) N=503	22.54 (8.98)	18.78 (6.43)	25.84 (9.62)	7.06	0.00
Avg mth 15-24 (kcal/day) N=503	691.78 (236.5)	681.78 (236.2)	700.55 (236.9)	18.77	0.38

Note: See Section ?? for discussions.

Table 2: Estimated model parameters ($\sigma_R = 0.50$ cm)

Parameter estimates (s.e.)					
<i>Preference</i>			<i>Production function</i>		
ρ	-0.0473	(0.0031)	A	4.1435	(0.0854)
γ	0.0325	(0.0059)	α_{H_0}	0.0220	(0.0088)
λ	-0.0257	(0.0056)	α_{male}	0.0086	(0.0016)
			β	0.0725	(0.0158)
			σ_ε	0.0097	(0.0010)
<i>Price discount</i>			<i>Measurement error</i>		
δ	0.3756	(0.0249)	σ_η	0.3823	(0.0151)
			σ_i	0.0427	(0.0017)

Note: See Section ?? for discussions.

Table 3: The fit of the estimated model's simulated choices to data

	Average protein choices (s.e)						Average height outcome (s.e)					
	Fresco			Atole			Fresco			Atole		
	<i>Model</i>	<i>Data</i>	<i>p-val</i>	<i>Model</i>	<i>Data</i>	<i>p-val</i>	<i>Model</i>	<i>Data</i>	<i>p-val</i>	<i>Model</i>	<i>Data</i>	<i>p-val</i>
Panel a: Average												
All	19.38 (0.45)	18.78 (0.42)	0.18	25.10 (0.64)	25.84 (0.59)	0.24	76.77 (0.26)	76.97 (0.23)	0.44	78.19 (0.17)	78.30 (0.20)	0.52
Panel b: Averages across genders												
Female	18.15 (0.63)	17.31 (0.61)	0.18	23.96 (0.80)	25.19 (0.81)	0.13	76.07 (0.37)	76.19 (0.30)	0.75	77.60 (0.25)	77.69 (0.25)	0.72
Male	20.48 (0.60)	20.10 (0.56)	0.52	26.16 (1.11)	26.43 (0.85)	0.81	77.40 (0.30)	77.67 (0.33)	0.36	78.74 (0.25)	78.86 (0.30)	0.63
Panel c: Averages across years												
1970-71	18.43 (0.95)	18.17 (0.80)	0.78	23.13 (1.18)	24.19 (1.12)	0.37	76.52 (0.43)	76.83 (0.41)	0.47	77.71 (0.39)	77.18 (0.40)	0.18
1972-73	19.30 (0.79)	18.97 (0.74)	0.68	24.51 (0.93)	25.68 (0.88)	0.21	76.80 (0.27)	76.73 (0.42)	0.79	78.07 (0.33)	78.57 (0.32)	0.13
1974-75	19.97 (0.74)	18.96 (0.66)	0.17	27.13 (1.14)	27.15 (1.06)	0.99	76.89 (0.33)	77.24 (0.36)	0.29	78.65 (0.31)	78.77 (0.33)	0.70

Note: For model columns, we report the model population mean predictions. Standard errors in model columns are based on the standard deviation of simulated sample means, with each simulated sample having the same sample size as the observed data. For data columns, we report sub-group observed sample means. Standard errors are based on the observed sample standard deviation and sample size. The p-value column reports the probability of results that deviate further from the observed sample mean occurring, given the sampling distribution given estimated parameters. See Section ?? for discussions.

ONLINE APPENDIX

You are What Your Parents Expect: Height and Local Reference Points

Fan Wang*

A Solution and Estimation Details (Online Publication)

A.1 Reference Point and Preference for Height

A.1.1 The Optimal Nutritional-Choice Problem

Given Equation (??) for the budget constraint, Equation (??) for the production function, Equation (??) for preferences, and the reference points distributions discussed in Section ??, the nutritional-choice problem is:

$$\begin{aligned} \max_N \left\{ C + \rho \cdot C^2 + \gamma \cdot H^{24} + \lambda \cdot \int_{R_{yv}} (H^{24} - R_{yv}) \mathbb{1} \{H^{24} \geq R_{yv}\} dF(R_{yv}) \right\} \\ C + p_{yv}^N \cdot (1 - \delta \cdot \mathbb{1}(v = Atole)) \cdot N = Y \\ H^{24} = \exp(A + X \cdot \alpha + \varepsilon) \cdot N^\beta \end{aligned} \quad (\text{A.1})$$

For notational convenience, we drop subscripts, let $\hat{A} = \exp(A + X \cdot \alpha + \varepsilon)$, replace H^{24} and C as functions of N , and rewrite Equation (A.1) as:

$$\max_N \left\{ \begin{aligned} & (Y - p \cdot N) + \rho \cdot (Y - p \cdot N)^2 \\ & + \gamma \cdot \hat{A} \cdot N^\beta \\ & + \lambda \cdot \left(\hat{A} \cdot N^\beta - \mu_R \right) \Phi \left(\frac{\hat{A} \cdot N^\beta - \mu_R}{\sigma_R} \right) \\ & + \lambda \cdot \sigma_R \cdot \phi \left(\frac{\hat{A} \cdot N^\beta - \mu_R}{\sigma_R} \right) \end{aligned} \right\} \quad (\text{A.2})$$

*. Fan Wang: Department of Economics, University of Houston, Houston, Texas, USA (email: fwang26@uh.edu)

In Equation (A.2), we analytically integrate the integral from Equation (A.1).^{A.1}

A.1.2 Optimal Nutritional Choices and σ_R

Panels a, c, and e of Figure A.1 show health (height) at age 24 months along the x-axis and household consumption along the y-axis. We plot the shape of indifference curves for a given value of γ and different values of λ and σ_R together with the consumption-health possibility frontier, which combines the budget constraint with the production function for height.

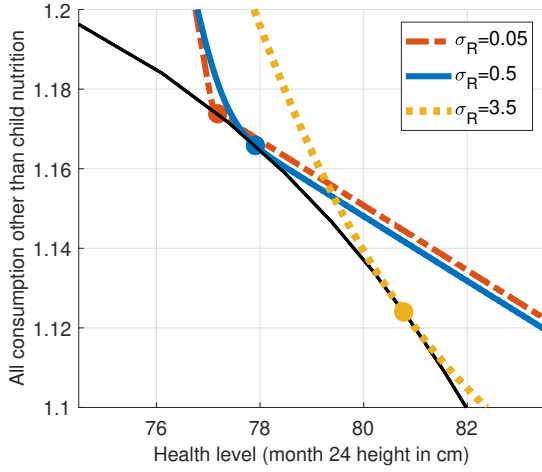
Panels b, d, and f show the difference between height at age 24 months and the mean beliefs about the reference height along the x-axis, and the component of utility that includes the reference point for height along the y-axis. We believe it is helpful to zero in on this component because it is the primary driver of our empirical analysis, and it shows the role that the parameters λ and σ_R play in our study of the determinants of height at age 24 months.

In panels a and b, we use our estimated values of γ and λ for the scenario in which $\sigma_R = 0.5$. In panels c and d, we set λ to be negative 1.5 times the estimated value of γ . In panels e and f, the λ value is lowered to be negative 2.5 times the estimated value of γ .

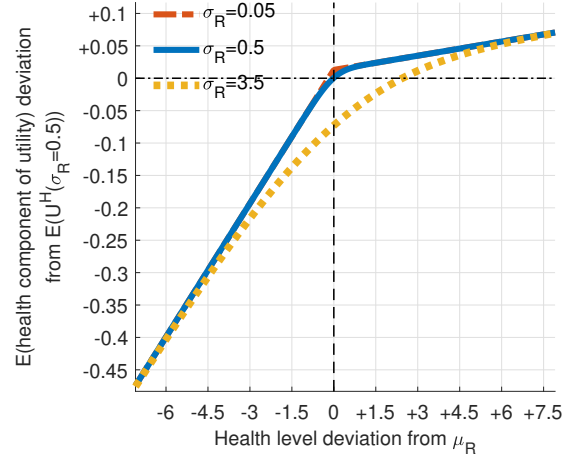
For the configuration of values of γ and λ in panel a, an increase in σ_R increases optimal height because the marginal benefits of height levels that exceed μ_R remains positive as panel b shows. However, for the configuration of values of γ and λ in panel c, we have a bliss point for height that varies just a little as we increase the value of σ_R . In panel e, as the value of λ continues to decrease, the disutility of height beyond the mean reference point decreases so fast that the uncertainty becomes costly. As a result, an exogenous increase in σ_R reduces the optimal level of height.

A.1. The expected utility function contained the expectation of a truncated normal random variable:

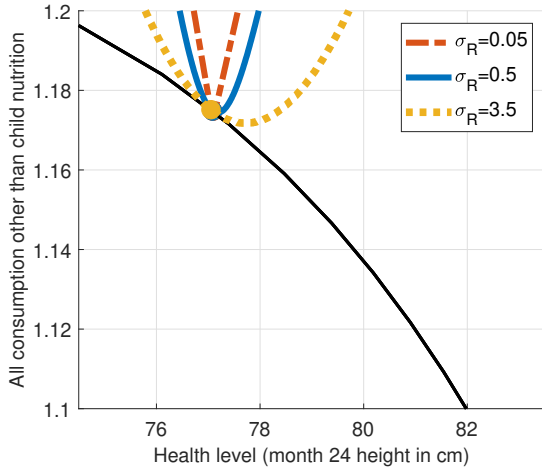
$$\int_{R_{yv}} (H^{24} - R_{yv}) \mathbb{1} \{H^{24} \geq R_{yv}\} dF(R_{yv}) = (H^{24} - \mu_{R_{yv}}) \cdot \left(\Phi \left(\frac{H^{24} - \mu_{R_{yv}}}{\sigma_{R_{yv}}} \right) \right) + \sigma_{R_{yv}} \phi \left(\frac{H^{24} - \mu_{R_{yv}}}{\sigma_{R_{yv}}} \right) \quad (\text{A.3})$$



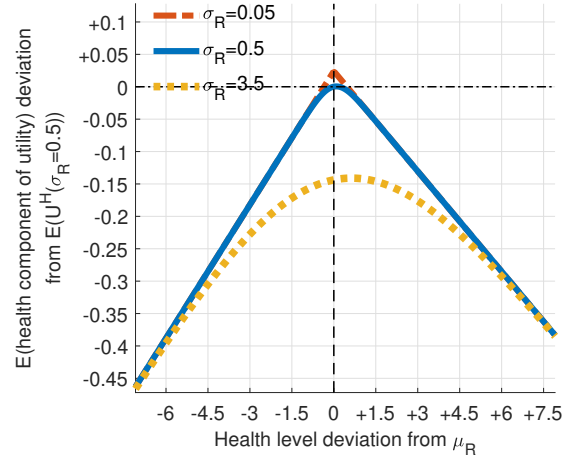
(a) Choices with $\gamma = 0.0325$, $\lambda = -0.0257$



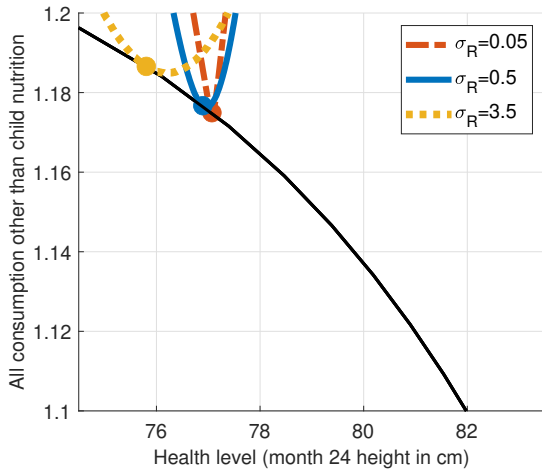
(b) Health preference with $\gamma = 0.0325$, $\lambda = -0.0257$



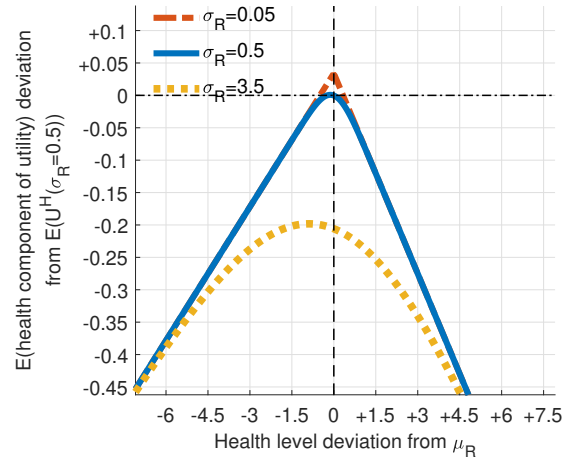
(c) Choices with $\gamma = 0.0325$, $\lambda = -\gamma \cdot 1.5$



(d) Health preference with $\gamma = 0.0325$, $\lambda = -\gamma \cdot 1.5$



(e) Choices with $\gamma = 0.0325$, $\lambda = -\gamma \cdot 2.5$



(f) Health preference with $\gamma = 0.0325$, $\lambda = -\gamma \cdot 2.5$

Fig. A.1. Reference points distribution standard deviation (σ_R) and health outcomes

A.2 Alternative Assumptions about Variance Beliefs

Our model assumes that parents adopt the expected height of two-year-old children in their village as the reference point. Parents estimate the expected height using the observations of the children in their village. We assume that their estimation of the expected height follows a normal distribution with mean belief parameter $\mu_{R_{y,v}}$ and variance belief parameter $\sigma_{R_{y,v}}^2$. In our empirical analysis, we assume that $\mu_{R_{y,v}}$ captures the uncertainty about the mean belief, so $\sigma_{R_{y,v}} = 0.5$.

In our sensitivity analysis, we adopt the assumption that the variance belief is equal to the variance of height. We argue that this approach is an upper bound for variance beliefs. Under this assumption, $\sigma_{R_{y,v}} = 3.5$. To complete our sensitivity analysis, we also investigate our model's decomposition findings for scenarios in which $\sigma_{R_{y,v}} = 1.5$ and $\sigma_{R_{y,v}} = 2.5$.

A.3 Data for Estimation

We include in the estimation children who were born between 1970 and 1975. We show summary statistics for these children in Sections ?? and ?. As discussed earlier, we do not observe both initial heights and heights at month 24 for children born before or after these years.

We use the months 15 to 24 average protein intakes, heights at month 24, protein prices, incomes, gender, and initial height variables shown in Table 1 and described in Section ?? as N , H^{24} , p_{yv}^N , Y , and components of X . We use protein intakes because Puentes et al. (2016) shows that proteins rather than non-protein components of calories matter for height growth in these INCAP data. Ideally, our intake variable should be averaged from month 0 to month 24. However, we do not observe protein values from month 0 to month 12 for close to half of our sample because data collection for this age group started in 1973. Also, it is not easy to estimate the protein component of breastmilk for children who rely on breastfeeding in the first year of life.

Table A.1: Model Predictions from estimating the model under different assumptions about the standard deviation of reference points distribution σ_R .

	Average protein choices				Average height outcome			
	$\sigma_R = 0.5$	$\sigma_R = 1.5$	$\sigma_R = 2.5$	$\sigma_R = 3.5$	$\sigma_R = 0.5$	$\sigma_R = 1.5$	$\sigma_R = 2.5$	$\sigma_R = 3.5$
Panel a: Average								
Fresco	19.38	19.24	19.25	19.29	76.77	76.77	76.78	76.77
Atole	25.10	25.30	25.30	25.47	78.19	78.32	78.33	78.39
Panel b: Averages across genders								
Fresco Female	18.15	18.14	18.16	18.15	76.07	76.11	76.12	76.11
Fresco Male	23.96	24.17	24.17	24.24	77.60	77.74	77.75	77.81
Atole Female	20.48	20.23	20.23	20.30	77.40	77.36	77.36	77.35
Atole Male	26.16	26.34	26.33	26.60	78.74	78.85	78.86	78.93
Panel c: Averages across years								
Fresco 1970-71	18.43	18.61	18.63	18.61	76.52	76.59	76.60	76.57
Atole 1970-71	23.13	23.52	23.46	23.46	77.71	77.85	77.85	77.88
Fresco 1972-73	19.30	19.22	19.24	19.25	76.80	76.79	76.79	76.79
Atole 1972-73	24.51	24.76	24.76	24.92	78.07	78.20	78.20	78.27
Fresco 1974-75	19.97	19.61	19.61	19.69	76.89	76.86	76.86	76.85
Atole 1974-75	27.13	27.13	27.15	27.46	78.65	78.78	78.80	78.87

A.4 Additional Estimation Information

A.4.1 Additional Model fit

In Table A.1, we provide additional information on the fits of the estimated model under varying assumptions about $\sigma_{R_{y,v}}$. Model fits with respect to protein choices and height outcomes are presented across the columns, and model fits with respect to overall averages in Atole and Fresco villages, as well as average by gender and time are presented in separate rows.

Estimated models across $\sigma_{R_{y,v}}$ assumptions – with estimates shown in Table ?? – are all able to provide similarly good fits between model predictions and key data moments.

A.4.2 Additional Estimates

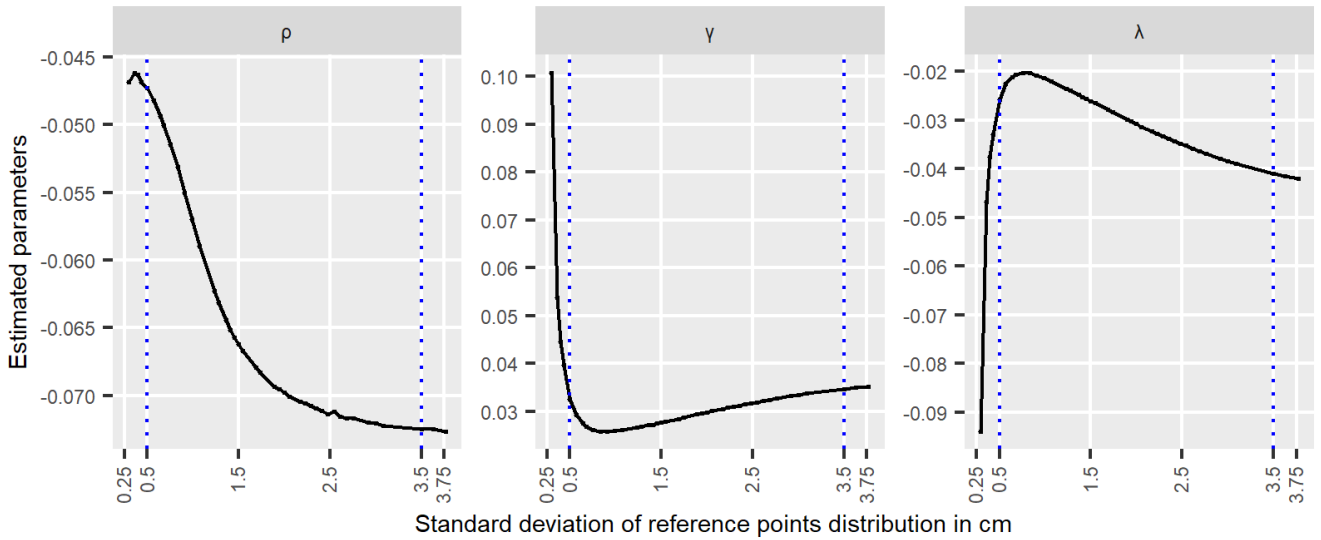


Fig. A.2. Preference parameter estimates from estimating the model under different assumptions about the standard deviation of reference points distribution $\sigma_{R_{y,v}}$

In Table ??, we present parameter estimates under four assumptions for $\sigma_{R_{y,v}}$ from 0.5 cm to 3.5 cm. In this section, for additional information, we provide key estimates from estimating the model at $\sigma_{R_{y,v}}$ values from 0.25 cm to 3.75 cm at 0.05 cm intervals. In Figure A.2, we visualize the point-estimates results for the core preference parameters of ρ , γ and λ in three sub-figures. We mark with dashed blue lines values of $\sigma_{R_{y,v}}$ along the x-axis where $\sigma_{R_{y,v}} = 0.5$ and $\sigma_{R_{y,v}} = 3.5$.

For ρ , the quadratic coefficient on non-nutritional consumptions, the estimated coefficient

becomes more negative as $\sigma_{R_{y,v}}$ increases. This finding implies lower marginal gains from additional non-nutritional consumption at higher $\sigma_{R_{y,v}}$.