Online Appendix: The Apple Does Not Fall Far From the Tree: Intergenerational Persistence of Dietary Habits

Frédéric Kluser *
University of Bern

Martina Pons †
University of Zurich

March 29, 2026

1 Additional Results – Other Dietary Measures

In this section, we examine the intergenerational transmission of different components of the grocery basket, alongside three indicators of dietary preferences. Specifically, we focus on the shares of meat and fish, eggs and dairy, and reliance on convenience foods. In addition, we investigate the persistence of certain dietary types: being a vegetarian, a meat lover, or a convenience-oriented shopper. We define the latter two categories as belonging to the top quartile of consumption within their respective groups.

Table 1 presents the intergenerational transmission results for these outcome variables. Panel (a) reveals strong intergenerational persistence across all dietary measures, though the magnitudes are somewhat smaller than those observed for fruits and vegetables. Nonetheless, all effects exceed those found for income. In particular, meat consumption shows a rank-rank slope of 0.192, indicating a substantial parental influence on their children's meat-purchasing behavior. Convenience food purchases exhibit slightly lower persistence, with a rank-rank slope of 0.179.

Transition probabilities provide additional support for these patterns. Children from low-meat-consuming households (Q1) have a 29.8% probability of remaining in the lowest meat consumption quartile, while those from high-meat-consuming households (Q5) have a 28.5% probability of staying at the top, underscoring the persistence of dietary habits. Convenience food follows a similar pattern.

For the binary outcomes in Panel (b), we only estimate average marginal effects using logistic regressions. Across all outcomes, the effects are close to 0.10. For instance, if a parent is vegetarian, the probability that their child is also vegetarian increases by approximately 10

*University of Bern, Department of Economics, Schanzeneckstrasse 1, CH-3001 Bern, frederic.kluser@unibe.ch.
†University of Zurich, Department of Economics, Schönberggasse 1 CH-8001 Zurich, martina.pons@econ.uzh.ch.

Table 1: Main Mobility Measures: Other Dietary Dimensions

	(a) Rank-Rank Reg.		(b) IGE	(c) CER		(d) Transition Prob.		
	Intercept	Slope	AME	25	75	Q1Q1	Q1Q5	Q5Q5
Panel a)								
Share Meat & Fish	40.671	0.192	0.195	46.534	54.937	29.808	13.645	28.461
	(0.097)	(0.002)	(0.002)	(0.688)	(0.652)	(0.168)	(0.136)	(0.171)
Share Convenience	41.330	0.179	0.177	46.621	53.862	28.824	14.019	27.738
	(0.093)	(0.002)	(0.002)	(0.699)	(0.678)	(0.165)	(0.140)	(0.169)
Share Dairy & Eggs	42.540	0.155	0.139	47.701	54.064	26.640	14.661	27.166
	(0.096)	(0.002)	(0.002)	(0.664)	(0.616)	(0.165)	(0.139)	(0.166)
Panel b)								
Vegetarian			0.105					
			(0.006)					
Meat Lovers			0.105					
			(0.002)					
Lazy			0.101					
			(0.002)					

Notes: The results are estimated using 270,957 observations. The IGE gives the average marginal effect for a quadratic specification. Standard errors are computed using 1,000 bootstrap replications.

percentage points.

It is important to highlight the key differences between these additional measures and the ones considered in the paper. First, fruit and vegetables form a homogeneous category in terms of health benefits: they are unprocessed, widely acknowledged as a cornerstone of a healthy diet, and increased consumption is generally beneficial for almost everyone.

By contrast, categories such as meat and fish encompass a wide range of products, from highly processed and red meats to fresh fish and chicken—items that fall under different classifications in dietary guidelines. As a result, these categories provide a noisier measure that does not capture a specific component of dietary choices, making interpretation involved.

2 Data

2.1 Matching Procedure

This section describes how we match the *customers* in the grocery transaction data with the *residents* in the administrative data. We begin by selecting all combinations of residents and customers with the same location grid cells and age. This generates 4.5 million matches between customers and residents, and we refer to them as *pairs*.[1] We take some additional steps to isolate the unique matches between *residents* and *customers*, proceeding as follows:

1. First, we exclude pairs where the customer's shopping behavior does not fit the resident's

[1]Note that some customers do not match any resident, which is most likely because their addresses in the grocery data are outdated. This is the case for 380,000 of the 2.8 million customers (13.5%), of which 260,000 are active customers (spending more than 50 Swiss francs monthly over our sample period).

locations of residence, since these residents are unlikely to be the actual holders of the linked loyalty card. To this end, we calculate the median annual road distance traveled between a resident's home location and the stores visited by the customer (weighted by trip expenditures). Then, we exclude customer-resident pairs with median shopping trips exceeding 20 kilometers in any year. This step excludes 191,000 pairs.

2. Customers can register in the loyalty program as a family if they have at least one child younger than 25. Hence, we delete all pairs where the customer is registered as a family, and the resident does not fulfill this criterion. This excludes 355,000 pairs.

3. Then, we select all customers that link to exactly one household (multiple residents can live in this household). This gives 1,585,204 unique customer-resident matches.

4. Although households can own multiple loyalty cards, the minimum age to register is 18. Hence, we exclude pairs with more customers than adult residents, eliminating 77,935 pairs.

5. We recover some additional unique matches by identifying consumers who have moved recently without notifying the retailer. To this end, we check whether these movers uniquely match a resident at their old location. This procedure identifies 47,571 additional unique pairs.

6. Removing the customers and residents matched in the previous step, we find an additional 3,845 unique matches at current locations. Steps (1) to (6) result in 1.55 million customers uniquely linked to a resident, accounting for 73% of active customers and 21% of Swiss adult residents.

7. For households owning multiple loyalty cards, we then aggregate expenditures within the household before calculating the relative fruit and vegetable share over the sample period.

8. Additionally, some children moved out recently. In this case, we exclude their expenditures in the periods they still lived with their parents when aggregating the expenditures over time, as these children may contaminate our measure of diet for their parents in the periods before they moved out.[2]

9. We assign the aggregated transaction data to all adult residents in the household. This provides grocery expenditures for 2,248,059 residents living in 1.17 million different households.

10. Finally, we select the 337,950 children for whom we observe at least one of their parents in the final data set.[3]

---

[2]Excluding them entirely leaves our estimates unchanged.

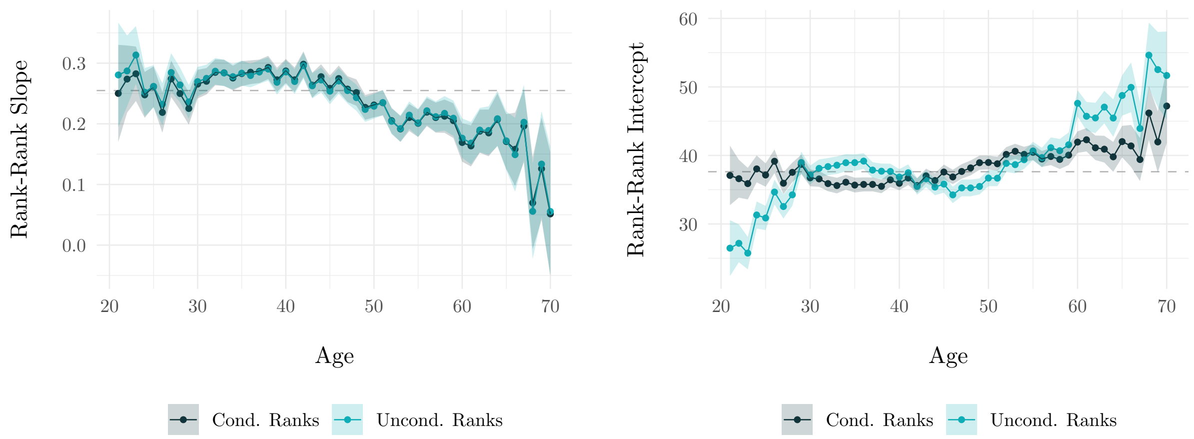
[3]Family linkages, including identifiers for mothers and fathers, have been collected since 2005. This information is available for all individuals unless their parents never lived in Switzerland, died before 2005, or if there was no civil status change either for them or their parents since the 1990s (for example, wedding, divorce, or birth). Consequently, the *Population and Households Statistics* includes information on the parents of 84% of the Swiss residents under age 60, and of 22% above age 60. The coverage for foreigners is lower because many of their parents live abroad. Yet, we include foreigners with known parents in our analysis.

2.2 Additional Information

Structural Survey. The survey questions a representative sample of 200,000 people above age 15 every year on housing, employment, mobility, and education. Participation is mandatory. Education is categorized as either primary, secondary, or tertiary education. Individuals who completed high school or an upper-secondary specialized school have a secondary education. The completion of any degree at a university, university of applied sciences, or university of teacher education results in a tertiary degree. As education stabilizes for most individuals after a certain age, we use educational variables only for individuals above the age of 25 at the time of the survey.

3 Additional Results

Figure 1: Rank-Rank Slope: Life Cycle



(a) Rank-Rank Slope

(b) Rank-Rank Intercept

Notes: Figure 1a shows the rank-rank slope by age group. The grey line is estimated using percentile ranks for children and parents conditional on their age in a variation of Equation (1) fully saturated in the children's age. The blue line is estimated using unconditional percentile ranks. Figure 1b shows the intercepts from the respective regressions. The dashed lines show the rank-rank slope and intercept reported in Table 2 in the paper. 95% confidence intervals are computed using bootstrapped standard errors (1,000 replications).

4 Model Setup

To discuss potential mechanisms explaining the origins of our findings, we introduce a simple framework on habit formation. We model the persistence of diet between generations as the result of a habit stock built during childhood and adjusting over a lifetime (see, for example, Campbell and Cochrane (1999), Fuhrer (2000), and Carroll et al. (2000) for some early work on habit formation models). Habit formation has been used to explain a variety of economic behaviors. For instance, there is evidence of habit formation in voting behavior (Fujiwara et al., 2016), digital addiction (Allcott et al., 2022), health behaviors, or handwashing (Hussam et al., 2022). Related to nutrition, Atkin (2013) finds that higher relative prices in the past shape current tastes, providing evidence of habit formation.

In our setting, individuals are born into families whose diet, skills, and nutritional knowledge exogenously determine their initial stock of habits for their adult life, h_1. We think about the origin of h_1 as a Beckerian parental investment into their children's diet through the transfer of skills and knowledge (see, for example, Becker and Mulligan, 1997). Other unobserved factors outside the household, such as childhood networks, including extended family, friends, and school, also determine habits without invalidating the framework. Individuals enter adulthood and start their own household in period $t = 1$ and live on forever. They maximize their lifetime utility by choosing their relative intake of healthy foods $c_t \in [0, 1]$ for $t = 1, 2, \ldots$, given their initial endowment of habits h_1 and the degree of habit persistence mapping current consumption and habits into future habits:

$$h_{t+1} = h_t + \phi(c_t - h_t), \tag{1}$$

where $\phi \in [0, 1]$ measures the strength of habit formation. Hence, through their consumption behavior, agents continuously update their habits as a weighted average of current habits and consumption. Low values of ϕ imply a high degree of habit persistence and a low degree of learning, and deviations in c_t only have little effect on h_{t+1}. In the extreme case with $\phi = 0$, habits do not adapt, while with $\phi = 1$, the habit at time t equals consumption in the previous period, and there is no habit persistence.

Instantaneous utility in each period takes the form

$$u(c_t, h_t) = g(c_t - c^*) + h(c_t - h_t), \tag{2}$$

where c^* denotes the optimal (healthy) intake of fruit and vegetables, which is assumed to be the same and known for all agents, and the functions $g(\cdot)$ and $h(\cdot)$ have the following properties:

$$\frac{\partial g(c_t - c^\star)}{\partial c} = \begin{cases} > 0, & if\ c_t < c^* \\ = 0, & if\ c_t = c^* \\ < 0, & if\ c_t > c^*, \end{cases} \tag{3}$$

and

$$\frac{\partial h(c_t - h_t)}{\partial c} = \begin{cases} > 0, & if\ c_t < h_t \\ = 0, & if\ c_t = h_t \\ < 0, & if\ c_t > h_t. \end{cases} \tag{4}$$

The two terms in Equation (2) account for two opposing forces. On the one hand, individuals want to eat healthily and be as close as possible to c^*. On the other hand, it is costly (painful) to deviate from one's habits h_t. Hence, any consumption different from $c_t = h_t$ causes disutility through adaptation costs.

To make the problem more concrete, we consider the following specification for the instantaneous utility function:

$$u(c_t, h_t) = -(c_t - c^*)^2 - \rho(c_t - h_t)^2, \tag{5}$$

where ρ is the importance of following one's habit relative to following a healthy diet. The quadratic specification means that small deviations from the optimal diet or one's habit cause little harm. However, large deviations are highly painful in utility terms. Intuitively, these deviations are costlier because they require additional preparation and shopping time, skills and information that need to be acquired (for example, by reading recipes), and new utensils.

Summarizing, each agent solves the following maximization problem:

$$\max_{c_t,h_{t+1}} U(c_t,h_t) = \max_{c_t,h_{t+1}} \sum_{t=1}^{\infty} \beta^{t-1} u(c_t,h_t)$$
$$\text{s.t. } h_{t+1} = h_t + \phi(c_t - h_t),$$
$$u(c_t,h_t) = -(c_t - c^*)^2 - \rho(c_t - h_t)^2,$$
$$h_1 \text{ given,}$$

where β is the discount factor. Solving the model, we find that the policy function $c_t(h_t)$ is a weighted average of the optimal diet c^* and the current habit stock h_t:

$$c_t(h_t) = wc^* + (1-w)h_t, \tag{6}$$

where the weight w is a function of the parameters (ϕ, β, ρ). Section 4.2 provides a detailed derivation of the solution for w. The weight w given to healthy eating increases in β and ϕ and decreases in ρ. Hence, if households are forward-looking (meaning, they care about future consumption), have amenable habits, and derive significant utility from a healthy diet, then they give more weight to following a healthy diet relative to habits.

4.1 Identification and Estimation

To estimate the model, we rely on the same data we use in the rest of the paper and treat children of different ages as people in different periods of their lives. We use data on children between the ages of 30 and 60, calibrate $\beta = 0.95$, and set $c^* = 0.23$. We set the latter equal to the 88th percentile of the distribution of c_t, as the national nutrition survey indicates that only 12% of households meet the recommended intake of fruits and vegetables. If we knew initial habits h_1, we could directly estimate $(1-w)$ in Equation (6). Since we do not directly observe habits, we proxy them with parents' diet denoted $\tilde{h}_1$, introducing a measurement error, and we express h_t and c_t as functions of initial habits h_1 for $t \geq 2$ by iterating backwards the law of motions for habits in Equation (1) and the policy function for consumption in Equation (6):

$$h_t = h_1 (1-w\phi)^{t-1} + c^* w\phi \sum_{j=0}^{t-2} (1-w\phi)^j \tag{7}$$

$$c_t = h_1(1-w)(1-w\phi)^{t-1} + c^* \left[(1-w)w\phi \sum_{j=0}^{t-2} (1-w\phi)^j + w \right]. \tag{8}$$

A regression of c_t on $\tilde{h}_1$ interacted with age dummies identifies $\xi \cdot (1-w)(1-w\phi)^{t-1}$ for all t, where the term $\xi \in (0,1)$ arises from the measurement error. However, using data from different cohorts, we can identify $(1-w\phi)$ and, therefore, the path for habits. We use a two-step estimator, where we first fit a saturated model of c_t on $\tilde{h}_1$ interacted with age fixed effects. Then, in the second step, we impose the structure $\xi \cdot (1-w)(1-w\phi)^{t-1}$ on the coefficients by fitting a linear model in t on the logarithm of the first step slope coefficients.[4] The resulting point estimate is

$$(1 - \hat{w}\hat{\phi}) = 0.988. \tag{10}$$

This expression does not separately identify ϕ and ρ because different values of the parameters are consistent with these results. As an example, consider an individual with $\rho = 1$ and $\phi = 0.021$, satisfying Equation (10). This individual gives a weight of $w = 0.57$ to healthy eating. Yet, the values $\rho = 2$ and $\phi = 0.028$ also satisfy Equation (10) and are, thus, observationally equivalent. This second individual values following a healthy diet less, and she assigns a lower weight to healthy eating ($w = 0.42$). However, she alters her habits faster. Hence, both of these individuals face the identical habit stock in the following periods, as the smaller deviation in consumption is coupled with more flexible habits such that Equation (10) holds.

Figure 2 pictures the continuum of compatible values for ϕ and ρ that satisfy Equation (10). We find that a higher valuation of a healthy diet (lower value of ρ) is consistent with our data if combined with stickier habits (lower ϕ). While if individuals value a healthy diet less (higher ρ), then habits have to be more amenable (higher ϕ). However, what is striking is that even for extremely high values of ρ, our model still implies sticky habits (for example, even a value of $\rho = 20$, implies $\phi = 0.105$). Hence, this provides evidence for the important role of habit formation and gives an explanation as to why most individuals do not meet the dietary recommendations.[5]

Reconciling the model with the empirical heterogeneities we estimate in the paper, we estimate our model for rich and poor individuals separately. Splitting the sample into income quartiles, we estimate $\hat{w}\hat{\phi} = 0.016$ for the top 25% and $\hat{w}\hat{\phi} = 0.012$ for the bottom quartile. Figure 3a shows the values of ϕ and ρ that are consistent with these results. The figure shows that as long as high-income individuals value healthy eating at least as much as low-income individuals, better-earning households face more amenable habits. If, however, low-income individuals value healthy eating more, it is possible that their habits adapt faster. Yet, this is unlikely to be the case as Lleras-Muney and Lichtenberg (2005) find that more educated individuals switch more easily to new drugs, suggesting their adaptation costs are lower. The difference in the estimated value of $w\phi$ for different income groups also implies that higher-income individuals have steeper habit trajectories. To give an illustration, Figure 3b shows the estimated habit trajectories of

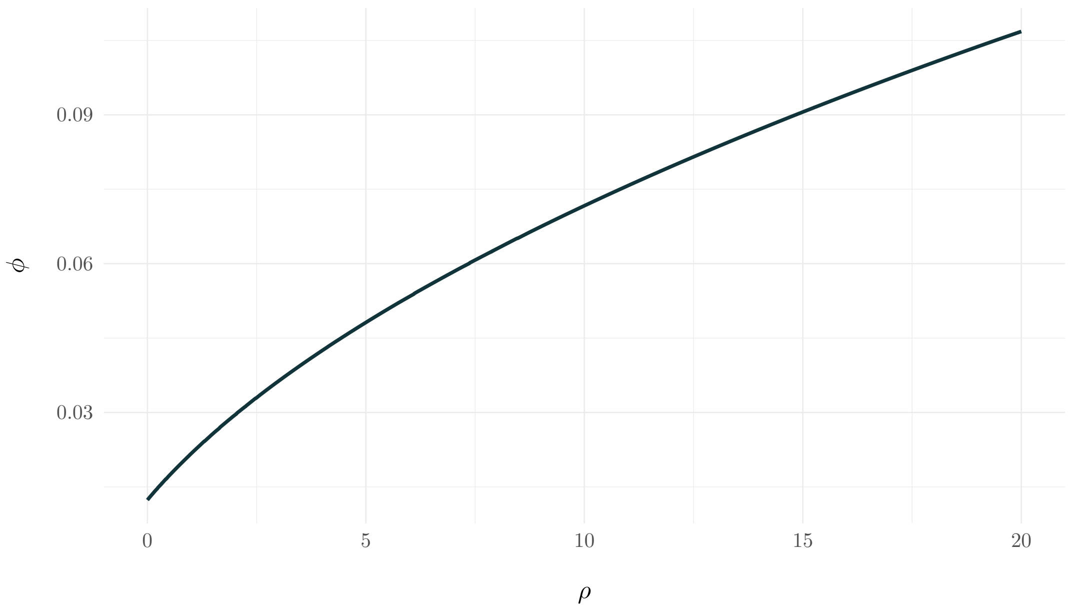
[4] One potential worry of this analysis is that the measurement error is not constant over time. More precisely, if the measurement error increases with age, it would imply that ξ is decreasing over time, consequently affecting the estimation of $log(1-w\phi)$. An alternative approach to estimate $(1-w\phi)$ would deal with the ratios of adjacent cohorts' slope coefficients:

$$\frac{Cov(c_{t+1}, \tilde{h}_1)}{Cov(c_t, \tilde{h}_1)} = (1 - w\phi), \ \forall t > 2, \tag{9}$$

and we can take the average of these ratios. In this way, only the coefficients of adjacent cohorts are compared, making this estimator more robust to potential cohort effects. However, this procedure does not entirely exploit the relationship between the coefficients implied by the model. Using this alternative approach, we find a coefficient of 0.991, suggesting that cohort effects should not invalidate the results.

[5] Regarding the role of discounting, habits are less sticky if the discount rate β is low, as people have lower incentives to invest in future habits and assign more weight to following their habits.

Figure 2: Habit Persistence Parameters



Notes: The figure shows the values of the habit persistence parameter ϕ and the relative utility weight ρ that are consistent with the result in Equation (10).

a low-income and a high-income individual, both with initial habits $h_1 = 0.10$. More affluent individuals build a habit stock that includes 1.25 percentage points more fruit and vegetables over fifty periods. All in all, these results are consistent with the finding of Cutler et al. (2006) that highly educated people are more likely to consume a healthy diet, exercise more, and take more preventive care.

4.2 Derivations

The Bellman equation $V_t(h_t)$ of the optimization problem takes the following form:

$$\begin{aligned} V_t(h_t) &= \max_{c_t} -(c_t - c^*)^2 - \rho\,(c_t - h_t)^2 + \beta V_{t+1}(h_{t+1}) \qquad \text{s.t. } h_{t+1} = h_t + \phi(c_t - h_t) \\ &= \max_{c_t} -\left(\frac{h_{t+1}}{\phi} - \frac{h_t}{\phi} + h_t - c^*\right)^2 - \rho\left(\frac{h_{t+1}}{\phi} - \frac{h_t}{\phi} + h_t - h_t\right)^2 + \beta V_{t+1}(h_{t+1}) \end{aligned} \tag{11}$$

with the resulting optimality conditions:

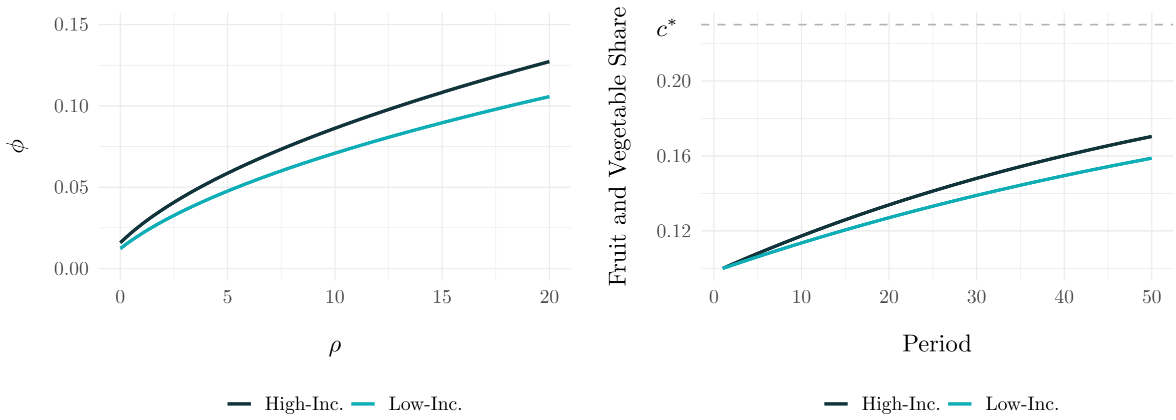
$$0 = -\frac{2}{\phi}(c_t - c^*) - \frac{2\rho}{\phi}(c_t - h_t) + \beta V'_{t+1}(h_{t+1}), \tag{12}$$

$$V'_t(h_t) = -\frac{2(\phi - 1)}{\phi}(c_t - c^*) - \frac{-2\rho}{\phi}(c_t - h_t). \tag{13}$$

Shifting the second FOC one period ahead and combining it with Equation (12) gives the following Euler equation:

$$(c_t - c^*) + \rho(c_t - h_t) = \beta(1 - \phi)(c_{t+1} - c^*) + \beta\rho(c_{t+1} - h_{t+1}). \tag{14}$$

Figure 3: Income Heterogeneities in the Model



(a) Habit Persistence Parameters (b) Evolution of the Habit Stocks

Notes: Figure 3a shows the values of the habit persistence parameter ϕ and the relative utility weight ρ for the best- and lowest-earning quartile of households in the sample. Figure 3b shows the evolution of the habit stock over 50 periods for the two income groups. The dashed grey line shows the optimal level of fruit and vegetable intake c^*.

In our setting with a quadratic utility function and a linear constraint, we can use a guess-and-verify approach. We guess that the policy function for $c_t(h_t)$ is a weighted average of the optimal healthy diet c^* and the current habit stock h_t ($w \in [0, 1]$):

$$c_t(h_t) = wc^* + (1 - w)h_t. \tag{15}$$

Inserting the guess into the Euler equation yields

$$[wc^* + (1 - w)h_t](1 + \rho + \beta\rho\phi) =$$
$$c^*[1 - \beta(1 - \phi)] + h_t [\rho - \beta\rho(1 - \phi)] + [c^*(w + \phi w - \phi w^2) + h_t(1 - w - \phi w + \phi w^2)](\beta(1 - \phi) + \beta\rho).$$

The method of undetermined coefficients provides the following two quadratic equations:

$$\begin{aligned} 0 = {} & \phi\beta(1 - \phi)w^2 + \phi\beta\rho w^2 + (1 + \rho + \beta\rho\phi - \beta(1 - \phi) - \beta\rho - \phi\beta(1 - \phi) - \phi\beta\rho)w \\ & - 1 + \beta(1 - \phi) \qquad (16) \\ 0 = {} & \phi\beta(1 - \phi)w^2 + \phi\beta\rho w^2 + (1 + \rho + \beta\rho\phi - \beta(1 - \phi) - \beta\rho - \phi\beta(1 - \phi) - \phi\beta\rho)w \\ & + \rho - \beta\rho(1 - \phi) + \beta(1 - \phi) + \beta\rho - 1 - \rho - \beta\rho\phi, \qquad (17) \end{aligned}$$

which both simplify to:

$$0 = (\phi\beta(1 - \phi) + \phi\beta\rho)w^2 + (1 + \rho - \beta - \beta\rho + \beta\phi^2)w - 1 + \beta(1 - \phi). \tag{18}$$

Solving this equation, we find that for any calibration, there is a single root satisfying the requirement $w \in [0,1]$:

$$w = \frac{-\phi^2\beta + (1+\rho)(\beta-1) + \sqrt{-4\phi\beta(-1+\beta-\phi\beta)(1-\phi+\rho) + (-\phi^2\beta + (1+\rho)(\beta-1))^2}}{2\phi\beta(1-\phi+\rho)}. \tag{19}$$

Under this value of w, the Euler equation and the resource constraint hold, justifying our initial guess.

References

Allcott, H., M. Gentzkow, and L. Song (2022): "Digital Addiction," *American Economic Review*, 112, 2424–2463.

Atkin, D. (2013): "Trade, Tastes, and Nutrition in India," *American Economic Review*, 103, 1629–1663.

Becker, G. S. and C. B. Mulligan (1997): "The Endogenous Determination of Time Preference," *The Quarterly Journal of Economics*, 112, 729–758.

Campbell, J. and J. Cochrane (1999): "By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior," *Journal of Political Economy*, 107, 205–251.

Carroll, C. D., J. Overland, and D. N. Weil (2000): "Saving and Growth with Habit Formation," *American Economic Review*, 90, 341–355.

Cutler, D., A. Deaton, and A. Lleras-Muney (2006): "The Determinants of Mortality," *Journal of Economic Perspectives*, 20, 97–120.

Fuhrer, J. C. (2000): "Habit Formation in Consumption and Its Implications for Monetary-Policy Models," *American Economic Review*, 90, 367–390.

Fujiwara, T., K. Meng, and T. Vogl (2016): "Habit Formation in Voting: Evidence from Rainy Elections," *American Economic Journal: Applied Economics*, 8, 160–188.

Hussam, R., A. Rabbani, G. Reggiani, and N. Rigol (2022): "Rational Habit Formation: Experimental Evidence from Handwashing in India," *American Economic Journal: Applied Economics*, 14, 1–41.

Lleras-Muney and Lichtenberg (2005): "Are the More Educated More Likely to Use New Drugs?" *Annals of Economics and Statistics*, 79/80, 671–69.