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

Frédéric Kluser †
University of Bern

Martina Pons ‡
University of Bern

First Version: October 2022
This Version: October 3, 2023
Link to the newest version

Abstract

Inadequate diets harm individual health, generate substantial healthcare costs, and reduce labor market income. Yet, the determinants of unhealthy eating habits remain poorly understood. We provide novel evidence of the strong intergenerational transmission of dietary choices from parents to children by exploiting unique grocery transaction records matched with administrative data. We find that children with parents spending one percentage point more on fruits and vegetables also spend 0.23 percentage points more on fresh produce. Our estimates exceed comparable measures for income transmission, indicating that dietary habits acquired during childhood are particularly persistent. Counterfactual analyses show that only 12% of the intergenerational persistence in diet can be explained by the transmission of income and education. Finally, we find substantial heterogeneities in diet transmission and introduce a habit formation model to discuss potential mechanisms.

Keywords: consumption inequality, intergenerational mobility, health behaviors
JEL-codes: D15, D83, I12, J12, L14.

*Please do not cite or circulate. We thank Blaise Melly and Maximilian von Ehrlich for all their support. We thank Simon Büchler, Dino Collalti, Gilles Duranton, Jessie Handbury, Lukas Hauck, Marcel Henkel, Pierre Magontier, Tobias Seidel, and seminar participants at the University of Bern and the ETH Zurich for their helpful comments and suggestions. Support for this project from the Migros Genossenschafts Bund, the Federal Statistical Office and the Central Cooperation Office, the Swiss National Science Foundation with grant ref. 187443 (NRP 77), and two grants from the University of Bern Doc.Mobility program is gratefully acknowledged.

†University of Bern, Department of Economics, Schanzeneckstrasse 1, CH-3001 Bern, frederic.kluser@unibe.ch.

‡University of Bern, Department of Economics, Schanzeneckstrasse 1, CH-3001 Bern, martina.pons@unibe.ch.

1 Introduction

Unhealthy eating habits not only impact our personal health and well-being but also put a substantial economic burden on our healthcare systems. A variety of health conditions, including obesity, cardiovascular diseases, and diabetes, has been linked to inadequate diet, accounting for 18% of all North American deaths (Afshin et al., 2019).[1] Additionally, these lifestyle-related diseases generate high medical costs. For example, according to the American Diabetes Association, every fourth healthcare dollar in the United States is spent on people with diabetes, and patients with diabetes generate more than twice as many medical costs as those without the disease. The detrimental consequences of poor dietary choices highlight the need to investigate the origins of unhealthy eating, opening the way for targeted interventions and policy recommendations. A growing literature has taken on the challenge of understanding determinants of dietary choices, and the general consensus is that eating patterns are highly persistent (see Hut, 2020, Hut and Oster, 2022, Atkin, 2013, 2016) and withstand major personal shocks and interventions (see Oster, 2018, Hut and Oster, 2022, Allcott et al., 2019a).

This paper studies the role of the family in determining dietary patterns by analyzing how parents transmit their nutritional choices to their children. To this end, we exploit unique grocery transaction records matched with Swiss administrative data to analyze the intergenerational persistence of diet. Switzerland is an insightful case to study dietary patterns, as almost everyone has sufficient access to healthy food.[2] Our data contains customer-linked spending by product categories from 1.5 billion shop visits between 2019 and 2020 at the largest Swiss retailer. We enrich this consumption data with family linkages and individual socio-demographic information from the Federal Statistical Office, allowing us to observe the shopping behavior of 220,000 children (10% of the population of children) and their parents. The main variable of interest and our measure of the healthiness of a household's diet is the expenditure share of fresh fruits and vegetables relative to total food expenditures.

Our findings show that family is a crucial determinant of dietary choices. We document an extensive intergenerational persistence in fruit and vegetable shares, indicating a strong transmission of eating choices from parents to children. We estimate a rank-rank slope of 0.23, and children whose parents spend one percentage point more on fruits and vegetables have a 0.23 percentage point higher spending themselves. Further, the children's probability of reaching the top quintile with parents at the bottom quintile is 12.2%. This is substantially smaller than

[1] The leading causes in 2017 were an excessive salt intake and an insufficient whole grain, fruit, and vegetable consumption. Globally, unhealthy diets were responsible for 11 million deaths in 2017 (Afshin et al., 2019).

[2] Switzerland has a high density of grocery stores such that households travel on average 600 meters to the nearest one, and 80% of the population have a store within 2 kilometers (Swiss Federal Statistical Office). In comparison, the median distance to the nearest food store in the United States is 0.9 miles (1,450 meters), and only 40% of the population live no more than a mile ($\approx$ 1,600 meters) from the closest store (USDA). In addition, healthy eating is also relatively affordable in Switzerland. According to the World Bank, less than 0.1% do not have the financial means to follow a healthy diet in Switzerland. In comparison, this is the case for 1.5% of households in the United States, 12% in China, and 97% in Madagascar. The World Bank considers a healthy diet as unaffordable if the lowest-cost basket fulfilling national guidelines for a healthy diet costs more than 52% of a household's income.

the probability that children with parents at the top quintile remain at the top of the distribution (30.7%). A comparison of our findings to income mobility suggests that intergenerational persistence of diet exceeds income transmission across all measures we consider, demonstrating that the development of dietary habits during childhood is a persistent channel through which parents impact their children's future.

One potential worry is that these patterns are simply due to the intergenerational transmission of income and education. More precisely, if highly educated and high-income individuals eat healthier, transmission of these socioeconomic variables could (at least partially) drive our results. To assess this possibility, we isolate the channel that does not pass through income and education with a counterfactual analysis proposed in Chernozhukov et al. (2013). The analysis shows that the transmission of these socioeconomic variables can explain only 12% of the persistence in diet between parents and children. This result indicates that parents impact their children's nutrition directly – for example, through the transfer of nutritional knowledge and dietary habits – rather than indirectly through socioeconomic variables.

Yet, the socioeconomic background of children may still be crucial to foster beneficial habits and to break unhealthy ones. Therefore, we look at different sub-samples and observe that the parents' influence is stronger among children with lower education and income, and the transmission mechanism weakens as the distance between parents and children increases. We find comparable results for mobility across language regions and observe much higher upward mobility in urban areas. Hence, a high socioeconomic status and the exposure to new environments foster healthy eating. To discuss potential mechanisms driving these findings, we introduce a simple model of dietary habit formation, in which agents inherit a habit stock from their parents. These habits influence the agents' consumption. On the one hand, they want to eat healthily while, on the other hand, deviating from one's habit causes disutility. We argue that rather than a lack of financial resources, differences in habit formation and adaptation costs are the most important determinants of dietary persistence.

The existing literature on intergenerational mobility predominantly focuses on income. For example, Chetty et al. (2014) document strong transmissions of income from parents to their children in the United States. Related papers show substantial spatial variation in mobility and disproportional disadvantages for non-white groups and Chetty et al. (2022a,b) argue that having a high share of friends with a high socioeconomic status strongly increases upward income mobility for low-income people.[3] In recent years, various papers conducted comparable analyses for other high-income countries (Acciari et al., 2022, Corak, 2020, Deutscher and Mazumder, 2020, Bratberg et al., 2017), including Switzerland (Chuard and Grassi, 2020).[4]

Yet, a much scarcer literature analyzes mobility in non-pecuniary dimensions like education,

[3]See also Chetty et al. (2016, 2020), and Chetty and Hendren (2018). Rothstein (2019) tries to disentangle the channels behind income persistence and concludes that job networks, as well as the local labor and marriage markets drive income mobility rather than a transmission of education or human capital.

[4]Some studies show that also accumulated wealth is persistent within families, even after four to five generations (Clark and Cummins, 2015, Adermon et al., 2018, or Charles and Hurst, 2003).

jobs, health, and consumption, which may partially be due to the limited data availability. For example, Halliday et al. (2020) analyze mobility in health and find striking gaps by race, region, and parent education, while Black et al. (2005) show that sons of better-educated mothers also attain higher education levels.[5] Nonetheless, the literature analyzing the behavior of consumers is surprisingly scarce. Exceptions rely all on self-reported survey data for small samples (less than 3,000 observations), including Charles et al. (2014) and Waldkirch et al. (2004) who use self-reported total food expenditures and imputed consumption based on the PSID. They find an intergenerational correlation in food expenditures from 0.14 to 0.20. Similarly, Bruze (2018), using the Danish Expenditure Survey, calculates an intergenerational elasticity of 0.41. To the best of our knowledge, our analysis is the first paper analyzing the transmission of dietary habits across generations.

This paper also contributes to the literature focusing on diet. This strain of the literature primarily focuses on evaluating the impact of policies promoting healthier eating behavior. These policies include food subsidies (Bailey et al., 2023, Goldin et al., 2022, Hastings et al., 2021), food labels (Barahona et al., 2023, Araya et al., 2022, Cook et al., 2005), sin taxes (Dickson et al., 2023, Aguilar et al., 2021, Dubois et al., 2020, Allcott et al., 2019b), carbon pricing of nutrition (Springmann et al., 2018), or school-food programs (Berry et al., 2021, Handbury and Moshary, 2021). However, such studies find results with low economic or statistical significance. Contrarily, we contribute to the understanding of eating behaviors' origins in the first place.

The paper is structured as follows. Section 2 introduces the data and presents summary statistics and Section 3 discusses our measures of intergenerational mobility. Section 4 presents our results, Section 5 and Section 6 show a counterfactual analysis and dive into heterogeneities. Section 7 introduces a simple framework on habit formation and discusses potential mechanisms of intergenerational transmission. Section 8 concludes.

2 Data

We analyze the intergenerational transmission of diet by combining (i) individual transaction data from the largest Swiss retailer with (ii) administrative data from the Federal Statistical Office. Throughout this paper, we refer to *children* as adult residents for which we observe at least one parent in the administrative data. They are our population of interest, and we treat their parents' characteristics as observable covariates. To introduce the data, we refer to individuals in the grocery data as *customers* and those in the administrative data as *residents*.

[5] Halliday et al. (2020) find a rank-rank slope of 0.11-0.15, while Andersen (2021) estimates a higher rank-rank slope of 0.28 from Danish register data. Further, intergenerational persistence has been documented for labor force participation (Fernandez et al., 2004) and tax evasion (Frimmel et al., 2019).

2.1 Data Sources

Grocery Transaction Data – The consumption data is from the loyalty program of the largest Swiss grocery retailer, holding a market share of 32.7% in 2020. The program participants identify themselves at the in-store checkout with their loyalty card in exchange for exclusive offers and discounts. 2.8 million individuals hold this loyalty card (i.e., 42% of all Swiss residents above legal age), and 2.1 million are active users spending at least 50 Swiss francs monthly.[6] The program is substantive and captures 79% of the retailer's sales. Also, the retailer charges the same prices throughout the country, independent of local purchasing power, wages, and costs, and stores of comparable size generally offer similar goods, except for local products.

The grocery data provides information on every consumer-linked purchase, including expenditures divided into 41 product categories. In this paper, we focus on the food product categories (*fruits and vegetables*, *meat and fish*, *milkproducts and eggs*, *conservables*, and *other food products*). The outcome of interest throughout this analysis is a child's share of fresh fruits and vegetables relative to total food expenditures. This is a suitable measure for a healthy diet because of four reasons. First, fruits and vegetables have a high correlation with the healthy eating index in Allcott et al. (2019a) of 0.57 and 0.41 and capture a more complex aggregation of individual products. Second, a diet low in fruits or vegetables is among the most frequent reasons for nutrition-related mortality in Afshin et al. (2019). Third, we observe that our measure correlates strongly with the intake of important micronutrients across age groups.[7] Fourth, this measure provides a transparent and objective approximation of dietary quality as it requires no weighting of different nutrients or products.
We use the universe of 1.5 billion customer-linked purchases for the period 2019-2021Q2. The grocery data also contains customer characteristics, including their residence location, age, gender, and household type.[8] The residence locations are coded on a grid of 100×100–meter cells. The grid contains 350,000 cells with a median population of 11 residents.

Administrative Data – We enrich this unique consumption data with administrative records for the Swiss population (8.7 million inhabitants in 2020). Pseudo social security numbers allow linking residents across different administrative data sets. We use three different data sets. The *Population and Households Statistics* provides socio-demographic characteristics for each resident for the years 2016–2020. This includes, among others, information on gender, age, marital status, residence location, household identifiers, and the pseudo-identifiers of spouses and kids. The residence locations are again approximated on the same 100×100–meter grid as in the grocery transaction data. Family linkages, including pseudo-identifiers for mothers

[6] 1 Swiss franc (CHF) equals approximately USD 1.10 on July 19, 2023, meaning CHF 50 ≈ USD 55.

[7] We compare the dispersion of our measure across age groups to the administrative National Nutrition Survey, inquiring 2,000 participants between the age of 18 and 75 about their previous day's diet. The expenditure share of fruits and vegetables has a correlation across age groups of 0.4 with the intake of fibers, 0.38 with phosphorus, 0.33 with zinc, 0.22 with Vitamin A, and 0.29 with magnesium.

[8] The household types include the categories *small households*, *young families*, *established families*, *golden agers*, and *pensioners*. To be a family, you have to register your children. This registration implies additional benefits related to family products.

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.[9] For the analysis, we consider all registered residents in 2020, plus those who moved away or died since 2019 or those who immigrated in 2021 (a total of 9 million people).[10]

The *The Old-Age and Survivors Insurance* dataset contains annual gross labor market income for every resident for the years 2016 to 2020.[11] Throughout this paper, we adjust average household income by the square root of household size.[12] Further, to reduce biases in estimating permanent income due to transitory shocks, we average annual household income for 2016–2020.

The *Structural Survey* for the years 2010–2020 provides information on housing, employment, mobility, and education. The representative survey selects 200,000 people above age 15 every year, and participation is mandatory. From this survey, we attain the highest completed education in a household and take the most recent survey they participated in for every individual. Education is categorized as either primary, secondary, or tertiary education.[13] As education and jobs stabilize for most individuals only after a certain age, we add the characteristics for individuals above age 25 at the time of the survey to the *Population and Households Statistics*.

2.2 Sample Construction

Matching – In a first step, we combine the food transaction data and the administrative data sets based on location grid cells, age, and gender. This generates 5.6 million matches between customers and residents. To identify the unique matches, we take some additional steps.[14]

[9] The coverage for foreigners is lower because many of their parents live abroad. Yet, we include foreigners with known parents in our analysis.

[10] Some customers who died or moved away before August 2021 will still be in the customer database, and we can analyze their diet in the previous period. In the same way, people who immigrated in 2021 may already be customers in our data but not yet residents in the 2020 administrative data.

[11] Contribution to this insurance is mandatory for everyone except for individuals younger than 25 with an annual income below 750 Swiss Francs. The contributions amount to a fixed share of the gross labor market income, including official awards, gifts, and bonuses, and are also mandatory for self-employed individuals.

[12] The calculation is $income\_adjusted = \frac{income\_total}{\sqrt{\#household\_members}}$, where we consider all household members, including small kids. The adjustment follows one of the equivalence scales suggested by the OECD. We compute $income\_total$ as the household's annual income by summing the income of all household members but excluding grown-up children who still live with their parents, as they likely do not contribute to the household's budget.

[13] Primary (or compulsory) education ends at the latest after around eleven mandatory years of school (including kindergarten). 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.

[14] Some customers do not match any resident, which is only possible if their location in the grocery data is outdated or due to measurement errors (for example, if the customer died or moved abroad). This is the case for 400,000 of the 2.8 million customers (14%), of which 330,000 are active customers, spending more than 50 Swiss francs monthly over our sample period.

The refinement proceeds as follows: (i) Using the transaction data itself, we calculate for every pair the median road distance traveled to spend one Swiss franc in a given year. Then, we require a customer to shop close enough to her home by excluding all pairs where the resident traveled more than 20 kilometers for an entire year, as this resident is likely not the owner of the loyalty card she links to. This step excludes 258,000 pairs. (ii) A customer can only be registered in the loyalty program as a family if she has kids younger than 25. Thus, we exclude the 530,213 customer-resident pairs where the customer is registered for the family program, but the resident has no kids fulfilling this criterion. (iii) From the remaining customer-resident combinations, we select customers that link to exactly one household (multiple residents can live in this household). This gives 1,244,071 unique customer-resident matches. (iv) The minimum age to register for the loyalty program is 18. Hence, we exclude households linking to more customers than household members aged 18 and older. This can happen if someone moved without changing their customer address. (v) As some consumers have moved recently without notifying the retailer, we check whether these movers uniquely match a customer at their old location. This procedure uniquely identifies 43,336 additional pairs. (vi) Removing the movers matched in the previous step, we find additional 4,093 matches at their actual location in the last step. This leaves a final sample of 1.4 million customers uniquely linked to a resident, accounting for 66% of active customers and 19% of Swiss adult residents.

Dietary Variables – For households owning multiple loyalty cards, we first aggregate expenditures within the household before calculating the relative share for each food product category at the household level. Additionally, some children moved out recently. In this case, we exclude their expenditures in the periods they still lived with their parents.[15] Then, we assign the aggregated transaction data to all adult residents in the household. This provides grocery expenditures for 2.2 million residents in 1.2 million households.

Sample Selection – In the final step, we generate the intergenerational linkages, selecting pairs of children and parents observed in the data and living in different households. This substantially reduces our sample size to 350,000 children linking to 208,000 of their fathers and 280,000 of their mothers. We adjust both household incomes and grocery expenditures by the square–root of household size. We restrict the sample to children and parents with average monthly grocery expenditures between CHF 50 and 1,000 per capita. This is because too-small monthly baskets might not capture the overall consumption accurately, while too-large monthly baskets are unlikely to suit personal use but are rather from business customers. We keep households with at most ten members to exclude, for example, large cohabiting arrangements and retirement homes. Ultimately, we focus on children between the ages of 21 and 70 and parents between the ages of 48 and 97 to avoid too small age groups in our estimation.[16] Further, we generate parents' variables as the average value of the father and mother weighted by their respective

[15] These children may contaminate our measure of diet for their parents in the periods before they moved out. Excluding them entirely leaves our estimates unchanged.

[16] Because we detect minor life cycles in diet, we provide all our results conditional on age groups and want to ensure that groups are large enough (see Section 3, for details).

food expenditures.[17] This gives a final sample of 220,000 children.

2.3 Summary Statistics

Table 1 displays summary statistics for the consumers' monthly food expenditures and the share allocated to fruits and vegetables. The average customer spends 350 Swiss francs per month (380 USD) and allocates 18% of this money to fresh fruits and vegetables. The last two columns of Table 1 compare these expenditures to the administrative *Household Budget Survey*, showing that our transaction data covers 57% of the average household grocery expenditures on food and beverages.[18]

Looking at different household characteristics, we observe that households increase their grocery expenditures throughout their life from a young age (262 Swiss francs) until age 45-54 (403 Swiss francs) before decreasing them again towards retirement (300 Swiss francs). Meanwhile, the share of these expenditures allocated to fruits and vegetables increases with age from 18% to 20%. This gives a first indication of a potential lifecycle in diet. Food expenditures also grow with income and education, such that, for example, the top income quintile spends 400 Swiss francs per month compared to 237 Swiss francs for the bottom quintile. Wealthier and better-educated households also consume relatively more fruits and vegetables, showing a nutritional inequality across different socioeconomic status as previously observed in Allcott et al. (2019a). Finally, we observe a larger fruit and vegetable share in urban than suburban or rural areas. One explanation could be that households in sparsely populated areas are more likely to buy fresh products from a farmer or own their own garden. Yet, households in rural areas spend with 333 Swiss francs only marginally less on grocery products than households in urban areas (349 Swiss francs), and we do not expect this to affect our results.

In addition, we want to assess the representativeness of our data. Table A1 shows summary statistics for the 192,000 matched children and compares them to the 2.2 million children in the population fulfilling the same selection criteria, while Figure A2 plots municipality-level sample averages against the population values. The average child in the final dataset is 44.1 years old with an adjusted household income of 83,000 Swiss francs. 54% of them are female and 65% married. Further, 53% hold a tertiary degree, and more than 90% live in multi-person households. Regarding geographical characteristics, 19% of the children in our sample live in the French-speaking part of Switzerland, 76% in the German- and 4% in the Italian-speaking region. Our sample resembles the population of children well, with some differences in marital status and the degree of urbanization. In the population, only 50% are married, compared

[17] If parents live together, their household characteristics and consumption behavior are identical, while individual variables vary. If parents have separate living arrangements, household characteristics, and consumption behavior differ, and we average all characteristics in the same way we average the shares of fruit and vegetables.

[18] This survey continuously selects 2,500 households each year, and participants take for an entire month notes on their income and expenditure. Note that as we do not observe beverages, our actual coverage of food products is even higher.

Table 1: Summary Statistics for Kids' Expenditures

	Total Spending			% Fruit & Vegetable			Budget Survey	
	Mean	p50	SD	Mean	p50	SD	Spending	Share
Overall	350	280	254	0.18	0.17	0.09	616	0.57
By Age								
< 34	262	209	183	0.18	0.17	0.09	459	0.57
35–44	374	311	256	0.19	0.18	0.08	654	0.57
45–54	403	330	289	0.18	0.17	0.08	728	0.55
55–64	343	277	245	0.19	0.18	0.09	663	0.52
65+	300	241	213	0.20	0.19	0.09	616	0.49
By Household Income								
< 4,530	237	186	170	0.17	0.16	0.10	409	0.58
4,530–6,717	263	206	195	0.17	0.15	0.09	485	0.54
6,718–9,288	329	267	233	0.17	0.16	0.08	604	0.54
9,289–12,855	365	301	253	0.18	0.17	0.08	713	0.51
12,856+	400	326	282	0.20	0.19	0.08	869	0.46
By Highest Education								
Elementary	246	193	179	0.15	0.14	0.08		
Secondary	330	264	238	0.17	0.16	0.08		
Tertiary	387	315	272	0.20	0.19	0.08		
By Pop. Density								
Rural	333	269	235	0.17	0.16	0.08		
Suburban	358	289	259	0.18	0.17	0.08		
Urban	349	269	266	0.21	0.20	0.09		

Notes: This table shows summary statistics for the transaction records of food expenditures of customers uniquely linked to a kid in the administrative data. The columns titled *Survey* show the average grocery expenditures for food and beverages from the administrative Household Budget Survey, 2015–2017, and the relative share between spending in our data and the survey. *Income Adjusted* adjusts household income by the square root of household size. *Highest Education* is the highest education anyone within the household completed, and *Pop. Density* is measured with the municipality's population density. *Age* and *Income Adjusted* use the respective quintiles in the Household Budget Survey for comparabilty.

to the 65% in the sample, and while 26.3% of the population live in densely populated areas, this only holds for 16.6% in our sample. The latter discrepancy is because we are less likely to identify unique combinations of customers and residents the more people live in a raster cell. We illustrate this in Figure A1a by plotting the share of residents in a municipality linked to a child against the number of children living within this municipality. While we link more than 10% of residents in smaller municipalities to a customer, this share declines as the population grows and lies around 5% for the largest cities. This result is not driven by the difference in penetration rates of the loyalty program across municipalities, as shown in Figure A1b. In summary, our sample represents the target population well, and our expenditures cover a large share of grocery expenditures.

3 Measuring Moblity

In the literature, there is not one single measure of mobility. Instead, many different statistics measure different aspects of mobility, which are not necessarily positively correlated.[19] For this paper, we need to consider that not all measures are interesting in our setting, as we focus on diet and not income, and the two outcomes exhibit important differences. First, our measure of diet is bounded from below and above, while income is not. Second, with our measure of diet, there is an optimal level or interval of consumption of fruits and vegetables, and an increase in vegetable consumption might not be beneficial above a certain threshold. Yet, most of the population seems to be on the left of this unknown threshold.[20] Differently, we usually assume a positive marginal utility of income so that more real income leads to better living standards and higher welfare. Hence, having a higher real wage than your parents seems to be a good thing in most cases, while this is not necessarily the case for the share of fruits and vegetables.

Previous papers analyzing intergenerational mobility faced two challenges: (i) how to approximate the lifetime outcome well enough to handle transitory fluctuations and (ii) how to deal with lifecycle issues. The general approach in the recent literature is to average the outcome of children and parents over some extended period and to restrict the analysis to certain age bins of children and parents, ensuring that children, in the case of income, are old enough to be a regular part of the labor market and that parents are not yet retired to avoid lifecycle and attenuation biases.[21]

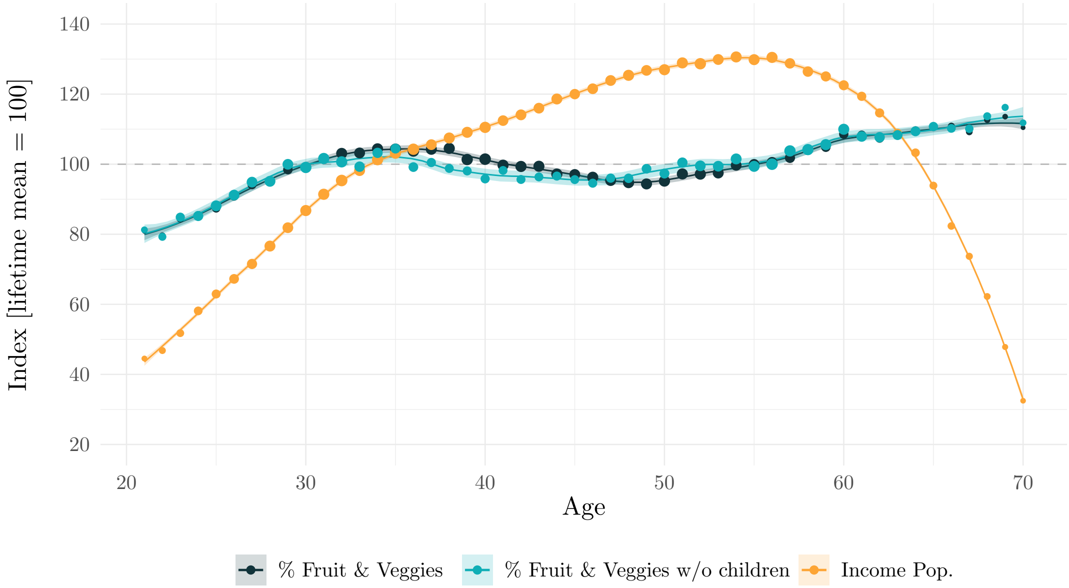
Figure 1 compares the lifecycle variation of diet and income, displaying the average income and the share of fruit and vegetable consumption as a function of age. Both are normalized to the respective lifetime mean to make the results comparable. While income and diet exhibit both some variation over the lifecycle, the variation in diet is substantially smaller than for income. Income more than doubles from age 21 to 60 before declining again towards retirement age. Diet exhibits an s-shaped pattern. Young people tend to have a relatively poor diet, which

[19] Deutscher and Mazumder (2023) provide an extensive discussion and clear classification of different mobility measures, discussing the relationship between them.

[20] According to governmental surveys, 87% the population does not follow the dietary recommendations to consume five portions of fruit and vegetables daily. In our data, 16% of the children spend less than 10% on fruit and vegetables, and 75% have a share smaller than 23%.

[21] There is a large variety of specific approaches. For example, Chetty et al. (2014) rank children's income at ages 29 and 30 within birth cohorts of children and compare it to their parents' five–year average family income when the children were 15 to 19 years old. Chetty and Hendren (2018) use children's income at the household level at age 26. Parents' income is measured as the five–year average household income from 1996 to 2000 (independent of their children's age), and ranks are conditional on birth cohorts. Corak (2020) measures children's individual income at age 38–45, arguing this age approximates average lifetime income very well. He compares this to parents' income as a five–year average when the child was 15–19 years old. Parents' income is defined as the father's and mother's income together. He addresses lifecycle concerns with robustness using children at ages 31 and 32. Acciari et al. (2022) restrict their analysis for Italian children's income at age 34–38 in 2016. Both parents' and children's income is measured as the average from 2016 to 2018. They compare the children's income to parents jointly and fathers and mothers separately. Acciari et al. (2022) address lifecycle issues with an error component model, simulating lifetime income. Similar strategies are also used in papers that do not concern income. For example, Andersen (2021) documents mobility in health, measuring parental health between the age of 60 and 70 and child's age between 36 and 50.

Figure 1: Life Cycle in Income and Diet of Children



Notes: The figure shows the average of three household variables for each age group between 21 and 70: (i) annual household income in the target population adjusted by the square root of household members (2.2 million observations), (ii) the households' expenditure share of fruits and vegetables in the sample (190,000), and (iii) the households' expenditure share of fruits and vegetables for households in the sample who currently do not live together with their children (72,000). All values are normalized to 100 for the lifetime average of each variable, and the points' size indicates the relative number of observations for this age group. The regression lines are estimates from a local regression with uniform confidence bands (with weights for the age groups' sizes).

improves by 30 percent until age 35. After that, there is a small decline of 10 percent until age 50, which then ameliorates again.[22] If we exclude instead households with children, the curve flattens, providing interesting insights. At the age where many households have small children, their diet improves above the lifetime mean. At the same time, they eat unhealthier around the age where they live together with older kids.[23] Given the visible, albeit small, lifecycle in diet, and since we observe children and parents at the same point in time, we will estimate ranks conditional on age as in Chetty et al. (2014) for the positional measures, and we control for age if the measure directly relies on the share of fruits and vegetables. If not indicated otherwise, we always compare a child's household diet to the weighted average of their parent's household diet.

[22]This effect toward the end of life could also be driven by higher survival rates of individuals following a healthy diet.

[23]For both variables, the graph shows the values of the variable at a point in time. Thus the changes could also be due to differences in diet across cohorts and not age effects.

3.1 Rank-Rank Slope

Our first measure of intergenerational mobility is the rank-rank slope (RRS). To take into account a potential lifecycle in diet, we consider the percentile rank of parents and children within each age category. Let r_{ci} denote child i's percentile rank (from 1 to 100) among children conditional on their age. Similarly, let r_{pi} be the percentile rank of their parents within their parents' age group. The rank-rank regression is estimated by regressing the children's rank on the parents' rank:

$$r_{ci} = \alpha + \beta r_{pi} + \epsilon_i, \tag{1}$$

where β is the rank-rank slope which provides a measure of transmission of the parents' position in their generation. For example, a $\beta = 0.3$ tells that if you compare two sets of parents one decile apart, their children are expected to be three percentiles apart. Then, the intercept α is the average rank for the lowest percentile ($r_{ci} = 1$). Without any correlation between r_{ci} and r_{pi}, the slope coefficient would be zero, and the intercept corresponds to the median rank. A steeper slope reflects a less mobile society (meaning more persistence). Hence, if each child were in the same percentile as their parents, the slope would be one, and the line would correspond to the 45-degree line.

3.2 Intergenerational Elasticity

As a second measure, we directly examine the relationship between children's diet and their parents. This measure is similar to the well-established intergenerational elasticity computed by regressing the logarithm of children's income on the logarithm of parents' income.[24] For our measure of diet, we do not take the logarithm, but we fit a quadratic model since it better fits the data. Further, we control for the lifecycle in diet by including parent and child age as well as their squares in the following regression:

$$s_{ci} = \delta_1 s_{pi} + \delta_2 s_{pi}^2 + x_i' \gamma + \nu_i, \tag{2}$$

where s_{ci} and s_{pi} are, respectively, the child's and parents' fruit and vegetable share, and x_i contains the age control variables. Since we fit a polynomial regression, the slope will change for different values of s_{pi}, and we will report the slope at $s_{pi} = \{25, 50, 75\}$.

[24] With a slight abuse of terminology, we refer to this measure as the *intergenerational elasticity*.

3.3 Transition Matrix

Transition matrices break down the children's and parents' distribution into groups of equal size. We group children and parents into quintiles, and the conditional probability that a child is in bin p_j given her parents are in bin p_k is defined as

$$TP_{j,k} = Pr(s_{ci} \in p_j | s_{pi} \in p_k). \tag{3}$$

This transition matrix answers questions like, *"What is the probability that an individual whose parents are in the bottom quintile of the distribution is in the top quintile?"* or "*What is the probability that this individual stays at the bottom of the distribution?"*. Hence, transition probabilities compare children to their parents at a fixed part of the parents' distribution.As for the previous measures, we compute quintiles again for each generation and age group separately. This implies that the bins p_j and p_k are age dependent.[25]

3.4 Conditional Expected Rank

The *Condtional Expected Rank* (CER) is the expected rank of children having parents at population percentile p:

$$CER(p) = \mathbb{E}(r_{ci} | r_{pi} = p). \tag{4}$$

We focus on the CER at the 25th and 75th percentiles, denoted CER25 and CER75. The CER can be estimated parametrically (using directly the information from the rank-rank regression) or nonparametrically. Both have different advantages. On the one hand, the parametric CER for children with parents at the 25th percentile also depends on the observations with parents at the top of the distribution as these observations influence both the intercept and slope of the regression. Hence, the parametric CER may be misspecified. On the other hand, with a large enough data set, one can calculate the CER directly from the subsample of parents at the percentile of interest, which is a fully nonparametric model. This measure is resilient against misspecification, but susceptible to larger variance. We opt for a middle way and use a nonparametric local linear regression evaluated at percentile p which fits our analysis best.

[25] We omit here the dependence of p_j and p_k on age to simplify notation.

Table 2: Comparison of Mobility Measures

	(a) Rank-Rank Reg.		(b) IGE			(c) CER		(d) Transition Prob.		
	Intercept	Slope	25	50	75	25	75	Q1Q1	Q1Q5	Q5Q5
Diet	38.87	0.227	0.252	0.231	0.206	45.41	53.79	30.27	12.16	30.66
	(0.1)	(0.002)	(0.004)	(0.003)	(0.002)	(0.81)	(0.79)	(0.21)	(0.16)	(0.21)
Income	41.81	0.171	0.128	0.126	0.124	43.67	54.43	28.73	15.30	27.96
	(0.19)	(0.004)	(0.005)	(0.005)	(0.006)	(1.20)	(1.13)	(0.30)	(0.26)	(0.27)

Notes: The diet results are estimated using 192,814 observations for children. The income results are estimated using 93,277 observations for children as we restrict the sample to children between the age of 32 and 38 and parents with an average age between 56 and 62 (We also restrict the mothers' age to be between 49 and 61 and the fathers' age between 50 and 62, such that both parents are at least two years away from retirement). The IGE for income uses the log of father's income as an explanatory variable and the log of children's income as a dependent variable. Therefore, we drop 365 observations with zero values.

4 Main Results

This section presents results on the overall persistence of dietary habits across generations. Table 2 reports coefficients and standard errors for all our results. Across all the reported mobility measures, we compute standard errors using 1,000 nonparametric bootstrap replications. Finally, to assess the magnitude of the persistence of dietary choices, we compare the findings to intergenerational mobility in income.

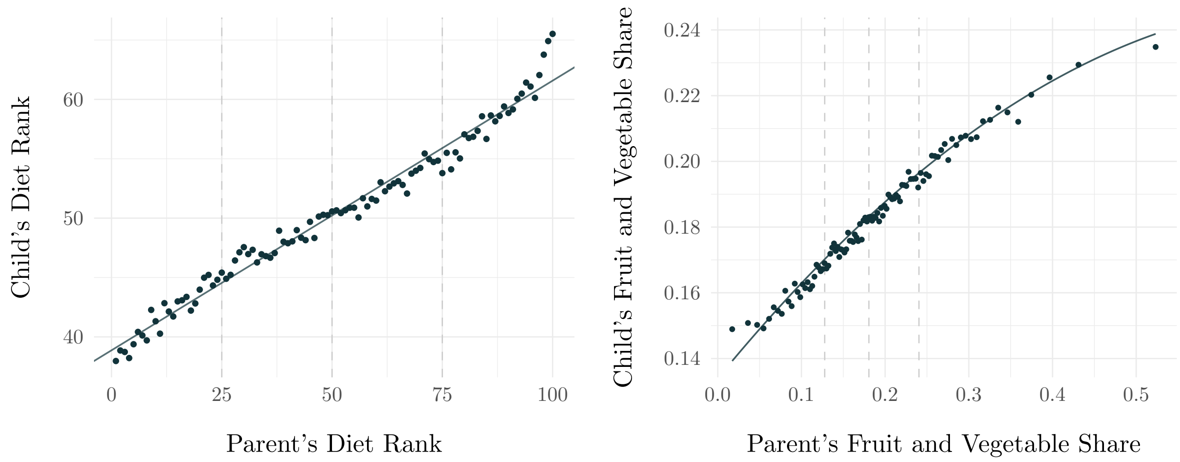
4.1 Dietary Mobility

Rank-Rank Regression – The estimated rank-rank slope in Panel (a) is 0.23, which shows that an increase in the parental percentile rank by one decile corresponds to an increase of 2.3 percentile ranks for the child. To put these results into perspective, it takes 2.98 generations to close the gap between two families at the first and the ninth decile.[26]

Figure 2 graphically illustrates the positional relationship between parents and children, plotting the estimated RRS regression line. The dots represent the average child percentile rank for each of the parents' percentile ranks. The linear model approximates dietary patterns particularly well in the rank-rank model, which aligns with previous findings on income mobility. To show that conditioning the percentile ranks on age solves the lifecycle issues, we allow the intercept and the slope to change over the lifecycle by saturating the model in children's age. While Figure 3a shows that the rank-rank slope is almost identical across both specifications, Figure 3b reveals that the intercept largely depends on the specification of the ranks, explaining the lifecycle observed in Figure 1. This observation supports our expectation that conditional ranks are a better measure of dietary mobility than their unconditional counterparts. The rank-rank slope is remarkably constant in early adulthood at around 0.25, showing that dietary habits acquired at an early age carry on far into adulthood. The rank-rank slope starts declining at around age

[26]The number of generations N to close the gap of $\Delta_{10,90} = 80$ percentile ranks between the first and ninth decile solves $\beta^N \Delta_{10,90} = 1$, such that $N = \frac{log(1/\Delta_{10,90})}{log(\beta)}$.

Figure 2: Intergenerational Diet: RRS and IGE



(a) Rank-Rank Slope

(b) Intergenerational Elasticity

Notes: This figure shows the global measures for intergenerational mobility in diet. Figure 2a shows the estimated the rank-rank regression line based on in Equation (1) and Figure 2b shows the estimation results for the intergenerational elasticity in Equation (2). The dots in both graphs are the average child's rank at each parent's percentile.

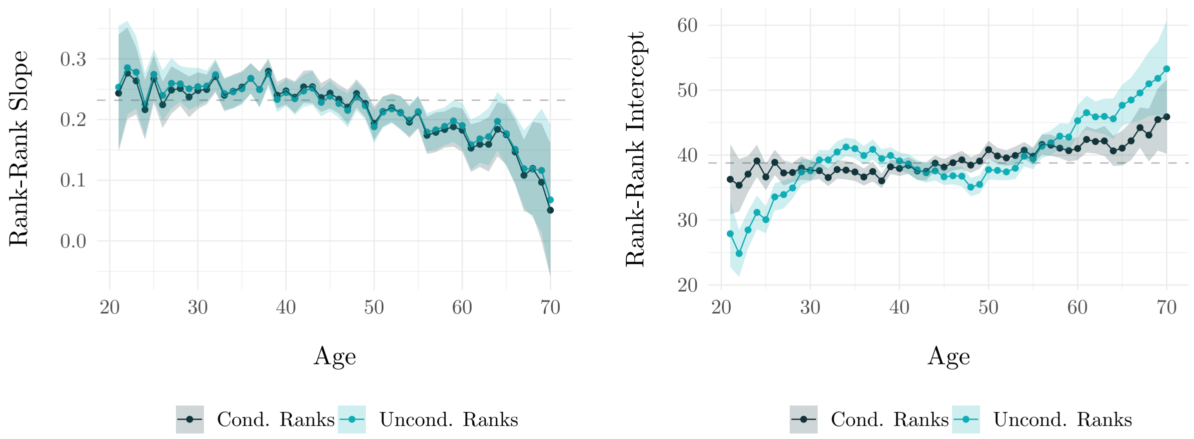
40, which could be explained through habit adaptation that takes several periods to form. Yet, the relationship remains sizable until later in life.

Intergenerational Elasticity – Panel (b) of Table 2 shows our estimates for the intergenerational elasticity in diet at different parental percentiles according to model (2) (namely, at the 25th, 50th, and 75th percentile). We observe in Figure 2b that the estimated slope decreases as the parents' share increases, following the data closely. The decreasing slope suggests that increasing the fruit and vegetable consumption of a parent with a low consumption has a stronger effect on their children's diet. For example, a one percentage point increase in parents' fruit and vegetable consumption is associated with a 0.26 percentage point increase in child consumption for parents at the 25th percentile. This relationship decreases to 0.21 when the parents are at the 75th percentile. Therefore, targeted policy interventions might have the largest benefits for unhealthily eating families, resulting in sizeable improvements in children's diets.

Conditional Expected Ranks – Panel (c) in Figure 2 shows the conditional expected rank estimated nonparametrically. We estimate a CER25 and CER75 of 45.6 and 54.1, respectively. Hence, a child with parents at the 25th percentile of the parents' distribution of fruits and vegetables is, on average, at the 46th conditional percentile of children. In contrast, children with parents at the 75th percentile can expect to reach the 54th percentile. Hence, although we observe strong persistence across generations in diet, there is still substantial reversion to the mean.

Transition Matrix – Figure 4 shows the estimated transition matrix with the corresponding confidence interval. We show selected key results of the transition matrix in Table 2 panel d). Without intergenerational persistence of diet across generations, the transition probabilities would not depend on parents' ranks, and we would observe 20% of children in each cell. The

Figure 3: Rank-Rank Slope: Life-Cycle



(a) Rank-Rank Slope

(b) Rank-Rank Intercept

Notes: Figure 3a shows the rank-rank slope for a given age. The estimation uses ranks for kids and parents conditional on their age according to Equation (1) (with interactions for age groups). The blue line adds the results from the same estimation using unconditional ranks. Figure 3b shows the intercepts (the expected rank for a child with parents at rank zero) from the respective regressions. The dashed lines show the average RRS slope and intercept reported in Table 2. Standard errors are estimated from 100 bootstrap replications.

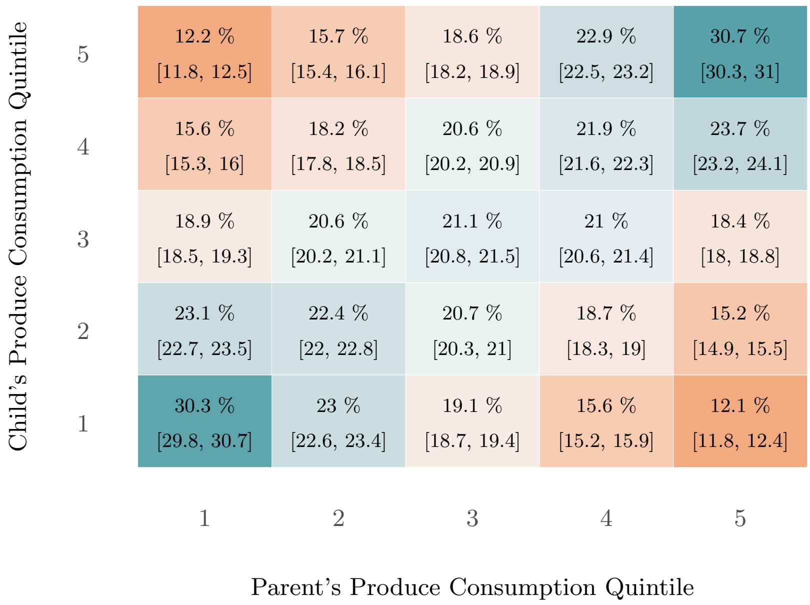
estimated transition probabilities reveal a strong persistence in diet between generations, as children are most likely to be in the same quintile as their parents. Focusing on the cells in the tails of the parents' distribution, 30.8% of children whose parents buy the least fruits and vegetables are also in the lowest quintile of children (corresponding to a Q1Q1 transition), while only 11.7% move up to the highest quintile (Q1Q5). If, on the other hand, a household's parents are among their generation's top 20% fruits and vegetable consumers, their children are most likely also in the fifth quintile (in 31.2% of the cases, Q5Q5). The estimation matrix is precisely estimated, and the confidence intervals are small.

Overall, we find a compelling persistence in healthy food consumption from our extensive supermarket data. Especially the "extreme" transition probabilities face the highest persistence, meaning that the so-called cycles of poverty and privileges are pronounced. At the same time, there is more mobility around the median of the distribution.

4.2 Comparison to Income Mobility

To put the magnitude of our findings into perspective, we compare them to intergenerational mobility in income. To this end, we generate a data set for all Swiss children fulfilling the sample restriction criteria applied to the final data (this sample corresponds to the one used in Table A1). We focus on the relationship between children and parents' income. Further, we average income between 2016 and 2021 to smooth out transitory fluctuations. Observations without income throughout this period have zero income, and if the parents are separated, we average their income. Figure 1 shows that lifecycle issues are far more pronounced for income

Figure 4: Intergenerational Diet



Notes: The figure shows the transition probabilities for children's ranks of fruit and vegetable consumption conditional on their parents' ranks based on Equation (3). We analyze transitions between quintiles and calculate the ranks for children and parents conditional on their age group within the respective subsample of parents and children. 95% confidence intervals in parenthesis are estimated from 100 bootstrap replications.

than for diet. Thus, we follow the procedure of the previous literature trying to select a subgroup of children and parents with stable income (see, among others, Chetty et al., 2014, Corak, 2020, or Acciari et al., 2022), and decide to restrict our analysis to children between the age of 30 and 40 with parents between 52 and 60. This restriction ensures that most children are already participating in the labor market and parents are not yet retired. Figure 1 shows that for these children, income only fluctuates slightly around the lifetime mean (all these age groups are within a maximum deviation from the average lifetime income of 10%), and parents' income is also stable. We estimate for income the same measures for intergenerational income mobility we use for diet, again calculating the ranks within children and parents conditional on age.

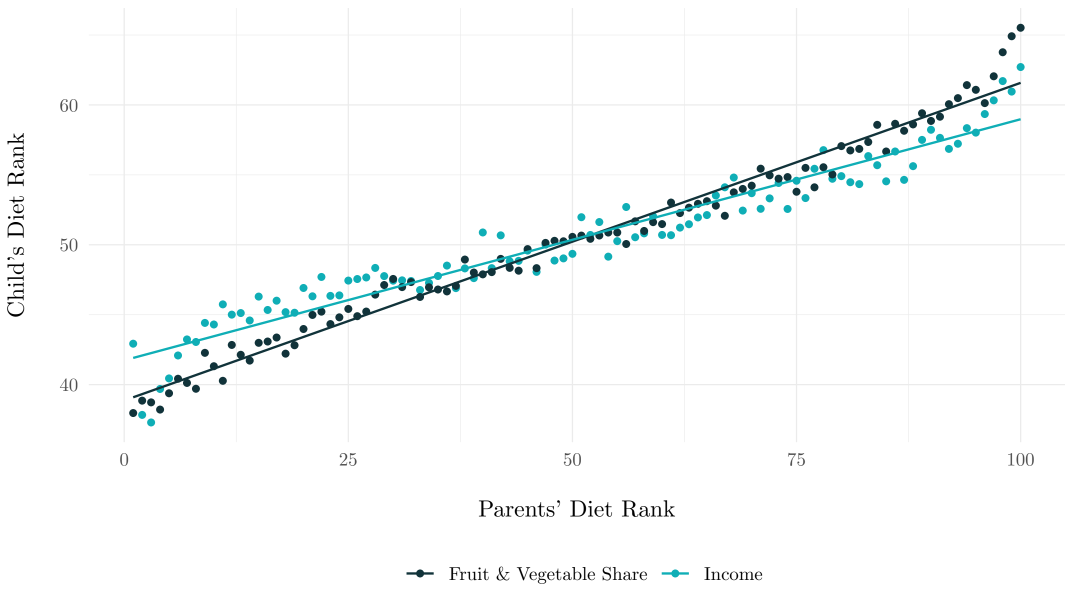
Table 2 shows an estimated RSS of 0.17 and an IGE of 0.16 at the 50th percentile.[27] The conditional expected ranks at the 25th and 75th percentile are 45.78 and 54.63. 28.5% of children with parents at the bottom quintile stay at the bottom, and 13.6% move up to the top.[28] Our estimates on income mobility in Switzerland are in the range of comparable analyses based on the same administrative data source (Chuard and Grassi, 2020).[29]

[27] We measure the intergenerational elasticity in income with a classical log-log specification, however, including a quadratic term.

[28] Different sample selection procedures and income definitions lead to comparable findings. Hence, we are confident that lifecycle issues are also contained for income after conditioning on age groups (as a covariate or in the form of conditional ranks), and our measures are robust to different sample selections.

[29] They use the administrative data for a different and longer time horizon and derive an RRS of 0.14 and an IGE of 0.22. They follow Chetty et al. (2014) and measure the parental income at child age 15 to 19.

Figure 5: Intergenerational Diet vs. Income: RSS



Notes: The figure shows the estimation results for the rank-rank regression in Equation (1) for intergenerational diet and income. The dots in both graphs are the expected child's rank if their parents are at a given percentile.

Comparing our estimated mobility measures between diet and income in Table 2, we observe that intergenerational transmission is more pronounced in the case of eating habits than income across all measures. Figure 5 illustrates this graphically and shows that the slope of the rank-rank regression for diet is substantially steeper. This relationship suggests that the development of dietary habits during childhood is a persistent channel through which parents impact their children's future in a magnitude that exceeds the parental influence on the economic outcomes of their children.

Nevertheless, it is important to note that income is particularly mobile in Switzerland in comparison with most other Western countries, and the relative persistence of diet and income may differ in other countries.[30]

5 Counterfactual Analyis

This section isolates the component of intergenerational transmission in diet that does not pass through the transmission of two important socioeconomic characteristics across generations, namely income and education. Isolating these channels is particularly important as Table 1 shows that better-earning and higher-educated individuals tend to consume more fruits and vegetables. Consequently, it is natural to ask whether and how much of the patterns we docu-

[30] Previous literature estimates, for example, a rank-rank slope income mobility of 0.34 for the United States (Chetty et al., 2014), 0.24 for Canada (Corak, 2020), 0.22 for Sweden and Norway (Bratberg et al., 2017), 0.25 for Italy (Acciari et al., 2022), and 0.21 for Australia (Deutscher and Mazumder, 2020).

ment in this paper are due to the intergenerational transmission of these socioeconomic variables only. To this end, we apply the method proposed in Chernozhukov et al. (2013) and compute counterfactual transition matrices to disentangle these socioeconomic drivers. The idea is to identify the counterfactual distribution by combining a population's conditional distribution function (cdf) with an alternative covariate distribution. Here, we are interested in the conditional distribution of children's diet (conditional on parents' diet) that we would observe if their income and education were independent of their parents' socioeconomic variables.[31] Since the ranks are conditional on age, we include the children's and parents' age in the conditioning set.

Let $F_{s_c|s_p,a_c,a_p}$ be the cdf of the children's diet s_c conditional on the parent's diet s_p as well as the age of children and parents, a_c and a_p. Further, note that from this distribution function, we can construct a transition matrix, provided we observe the marginal distribution of the parent's diet conditional on age. Let x_c denote a vector containing income and education of children, and let x_p contain the corresponding parents' variables. Our main object of interest is the counterfactual distribution of the children's diet that we would observe if we change the covariate distribution $F_{x_c|s_p,a_c,a_p,x_p}$ to $F_{x'_c|s_p,a_c,a_p,x_p}$ denoted $F_{s_c|s_p,a_c,a_p}\langle x|x'\rangle(s_c|s_p,a_c,a_p)$.

Starting from the conditional cdf of children's diet conditional on (s_p,a_c,a_p,x_p,x_c) we can attain $F_{s_c|s_p,a_c,a_p,x_p}\langle x|x'\rangle(s_c|s_p,a_c,a_p,x_p)$ by integrating the conditional cdf over a different covariates distribution:

$$\begin{aligned} &F_{s_c|s_p,a_c,a_p,x_p}\langle x|x'\rangle(s_c|s_p,a_c,a_p,x_p) = \\ &\qquad \int_{\mathcal{X}'_c} F_{s_c|s_p,a_c,a_p,x_c,x_p}(s_c|s_p,a_c,a_p,x_c,x_p)dF_{x'_c|s_p,a_c,a_p,x_p}(x_c|s_p,a_c,a_p,x_p), \end{aligned} \tag{5}$$

where $\mathcal{X}_j$ denotes the support of the covariates x_j for $j=\{c,p\}$ conditional on the other variables. Then, integrating $F_{s_c|s_p,a_c,a_p,x_p}\langle x|x'\rangle(s_c|s_p,a_c,a_p,x_p)$ over the distribution of the parents' covariates yields the desired result:

$$\begin{aligned} &F_{s_c|s_p,a_c,a_p}\langle x|x'\rangle(s_c|s_p,a_c,a_p) = \\ &\qquad \int_{\mathcal{X}_p} F_{s_c|s_p,a_c,a_p,x_p}\langle x|x'\rangle(s_c|s_p,a_c,a_p)dF_{x_p|s_p,a_c,a_p}(x_p|s_p,a_c,a_p). \end{aligned} \tag{6}$$

In the counterfactual scenario that we consider, children's income and education are independent of the parental socioeconomic variables. Also, we assume that parents' age and parents' diet do

[31] We expect no direct channel from parents' income or education on their children's diet. Instead, such effects are more likely to pass through the parents' diet directly. This assumption is consistent with the seminal work of Altonji et al. (1992), which shows that the economic resources of the extended family have no impact on an individual's consumption. This rejects the classical altruism theory of perfect risk and consumption sharing within the extended family.

not affect children's characteristics. Thus,

$$F_{x_c'|s_p,a_c,a_p,x_p}(x_c|s_p,a_c,a_p,x_p) = F_{x_c'|a_c}(x_c|a_c),$$

where we keep the children's age in the conditioning set to account for the lifecycle changes in income and different education distribution over cohorts. Thus, this counterfactual scenario closes the path going from the parents' to the children's diet through the intergenerational transmission of education and income.

The implementation follows the plug-in approach. We obtain the conditional distribution function $F_{s_c|s_p,a_c,a_p,x_c,x_p}$ by inverting the conditional quantile function:[32]

$$F_{s_c|s_p,a_c,a_p,x_c,x_p}(s_c|s_p,a_c,a_p,x_c,x_p) = \int_{(0,1)} 1\left\{Q(u,s_c|s_p,a_c,a_p,x_c,x_p) \le s\right\} du, \quad s \in \mathcal{S} \tag{7}$$

where $Q(\tau, s_c|s_p,a_c,a_p,x_c,x_p)$ is the τ conditional quantile function of s_c given the covariates. For this first step, we estimate flexible quantile regressions for $\tau = \{0.005, 0.015, \ldots, 0.995\}$. The regressors include a second-order polynomial of parents' diet, and for both parents and children we include age and education dummies as well as household income (and its square) interacted with age and a dummy for age ≥ 65. This last term allows income to have a different effect over the life-cycle, which is discontinuous after reaching retirement age.[33] For the estimation of the covariate distribution $F_{x_c'|a_c}$, we use the empirical distribution function:

$$\hat{F}_{x_c'|a_c=k} = \frac{1}{n_k}\sum_{i=1}^{n_k} 1\{X_{ci} \le x\}, \tag{8}$$

where n_k is the number of children in a given age group.

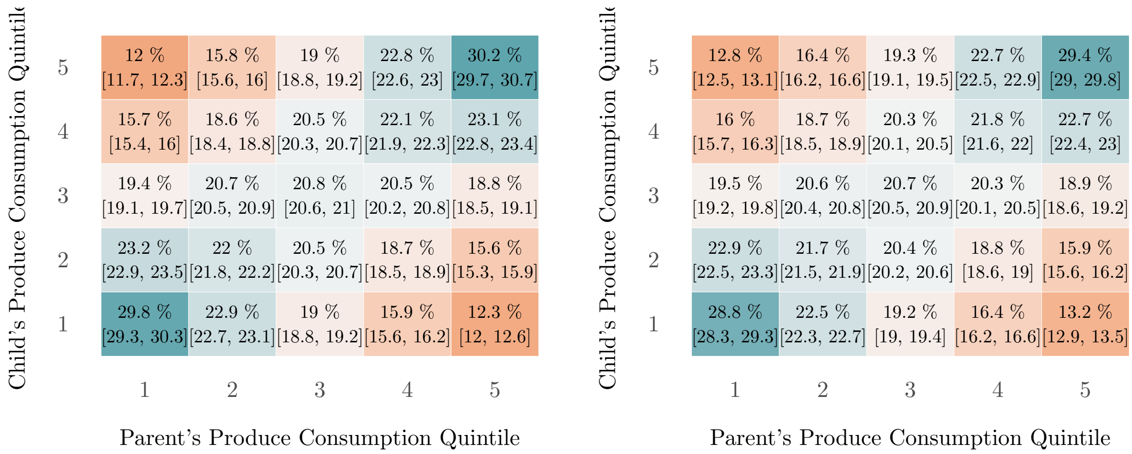
For this analysis, we restrict the sample to the 138,477 children for which we observe their as well as their parents' education. The procedure in this section relies on the correct specification of the conditional quantile function. While we fit a flexible model, to further ensure a meaningful comparison, we re-estimate the transition probabilities in this smaller sample using the same linear quantile model.

Figure 6 shows the estimated transition probabilities with the corresponding bootstrap confi-

[32] For this step, both a quantile regression or a distribution regression can be used (see Chernozhukov et al., 2013). One of the main advantages of a distribution regression is that it does not require a continuous outcome and allows for mixed and discrete ones. However, this does not pose a problem in our case, as our outcome variable exhibits a smooth conditional density. On the other hand, the quantile regression coefficient provides a more natural interpretation.

[33] During the sample period, the retirement age in Switzerland is 65 for men and 64 for women.

Figure 6: Intergenerational Diet: The Role of Income and Education



(a) Transition Matrix

(b) Counterfactual Transition Matrix

Notes: Figure 6a shows the transition matrix and Figure 6b shows the counterfactual transition matrix based on Eqaution (6) and Eqaution (5). The counterfactual considers the case where children's income and education are assigned independently from their parents' values. Bootstrap confidence intervals are in parenthesis. The results are estimated using the sample of 138,477 children for which we observe their as well as their parents education.

dence bands. First, Panel a) displays the transition probabilities estimated with the procedure described above, however, without changing the covariates' distribution. The results in panel a) are not statistically different from the transition probabilities computed nonparametrically for the entire sample in Figure 4. This observation indicates that our estimated quantile regression is flexible enough.

Second, Panel b) shows the counterfactual transition probabilities. The transition matrix is similar to the one in Panel a). However, mobility is statistically significantly higher, mostly in the extremes. For example, the Q1Q1 and Q5Q5 decrease, and the Q1Q5 probability increases. Consider the Q5Q5 cell. Individuals whose parents are in the fifth quintile are ten percentage points (= 30-20) more likely to be themselves in the fifth quintile than if there was no intergenerational transmission of diet. We refer to this as an excess probability. In the counterfactual scenario where we close the channel going through income and education, this number declines to nine percentage points (= 29-20). This change suggests that the transmission of income and education over generations explains around 10% of the excess probability. A similar calculation indicates that around 10% of the excess probability of remaining trapped at the bottom of the distribution can be attributed to income and education transmission.

In order to break down these transition matrices into a unique number, we compute the normalized trace similarly to Jäntti and Jenkins (2015). For the transition matrix in panel a), we find a normalized (off-diagonal) trace of 25.04 which is 5.04 percentage points from the 20 of a perfectly mobile society. For panel b) we have a normalized trace of 24.42. This means that even after closing the path through education and income, there is a gap of 4.42 percentage points, again suggesting that income and education might drive only 12% of the intergenerational transmission of diet. If we do the same exercise for the diagonal elements, we find that only 10% can be explained by education and income.

Overall, these results suggest that only between 10% and 12% of intergenerational persistence of diet can be explained by intergenerational transmission of income and education. This result indicates that parents impact their children's nutrition directly – for example, through the transfer of nutritional knowledge and dietary habits – rather than indirectly through socioeconomic variables.

6 Heterogeneities

Heterogeneities in the persistence of dietary habits across socioeconomic variables might enable dietary changes for some individuals while trapping others. This section unfolds heterogeneities between income classes, education levels, degrees of urbanization, and the distance to parents. Throughout this analysis, we use percentile ranks based on the entire sample and reweight the observations in each group such that the parents' distribution imitates the one in the entire sample. This approach eliminates the mechanical result that children belonging to an unhealthy group have a higher chance of making it to the top.[34]

Table 3 shows the rank-rank slopes, conditional expected ranks, intergenerational elasticities, and transition probabilities for the different subgroups. The second column contains the P-value of the Wald statistic testing for equality of the rank-rank slope between all the subgroups. Bootstrapped standard errors are in parentheses.

First, panel (a) shows the results for the three education levels: primary, secondary, and tertiary. The rank-rank slopes are close to 0.21 in all groups and not statistically different from each other. This means that a higher education for children does not impact how parents transfer their diet. Instead, Figure 7a reveals that the intercept increases with education such that higher-educated children consume more fruits and vegetables. Therefore, education allows children to break out of unhealthy dietary habits, not through a change in the transmission of these habits but through the simple fact that higher-educated households systematically follow a healthier diet, independent of their parents. Multiple reasons may explain this observation. For example, higher-educated individuals may have a more profound nutritional knowledge, a better assimilation of dietary information, or a higher patience, and we discuss such channels in a conceptual framework in Section 7.

Second, panel (b) digs into differences between income groups. To account for the lifecycle in income, the income quartiles are conditional on age, and we keep only children of working age

[34]This happens because, in unhealthy groups, children are more likely to surpass their parents' outcomes through mean reversion. The reweighting procedure gives equal weights to all percentiles in the rank-rank regression and the conditional expected rank. For the transition matrix, the reweighting changes the distribution of children conditional on their parents' bins and, therefore, also changes the children's ranks. For an extensive discussion of weighting approaches in these settings, see Deutscher and Mazumder (2023).

Table 3: Heterogeneities

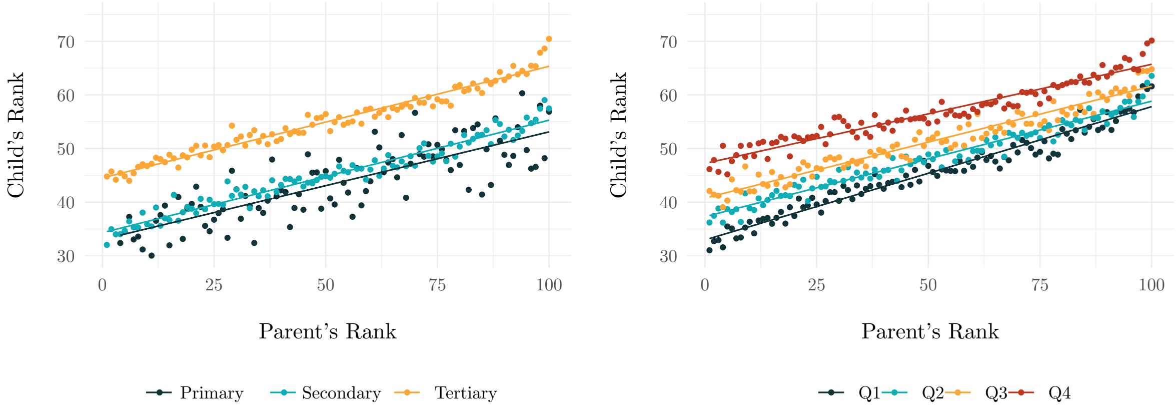
	Rank-Rank		IGE			CER		Transition Prob.			
	RRS	P-value	25	50	75	25	75	Q1Q5	Q1Q1	Q5Q5	N
(a) Kid's Education											
Primary	0.217	0.876	0.264	0.229	0.190	36.76	51.02	7.69	42.71	22.33	4,991
	(0.019)		(0.023)	(0.018)	(0.021)	(4.30)	(4.40)	(1.06)	(1.17)	(1.73)	
Secondary	0.210		0.233	0.208	0.179	39.81	48.80	8.89	36.46	23.77	59,222
	(0.004)		(0.006)	(0.004)	(0.004)	(1.37)	(2.32)	(0.33)	(0.40)	(0.45)	
Tertiary	0.208		0.220	0.208	0.193	50.64	59.39	16.35	22.77	35.34	74,264
	(0.004)		(0.008)	(0.005)	(0.004)	(1.28)	(1.44)	(0.34)	(0.39)	(0.34)	
(b) Kid's Income											
1th Quartile	0.246	0.000	0.261	0.240	0.217	40.21	48.58	8.39	39.65	26.54	45,962
	(0.004)		(0.007)	(0.005)	(0.005)	(1.56)	(1.91)	(0.28)	(0.42)	(0.49)	
2nd Quartile	0.218		0.233	0.214	0.194	45.26	53.36	9.82	31.34	26.22	45,954
	(0.005)		(0.006)	(0.005)	(0.005)	(1.62)	(1.61)	(0.32)	(0.41)	(0.41)	
3rd Quartile	0.203		0.223	0.202	0.178	46.57	53.09	13.19	26.00	30.68	45,940
	(0.004)		(0.007)	(0.005)	(0.005)	(1.85)	(1.60)	(0.37)	(0.47)	(0.46)	
4th Quartile	0.193		0.213	0.201	0.187	52.68	58.04	19.17	21.87	36.99	45,934
	(0.005)		(0.007)	(0.005)	(0.005)	(1.89)	(1.76)	(0.45)	(0.47)	(0.48)	
(c) Kid's place of residence											
Rural	0.217	0.000	0.247	0.219	0.186	41.73	48.75	8.32	34.99	24.25	50,571
	(0.004)		(0.007)	(0.005)	(0.005)	(1.42)	(1.58)	(0.25)	(0.44)	(0.50)	
Suburban	0.211		0.230	0.212	0.192	45.59	53.78	12.12	29.77	28.58	110,098
	(0.003)		(0.005)	(0.004)	(0.003)	(1.16)	(0.97)	(0.21)	(0.26)	(0.28)	
Urban	0.186		0.203	0.192	0.180	53.68	60.33	23.38	20.90	41.65	31,953
	(0.006)		(0.010)	(0.007)	(0.006)	(2.22)	(1.79)	(0.62)	(0.60)	(0.52)	
(d) Distance to Parents											
1th Quartile	0.254	0.000	0.271	0.252	0.231	44.21	52.85	9.83	33.17	29.48	48,204
	(0.004)		(0.007)	(0.005)	(0.004)	(1.59)	(1.41)	(0.30)	(0.51)	(0.50)	
2nd Quartile	0.239		0.266	0.243	0.216	45.57	52.90	10.52	31.48	29.40	48,204
	(0.005)		(0.007)	(0.005)	(0.005)	(1.64)	(1.66)	(0.29)	(0.45)	(0.48)	
3rd Quartile	0.213		0.240	0.218	0.193	45.01	53.73	13.09	29.91	30.29	48,203
	(0.005)		(0.007)	(0.005)	(0.005)	(1.57)	(1.37)	(0.31)	(0.46)	(0.46)	
4th Quartile	0.193		0.219	0.199	0.176	47.68	56.11	15.52	26.48	32.82	48,203
	(0.005)		(0.008)	(0.006)	(0.005)	(1.64)	(1.65)	(0.39)	(0.53)	(0.48)	

Note: The table shows the results for difficult subsamples defined by education, income, residence, and distance to their parents. The second column gives the P-value of the null hypothesis that the rank-rank slope is the same for all subgroups. Bootstrap standard errors in parenthesis are computed using 1,000 replications. The number of observations in each subgroup is shown in the last column.

(25-64).[35] As shown, the rank-rank slope and intergenerational elasticity monotonically decrease as children's income increases. For children in the first income quartile, we find a rank-rank slope of 0.25 compared to 0.185 for individuals in the fourth quartile. Thus, percentile ranks are more persistent over generations among low-income children, meaning that their parents' diet has a stronger and more persistent influence. Figure 7b shows the rank-rank slope and expected ranks for all four income quartiles. Contrasting our findings on education, we observe a statistically significant slope difference in addition to an intercept shift. This means that high-earning children are able to break out of unhealthy childhood habits and to strengthen beneficial ones. In addition to the learnings on education, a higher income appears to incentivize households to

[35]The results are not affected if we use all observations.

Figure 7: Intergenerational Diet: Heterogeneous RRS



(a) RRS: Education

(b) RRS: Income

Notes: This figure shows estimation results from a rank–rank regression (Equation (1)) for different subpopulations, complementing the results in Table 3. Figure 7a displays the RRS for different education levels and Figure 7b for different income levels. The dots in both graphs are the expected child's rank if their parents are at a given percentile.

allocate more money to a healthy consumption.

These findings also translate into geographical differences. Panel (c) shows that mobility is highest in urban areas and lowest in rural areas. The transition probabilities show that children living in urban areas have an outstanding possibility of moving up in the distribution. Strikingly, a child born to parents in the first quintile of the distribution is more likely to find himself at the top of the distribution than in the first quintile. It appears that, in addition to higher education and income, moving to urban areas exposes individuals to new social networks and an abundant grocery supply that favor healthy behaviors.

Lastly, the parental transmission of diet may weaken if households move away from their family and the social network in which they grew up. Hence, panel (d) analyzes the role of the distance between the children's and parents' residences. We observe that nutritional persistence remains high even if children live far away from their parents. However, the further the children move away from their parents, the lower the patterns' persistence. Striking is especially that households are seven percentage points less likely to be trapped at the bottom if they live far away. This finding suggests that new social networks and environments play a decisive role in breaking old habits and is consistent with previous findings on diminishing social interactions and responses to family-related shocks with increasing distance (see, for example, Büchel et al., 2020 and Fadlon and Nielsen, 2019).

7 Conceptual Framework

7.1 Model Setup

To discuss potential mechanisms explaining the origins of our findings, we introduce a simple framework on habit formation and model the persistence in diet between generations as a habit stock inherited from one's parents and adjusted over a lifetime (see, for example, Fuhrer (2000), Carroll et al. (2000), and Campbell and Cochrane (1999) for some early work). In this setting, individuals live for three periods. Period 0 is their childhood, in which their early diet, as well as nutritional knowledge and skills, develop within their family environment, forming the initial stock of habits h_1. We think about the origin of h_1 as in Becker and Mulligan (1997), where parents not only invest financially into their children's future but also invest time and effort to shape patience as a non-pecuniary value. Similarly, parents invest in our model in their children's diet, shaping their initial habits h_1.

Agents then live on during two periods, $t = \{1, 2\}$, and decide in each period the share of healthy food consumption $c_t \in [0, 1]$ by maximizing their lifetime utility function. In period 1, grown-up individuals are endowed with a set of eating habits h_1 inherited from their parents and start their own household. Through their consumption behavior in period $t = 1$, agents will change the stock of habits for period 2, h_2. More precisely, current habits are a weighted average of past habits and past consumption:

$$h_t = h_{t-1} + \phi(c_{t-1} - h_{t-1}), \tag{9}$$

where ϕ measures the strength of habit formation. Low values of $\phi \in [0, 1]$ imply a high degree of habit persistence, and deviations in c_{t-1} only have little effect on h_t. 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.

Instantaneous utility in period $t \in \{1, 2\}$ takes the form

$$u(c_t, h_t) = g(c_t - c^*) + h(c_t - h_t), \quad t = 1, 2 \tag{10}$$

where c^* denotes the optimal (healthy) level of fruits 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:

$$\frac{\partial g(c_t - c^\star)}{\partial c} = g'(c_t - c^*) = \begin{cases} > 0, & if\ c_t < c^* \\ = 0, & if\ c_t = c^* \\ < 0, & if\ c_t > c^*, \end{cases} \tag{11}$$

and

$$\frac{\partial h(c_t - h_t)}{\partial c} = h'(c_t - h_t) = \begin{cases} > 0, & if\ c_t < h_t \\ = 0, & if\ c_t = h_t \\ < 0, & if\ c_t > h_t. \end{cases} \tag{12}$$

The two terms in Equation (10) account for two opposing forces. On the one hand, all 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:

$$u(c_t, h_t) = -(c_t - c^*)^2 - \rho(c_t - h_t)^2, \tag{13}$$

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, they are costlier because larger changes in diet require additional preparation and shopping time, skills and information that need to be acquired (for example, by reading recipes), and new utensils.

Each agent solves the following maximization problem:

$$\max_{c_1, c_2}\ u(c_1, h_1) + \beta u(c_2, h_2) \quad \text{s.t.} \quad \begin{aligned} & h_2 = h_1 + \phi(c_1 - h_1) \\ & h_1 \text{ given.} \end{aligned} \tag{14}$$

Solving the maximization problem yields the following optimal levels of consumption:

$$c_1(h_1, c^*) = \alpha c^* + (1 - \alpha)h_1, \tag{15}$$

$$c_2(h_1, c^*) = \gamma c^* + (1 - \gamma)h_1, \tag{16}$$

with $\alpha = \frac{1+\rho+\beta\rho\phi}{1+2\rho+\rho^2+\beta\rho\phi^2}$, $\gamma = \frac{1+\rho\phi\alpha}{1+\rho}$, and $\gamma > \alpha$.

Hence, the optimal consumption is in both periods a weighted average between one's endowment of habits and the known healthy diet c^*. The weights given to following a healthy diet both become larger as ϕ and β increase but decrease in ρ. We can show that $\gamma > \alpha$ for any calibration, signifying that healthy eating becomes more important over time. Hence, low adjustment costs, flexible habits, and high patience encourage healthier nutrition in the future.

7.2 Discussion

The mechanisms generated by this illustrative framework allow us to replicate our key empirical findings: (i) individuals tend to eat better later in life, and (ii) the intergenerational persistence of diet is highly persistent but decreases with age. We use the model to discuss specific channels driving a wedge between socioeconomic groups.

Habit Persistence (ϕ) – The weak intergenerational mobility we document, even late in life, suggests a small ϕ generating persistent dietary habits. This means that the formation of new habits is slow and that deviations from one's habit only have small effects on the future habit stock. If we allow ϕ to vary across individuals, agents with more amenable habits (high ϕ) face a higher incentive to follow a healthy diet early in life. This behavior positively affects their next periods' utility through the habit formation equation. Under rigid dietary habits (low ϕ) instead, childhood habits are more difficult to change, so the future benefits of eating healthily today are smaller.

Significant heterogeneities in the rank-rank slope in Table 3 indicate that higher-income individuals have a lower ϕ, translating more of their current consumption into future habits. This suggests that they make more deliberate eating choices and learn from them. Further, moving away from one's childhood surroundings exposes individuals to new environments, encouraging a faster adaptation of habits.

Adjustment Costs (ρ) – While habit persistence prevents individuals from sticking to their short-term consumption changes, high adjustment costs discourage such changes in the first place. In addition, dietary changes can be particularly unattractive for individuals with a low ϕ. If changing one's diet is costly and has little effect on future habits, it becomes particularly unattractive.These adaptation costs include affordability and acquiring nutritional information and knowledge, and these barriers are likely easier to overcome for richer and more educated individuals.

For example, Lleras-Muney and Lichtenberg (2005) find that more educated individuals switch more easily to new drugs, suggesting their adaptation costs are smaller. Further supporting the hypothesis that highly educated individuals face lower adaptation costs, Lakdawalla and Goldman (2001) and Case et al. (2005) document that when learning is possible (for example, for larger chronic diseases), education and health have a larger association.

A better access to groceries might also lower adaptation costs in richer, urban areas. However, this is unlikely to be an important channel in a country with a high density of supermarkets. Further, Allcott et al. (2019a) find that even in the United States, where food desert is a prevalent issue, access differentials explain only 10% of the nutritional inequality while differences in demand account for 90%. An additional explanation for lower ρ among agents with a higher SES is that they might enjoy more utility from consuming a healthy diet relative to the adaptation cost, for example, because they assign a higher value to health. This is 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 cases.

Discounting (β) – Complementing the other channels, more patient individuals with a higher β follow a healthier diet already in the first period because of the positive effect on their habit stock. The relative weights given to habits in the two periods in Equation (15) and (16) reflect this, giving more weight to habits in the first period ($\alpha < \gamma$). Individuals of higher economic status may be more patient and value future utility more. This argument relates to the Fuchs (1980) hypothesis, which states that patience might be the reason for the correlation between health and education.

8 Conclusions

The detrimental consequences of bad dietary habits are responsible for a sizeable social and economic burden, while the origins of these harmful eating habits are so far greatly understudied. This paper sheds light on the intergenerational transmission of dietary habits from parents to their children. We do so by combining unique supermarket transaction data with administrative records, including family linkages. We contribute to the literature with novel evidence showing that one's family background is, in fact, a crucial determinant of persistent eating patterns, suggesting that the diet consumed early on in life at one's parents' dinner table shapes our nutritional tastes and preferences throughout our lives. Isolating the part of diet transmission going through education and income, we show that these socioeconomic variables only explain 12% of the intergenerational persistence in diet. Although other unobserved variables of children likely influence eating habits throughout their lives, the direct effect of childhood diet is presumably large. Further, we show the intergenerational transmission of diet varies across observable covariates. Higher-educated and better-earning children generally eat better, independent of their parents. While the transmission mechanism (in terms of the rank-rank slope) does not vary between educational levels, it grows significantly weaker as income rises. Hence, low-income individuals are particularly vulnerable to getting stuck in a cycle of unhealthy diets. Further, our results show even larger upward mobility among children living in urban areas. Among those living in urban areas, children with parents in the first quintile have a higher probability of reaching the highest quintile than remaining trapped in the lowest quintile themselves. We also observe that the transmission also becomes weaker as the distance between children and their parents increases, suggesting that breaking out of one's childhood environment can be a valid way to break unhealthy patterns.

These findings have important implications for public health and policymakers. Recognizing the influence of family on dietary choices helps to design targeted interventions and formulate policy recommendations aimed at promoting healthier eating habits. By understanding the origins of unhealthy eating patterns and the mechanisms through which they are transmitted across generations, policymakers and healthcare professionals can develop effective strategies to combat the rising prevalence of diet-related diseases. For example, policy interventions targeting school food programs, nutritional education for children, and information campaigns at schools and doctors' offices may be particularly effective. We suggest that habit formation is an important mechanism suggesting that not only does the apple not fall far from the tree but also that it doesn't roll far away after falling.

References

Acciari, P., Polo, A., Violante, G.L., 2022. And Yet It Moves: Intergenerational Mobility in Italy. American Economic Journal: Applied Economics 14, 118–163. doi:10.1257/app.20210151.

Adermon, A., Lindahl, M., Waldenström, D., 2018. Intergenerational Wealth Mobility and the Role of Inheritance: Evidence from Multiple Generations. The Economic Journal 128, F482–F513. doi:10.1111/ecoj.12535.

Afshin et al., 2019. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet 393, 1958–1972. doi:10.1016/S0140-6736(19)30041-8.

Aguilar, A., Gutierrez, E., Seira, E., 2021. The effectiveness of sin food taxes: Evidence from Mexico. Journal of Health Economics 77, 102455. doi:10.1016/j.jhealeco.2021.102455.

Allcott, H., Diamond, R., Dubé, J.P., Handbury, J., Rahkovsky, I., Schnell, M., 2019a. Food Deserts and the Causes of Nutritional Inequality. The Quarterly Journal of Economics 134, 1793–1844. doi:10/ggfq88.

Allcott, H., Lockwood, B.B., Taubinsky, D., 2019b. Should We Tax Sugar-Sweetened Beverages? An Overview of Theory and Evidence. Journal of Economic Perspectives 33, 202–227. doi:10/gjhkhg.

Altonji, J., Hayashi, F., Kotlikoff, L., 1992. Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data. American Economic Review 82. doi:10.3386/w3046.

Andersen, C., 2021. Intergenerational health mobility: Evidence from Danish registers. Health Economics 30, 3186–3202. doi:10.1002/hec.4433.

Araya, S., Elberg, A., Noton, C., Schwartz, D., 2022. Identifying Food Labeling Effects on Consumer Behavior. Working Paper doi:10.2139/ssrn.3195500.

Atkin, D., 2013. Trade, Tastes, and Nutrition in India. American Economic Review 103, 1629–1663. doi:10.1257/aer.103.5.1629.

Atkin, D., 2016. The Caloric Costs of Culture: Evidence from Indian Migrants. American Economic Review 106, 1144–1181. doi:10.1257/aer.20140297.

Bailey, M.J., Hoynes, H., Rossin-Slater, M., Walker, R., 2023. Is the Social Safety Net a Long-Term Investment? Large-Scale Evidence from the Food Stamps Program. Review of Economic Studies doi:10.1093/restud/rdad063.

Barahona, N., Otero, C., Otero, S., 2023. Equilibrium Effects of Food Labeling Policies. Econometrica 91, 839–868. doi:10.3982/ECTA19603.

Becker, G.S., Mulligan, C.B., 1997. The Endogenous Determination of Time Preference. The Quarterly Journal of Economics 112, 729–758. doi:10.1162/003355397555334.

Berry, J., Mehta, S., Mukherjee, P., Ruebeck, H., Shastry, G.K., 2021. Crowd-out in school-based health interventions: Evidence from India's midday meals program. Journal of Public Economics 204, 104552. doi:10.1016/j.jpubeco.2021.104552.

Black, S.E., Devereux, P.J., Salvanes, K.G., 2005. Why the Apple Doesn't Fall Far: Understanding Intergenerational Transmission of Human Capital. American Economic Review 95. doi:10.1257/0002828053828635.

Bratberg, E., Davis, J., Mazumder, B., Nybom, M., Schnitzlein, D.D., Vaage, K., 2017. A Comparison of Intergenerational Mobility Curves in Germany, Norway, Sweden, and the US. The Scandinavian Journal of Economics 119, 72–101. doi:10.1111/sjoe.12197.

Bruze, G., 2018. Intergenerational mobility: New evidence from consumption data. Journal of Applied Econometrics 33, 580–593. doi:10.1002/jae.2626.

Büchel, K., Ehrlich, M.V., Puga, D., Viladecans-Marsal, E., 2020. Calling from the outside: The role of networks in residential mobility. Journal of Urban Economics 119, 103277. doi:10.1016/j.jue.2020.103277.

Campbell, J., Cochrane, J., 1999. By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. Journal of Political Economy 107, 205–251. doi:10.1086/250059.

Carroll, C.D., Overland, J., Weil, D.N., 2000. Saving and Growth with Habit Formation. American Economic Review 90, 341–355. doi:10.1257/aer.90.3.341.

Case, A., Fertig, A., Paxson, C., 2005. The lasting impact of childhood health and circumstance. Journal of Health Economics 24, 365–389. URL: https://linkinghub.elsevier.com/retrieve/pii/S0167629604001262, doi:10.1016/j.jhealeco.2004.09.008.

Charles, K., Danziger, S., Li, G., Schoeni, R., 2014. The Intergenerational Correlation of Consumption Expenditures. American Economic Review 104, 136–140. doi:10.1257/aer.104.5.136.

Charles, K., Hurst, E., 2003. The Correlation of Wealth across Generations. Journal of Political Economy 111, 1155–1182. doi:10.1086/378526.

Chernozhukov, V., Fernández-Val, I., Melly, B., 2013. Inference on Counterfactual Distributions. Econometrica 81, 2205–2268. doi:10.3982/ECTA10582.

Chetty, R., Hendren, N., 2018. The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates. The Quarterly Journal of Economics 133, 1163–1228. doi:10.1093/qje/qjy006.

Chetty, R., Hendren, N., Jones, M.R., Porter, S.R., 2020. Race and Economic Opportunity in the United States: an Intergenerational Perspective. The Quarterly Journal of Economics 135, 711–783. doi:10.1093/qje/qjz042.

Chetty, R., Hendren, N., Katz, L.F., 2016. The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. American Economic Review 106, 855–902. doi:10.1257/aer.20150572.

Chetty, R., Hendren, N., Kline, P., Saez, E., Turner, N., 2014. Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility. American Economic Review 104, 141–147. doi:10.1257/aer.104.5.141.

Chetty, R., Jackson, M.O., Kuchler, T., Stroebel, J., Hendren, N., al., 2022a. Social capital I: measurement and associations with economic mobility. Nature 608, 108–121. doi:10.1038/s41586-022-04996-4.

Chetty, R., Jackson, M.O., Kuchler, T., Stroebel, J., Hendren, N., al., 2022b. Social capital II: determinants of economic connectedness. Nature 608, 122–134. doi:10.1038/s41586-022-04997-3.

Chuard, P., Grassi, V., 2020. Switzer-Land of Opportunity: Intergenerational Income Mobility in the Land of Vocational Education. Working Paper doi:10.2139/ssrn.3662560.

Clark, G., Cummins, N., 2015. Intergenerational Wealth Mobility in England, 1858-2012: Surnames and Social Mobility. The Economic Journal 125, 61–85. doi:10.1111/ecoj.12165.

Cook, P.J., Ostermann, J., Sloan, F.A., 2005. The Net Effect of an Alcohol Tax Increase on Death Rates in Middle Age. American Economic Review 95, 278–281. doi:10.1257/000282805774670419.

Corak, M., 2020. The Canadian Geography of Intergenerational Income Mobility. The Economic Journal 130, 2134–2174. doi:10.1093/ej/uez019.

Cutler, D., Deaton, A., Lleras-Muney, A., 2006. The Determinants of Mortality. Journal of Economic Perspectives 20, 97–120. doi:10.1257/jep.20.3.97.

Deutscher, N., Mazumder, B., 2020. Intergenerational mobility across Australia and the stability of regional estimates. Labour Economics 66, 101861. doi:10.1016/j.labeco.2020.101861.

Deutscher, N., Mazumder, B., 2023. Measuring Intergenerational Income Mobility: A Synthesis of Approaches. Journal of Economic Literature 61, 988–1036. doi:10.1257/jel.20211413.

Dickson, A., Gehrsitz, M., Kemp, J., 2023. Does a Spoonful of Sugar Levy Help the Calories Go Down? An Analysis of the UK Soft Drinks Industry Levy. Review of Economics and Statistics , 1–29doi:10.1162/rest_a_01345.

Dubois, P., Griffith, R., O'Connell, M., 2020. How Well Targeted Are Soda Taxes? American Economic Review 110, 3661–3704. doi:10.1257/aer.20171898.

Fadlon, I., Nielsen, T.H., 2019. Family Health Behaviors. American Economic Review 109, 3162–3191. doi:10.1257/aer.20171993.

Fernandez, R., Fogli, A., Olivetti, C., 2004. Mothers and Sons: Preference Formation and Female Labor Force Dynamics. The Quarterly Journal of Economics 119, 1249–1299. doi:10.1162/0033553042476224.

Frimmel, W., Halla, M., Paetzold, J., 2019. The Intergenerational Causal Effect of Tax Evasion: Evidence from the Commuter Tax Allowance in Austria. Journal of the European Economic Association 17, 1843–1880. doi:10.1093/jeea/jvy033.

Fuchs, V., 1980. Time Preference and Health: An Exploratory Study, in: Economic Aspects of Health. NBER Books, Cambridge, MA, pp. 93–120. doi:10.3386/w0539.

Fuhrer, J.C., 2000. Habit Formation in Consumption and Its Implications for Monetary-Policy Models. American Economic Review 90, 367–390. doi:10.1257/aer.90.3.367.

Goldin, J., Homonoff, T., Meckel, K., 2022. Issuance and Incidence: SNAP Benefit Cycles and Grocery Prices. American Economic Journal: Economic Policy 14, 152–178. doi:10.1257/pol.20190777.

Halliday, T.J., Mazumder, B., Wong, A., 2020. The intergenerational transmission of health in the United States: A latent variables analysis. Health Economics 29, 367–381. doi:10.1002/hec.3988.

Handbury, J., Moshary, S., 2021. School Food Policy Affects Everyone: Retail Responses to the National School Lunch Program. NBER Working Paper 29384. doi:10.2139/ssrn.3897936.

Hastings, J., Kessler, R., Shapiro, J.M., 2021. The Effect of SNAP on the Composition of Purchased Foods: Evidence and Implications. American Economic Journal: Economic Policy 13, 277–315. doi:10/gnp9t9.

Hut, S., 2020. Determinants of Dietary Choice in the US: Evidence from Consumer Migration. Journal of Health Economics 72, 102327. doi:10/gnp9t7.

Hut, S., Oster, E., 2022. Changes in household diet: Determinants and predictability. Journal of Public Economics 208, 104620. doi:10.1016/j.jpubeco.2022.104620.

Jäntti, M., Jenkins, S.P., 2015. Income Mobility. Handbook of Income Distribution 2, 807–935. doi:10.1016/B978-0-444-59428-0.00011-4.

Lleras-Muney, Lichtenberg, 2005. Are the More Educated More Likely to Use New Drugs? Annals of Economics and Statistics , 671–696doi:10.2307/20777592.

Oster, E., 2018. Diabetes and Diet: Purchasing Behavior Change in Response to Health Information. American Economic Journal: Applied Economics 10, 308–348. doi:10.1257/app.20160232.

Rothstein, J., 2019. Inequality of Educational Opportunity? Schools as Mediators of the Intergenerational Transmission of Income. Journal of Labor Economics 37, 85–123. doi:10.1086/700888.

Springmann, M., Sacks, G., Ananthapavan, J., Scarborough, P., 2018. Carbon pricing of food in Australia: an analysis of the health, environmental and public finance impacts. Australian and New Zealand Journal of Public Health 42, 523–529. doi:10.1111/1753-6405.12830.

Waldkirch, A., Ng, S., Cox, D., 2004. Intergenerational Linkages in Consumption Behavior. The Journal of Human Resources 39, 355. doi:10.2307/3559018.

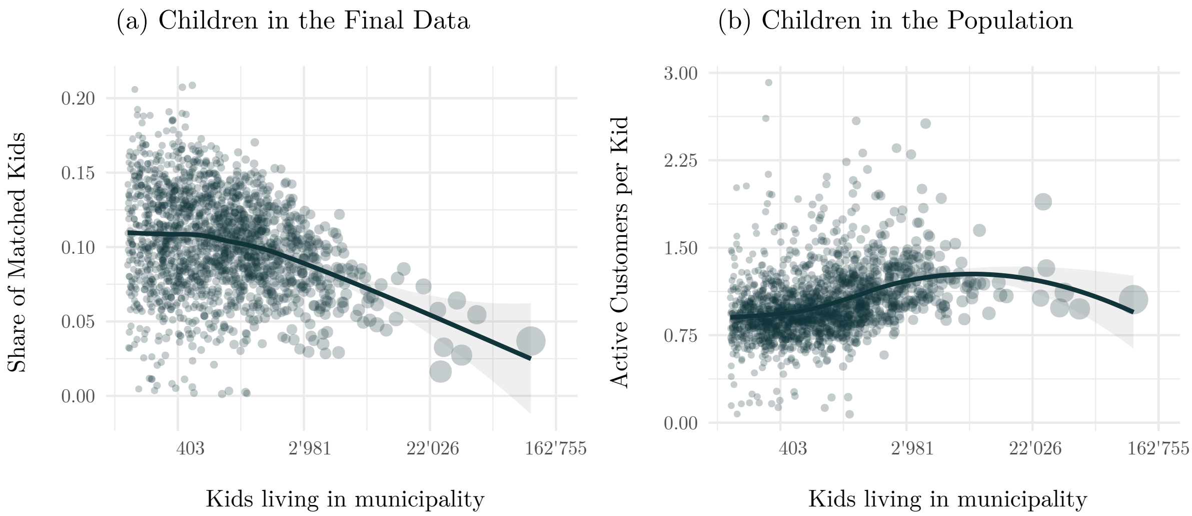
A Data: Additional Summary Statistics

Table A1: Summary Statistics for Kids

	Final Sample		Population	
Panel a)	Mean	SD	Mean	SD
Age	44.10	10.85	43.82	11.70
Age father	72.05	9.81	71.18	10.35
Age mother	71.45	10.66	70.97	11.36
Income Total	144.20	129.54	129.68	109.05
Income Adjusted	83.30	79.61	81.60	64.85
Panel b)	Pct.	N	Pct.	N
Gender		192,814		2,276,376
Female	54.4	104,844	50.8	1,155,500
Male	45.6	87,970	49.2	1,120,876
Marriage		192,814		2,276,376
Married	65.1	125,443	50.3	1,144,923
Not Married	34.9	67,371	49.7	1,131,453
Highest Education		138,477		1,554,457
Tertiary	53.6	74,264	50.0	777,526
Secondary	42.8	59,222	44.6	694,008
Elementary	3.6	4,991	5.3	82,923
Language Region		192,622		2,273,913
French	19.5	37,647	22.0	500,058
German	76.7	147,815	72.3	1,643,900
Italian	3.7	7,160	5.7	129,955
Pop. Density		192,622		2,273,913
Rural	26.3	50,571	21.6	490,575
Suburban	57.2	110,098	52.2	1,186,021
Urban	16.6	31,953	26.3	597,317
Household Size		192,814		2,276,376
1	8.5	16,465	21.0	478,402
2	26.4	50,994	33.1	754,521
3-4	51.7	99,773	37.2	846,187
5+	13.3	25,582	8.7	197,266
Observations		192,814		2,276,376

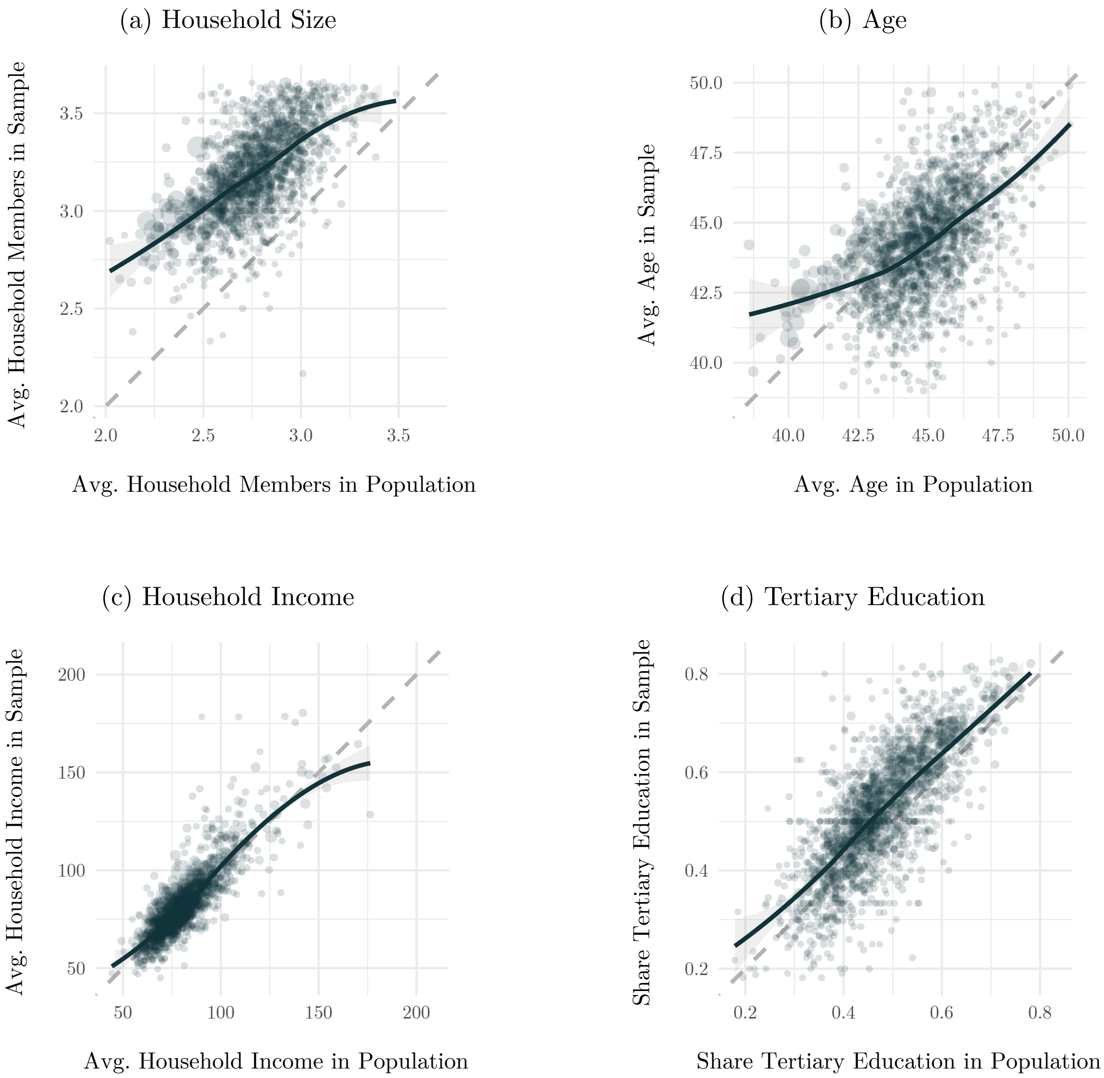
Notes: This table shows summary statistics for the customers uniquely linked to a kid in the administrative data and for the entire population of kids. *Income Total* is a household's average labor market income 2016-2020 in 1,000 CHF, and *Income Adjusted* adjusts household income by the square root of household size. *Highest Education* is the highest education anyone within the household completed, and *Pop. Density* is defined by the municipality's population density.

Figure A1: Match Rate



Notes: The figure illustrates the representativeness of the retailer's loyalty program. To this end, Figure A1a shows the share of matched kids relative to all kids living in this municipality. Figure A1b shows for the full customer data the number of active customers relative to their municipality's number of children. Each dot represents a municipality's value, while the size indicates the municipality's population. The solid line shows a local regression.

Figure A2: Municipality Averages: Kids in the Sample vs. Population



Notes: The figure illustrates the representativeness of the matched final data. To this end, we compare kids that could be uniquely matched between the administrative and consumption data to the entire Swiss population of kids. Each dot represents a municipality's average, while the dot's size indicates the municipality's population. The blue line shows a local regression. The dashed line is the 45-degree line. *Household Size* is the count of members living in an average household, *Age* is the average age of all kids in this municipality, *Household Income* is the average household labor market income, and *Tertiary Education* is the average share of households with at least one member having a tertiary degree.