

A Service of

ZBW

Leibniz-Informationszentrum Wirtschaft Leibniz Information Centre for Economics

Bühler, Jonas; Schaltegger, Christoph A.; Häner, Melanie

## Conference Paper Social Persistence and its drivers - An Analysis of Sibling Correlation in Switzerland

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2024: Upcoming Labor Market Challenges

### Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

*Suggested Citation:* Bühler, Jonas; Schaltegger, Christoph A.; Häner, Melanie (2024) : Social Persistence and its drivers - An Analysis of Sibling Correlation in Switzerland, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2024: Upcoming Labor Market Challenges, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

This Version is available at: https://hdl.handle.net/10419/302378

#### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

#### Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



# WWW.ECONSTOR.EU

# Social Persistence and its drivers - An Analysis of Sibling Correlation in Switzerland

Jonas Bühler <sup>1</sup>, Melanie Häner <sup>1</sup>, and Christoph A. Schaltegger <sup>1</sup>, \*\*\*\*

<sup>1</sup>Institute for Swiss Economic Policy (IWP) at the University of Lucerne <sup>2</sup>University of St. Gallen

Preliminary version - February 28, 2024

#### Abstract

This study unravels the impact of family background on long-run income based on sibling correlation in Switzerland. Utilizing variance decomposition analysis and two-level linear mixed models, we demonstrate that family background accounts for 14% of the variation in long-run income, with significant age and gender disparities. Our comprehensive approach sheds light on the underlying factors influencing sibling correlations. Subsequently and simultaneously testing the effect of various individual and family-specific factors on the similarity among siblings enriches the existing literature. Notably, parental income emerges as a statistically significant driver of sibling correlations, while other family-specific effects such as parental civil status or nationality have limited explanatory power.

*Keywords*: social mobility, sibling correlation, family background, inequality of opportunity

JEL Classification: I30, J62, J12

\*Authors are listed in alphabetical order. Corresponding authors: Institute for Swiss Economic Policy at the University of Lucerne, Obergrundstrasse 9, CH-6003 Lucerne, Switzerland. E-mail: jonas.buehler@iwp.swiss, melanie.haener@iwp.swiss. We thank participants of the Silvaplana Workshop on Political Economy 2023, the Economics Research Seminar of the University of Bayreuth and the Young Swiss Economists Meeting 2024 for their valuable comments on this paper.

## 1 Introduction

The 35<sup>th</sup> president of the United States, John F. Kennedy, had three brothers among his eight siblings. Joseph Junior, the eldest brother, tragically perished in a plane crash in 1944. However, the remaining brothers all embarked on successful political careers. Robert Kennedy achieved distinction as the Attorney General and U.S. Senator from New York, while Edward M. Kennedy left an indelible mark with his 47-year tenure as a U.S. Senator. The intriguing question arises: How do siblings follow such parallel paths and attain success? Does it reflect unequal opportunities across families or is it simply a result of their collective effort and talent? Sociologists and economists have long pondered these and similar inquiries. Solon (1999) concluded: "The mystery of what underlies the considerable resemblance between brothers in their long-run earnings remains a fascinating puzzle and should be a priority for continuing research." Today, almost 25 years later, this mystery is still largely unsolved.

While intergenerational parent-child mobility<sup>1</sup> is often measured through indicators like income or educational attainment, an alternative approach involves examining sibling correlations. They are considered as an omnibus variable describing the importance of family background (Solon, 1999). Siblings-based estimates usually show a stronger family effect than single-variable parent-child mobility estimates because siblings share arguably more immutable circumstances than e.g. parental income, such as schools, neighborhoods or friends (Björklund and Jäntti, 2020).

While most studies on social mobility focus on assessing societal permeability, understanding the drivers of family influence is equally important and politically relevant. Decomposing sibling correlations helps uncover the factors contributing to upward and

<sup>&</sup>lt;sup>1</sup>See e.g., Chetty et al. (2014), Jäntti and Jenkins (2015), or Corak (2013) for income mobility and Hertz et al. (2008), or Black and Devereux (2010) for educational mobility.

downward mobility within a society.

In this study, we examine the impact of family background on long-run income outcomes among siblings in Switzerland. Our goal is to gain insights into the factors that contribute to the similarity of siblings' incomes compared to individuals outside the family. We make a twofold contribution to the existing literature:

In our first contribution, we examine the baseline variance estimations without any further covariates to assess the distribution of long-run incomes within and between families. Our analysis uncovers that family background accounts for 14% of the total income variance among siblings. Moreover, we find gender-specific variations in the sibling correlation coefficient, with brothers exhibiting a higher correlation (32%) compared to sisters (22%). Additionally, educational attainment shows a stronger correlation among siblings (ICC of 0.26), highlighting the impact of family background on educational outcomes.

Second, we delve deeper into the drivers of the sibling correlation by controlling for family specific factors such as parental income or parental civil status. This decomposition approach allows us to better understand the factors responsible for income similarities among siblings. We extend the baseline model by sequentially introducing various variables to the model. While we selectively test some of these variables together, our approach allows us to examine their joint influence on income similarities among siblings.

Our analysis uncovers significant insights into the drivers of sibling correlations. Parental income explains approximately 4.3% of the sibling correlation and around 3.5% of the variance between families. Notably, the influence of parental income in Switzerland is comparatively lower than in other countries, indicating country-specific variations in the relationship between parental income and sibling outcomes. We further investigate the impact of family-specific factors, such as parental income, nationality, and parental marital status, by simultaneously including them in the baseline model. However, their joint explanatory power is limited, explaining less than 2% of the overall variance in long-run income. By systematically examining various family and individual factors, our study offers a nuanced understanding of the drivers of sibling correlations and contributes to the existing literature on income outcomes among siblings.

Our analysis is based on data from the Swiss Household Panel covering the period from 1999 to 2021, enabling us to account for individuals' childhood circumstances. To investigate the drivers of family background effects, we employ linear mixed models with both random and fixed effects, utilizing restricted maximum likelihood (REML) estimation. This approach allows us to model and decompose the factors contributing to family background effects and obtain robust estimates of their impact on long-run income outcomes among siblings.

The remainder of this paper is structured as follows: Section 2 consists of a literature review, followed by the methodological part (Section 3), which also describes the data. Section 4 contains the results and evaluations of the models. Finally, Section 5 discusses the results, draws conclusions, and grants an outlook for future research.

## 2 Literature Review

Studies on intergenerational social mobility evaluate the degree of intergenerational transmission of social status by examining various indicators, such as income, wealth, occupational position, education, and political position (Black and Devereux, 2011; Corak, 2013; Solon, 2018). However, traditional economic research primarily investigates intergenerational associations between parents and children rather than between siblings (Bügelmayer and Schnitzlein, 2018). Comparable results of the estimated intergenerational elasticity of income (IGE) for a large number of countries range from about 0.15 for Scandinavian countries to about 0.5 for the United States (Corak, 2013). Björklund and Jäntti (2020) conduct a comprehensive review of studies examining both sibling correlations and parentchild correlations. Their findings indicate that sibling correlations in education and income tend to be considerably higher than parent-child correlations. They conclude that parental socioeconomic status can only partially account for the observed sibling correlation. The measured sibling correlations vary in dependence of the siblings' sexes. E.g., Österbacka (2001) estimates brother correlations in earnings in Finland to be 0.26, sister correlations to be 0.11, and mixed sibling correlation to be 0.13. Most studies focus either on brothers or sisters only due to the differences in labor market attachment between brothers and sisters (Schnitzlein, 2014).<sup>2</sup>

A few studies have already attempted to investigate the drivers of sibling correlation. However, they are limited to explore the explanatory content of parental income, educational attainment, or occupation status. E.g., Mazumder (2008) shows that about 36% of brother correlations are explained by paternal income. Björklund et al. (2010) find similar evidence for Sweden: Their sibling correlation implies around 20% of the variation in longrun income is attributed to factors mixed siblings share. After controlling for the father's income, the study reveals a 13% decrease in the family variance component and a 10% decrease in the sibling correlation. Simultaneously considering parental income, education, and occupation results in a 26% decrease in the family variance component and a 22% decrease in the sibling correlation. These findings suggest that parental socioeconomic status explains only about one-quarter of the observed sibling correlation, suggesting that additional familial factors play a significant role in shaping the similarities in long-term incomes among siblings (Björklund et al., 2010).

For Germany, Schnitzlein (2014) finds that 43% of inequality in permanent income, expressed as brother correlation, is attributable to family background. In a country com-

<sup>&</sup>lt;sup>2</sup>Assuming, the equalization in the labor market between men and women will continue to develop in the direction of gender equality, differences in labor market attachment will probably narrow. Therefore, in our opinion, looking exclusively at brothers or sisters and omitting mixed sexes siblings would deprive the whole research of an exciting and informative dimension of how family background and environment shape sibling similarities.

parison including Germany, the United States, and Denmark, the latter displays the lowest values for the family variance component as well as for the sibling correlations for brothers and sisters<sup>3</sup> (Schnitzlein, 2014).

There is also evidence regarding the family's influence on children's educational attainment: In most Western countries, the family background is responsible for 40% to 60% of total inequality in years of schooling (Björklund and Salvanes, 2011; Bredtmann and Smith, 2018).

Research on intergenerational social mobility in Switzerland is limited (Häner and Schaltegger, 2021b). Chuard and Grassi (2020) show that income mobility in Switzerland is higher (0.14) than in the US and even higher than in Nordic countries with significant regional differences, expressing parent-child correlation in percentile rankings. One can use the latter findings for international intergenerational income mobility comparison since there are few studies for Switzerland. In this context, we also refer, as Björklund and Jäntti (2020), to Corak et al. (2014) and Jäntti and Jenkins (2015), focusing on differences between the intergenerational correlation (IGC) and the rank correlation. At the same time, educational mobility is significantly lower (Bauer, 2006; Chuard and Grassi, 2020). Furthermore, Häner and Schaltegger (2021a), examining multigenerational social mobility in Switzerland over 15 generations based on surnames, show that the family influence on social status dissipates over three generations. To the best of our knowledge, sibling correlations have not been derived for Switzerland so far.

With our study, however, we are not only the first in Switzerland to look at similarity between siblings. Rather, we want to contribute more to the literature on social mobility by disentangling the influence of family background by decomposing the variation in long-run income. Our study aims to extend beyond examining the explanatory influence of parental income alone. Instead, we aim to investigate additional factors that contribute to sibling

<sup>&</sup>lt;sup>3</sup>Brother correlation in income: US=0.450, Germany=0.432, Denmark=0.202

similarity, thereby enabling us to identify the key determinants of social mobility.

## 3 Methods and Data

#### 3.1 Sibling Correlation

To estimate the sibling correlation, we use a linear mixed model based on the framework proposed by Solon et al. (1991) and Solon (1999). The sibling correlation serves as an omnibus variable measuring the importance of belonging to a particular family (Solon, 1999). It includes factors siblings share, such as parental income, parental education, the mother's age at birth of the first child, common neighborhood, or the family structure.

Income y of the  $i^{th}$  sibling in the  $j^{th}$  family can be decomposed according to the following equation:

$$y_{ij} = \beta_{0j} + \epsilon_{ij},\tag{1}$$

where  $\beta_{0j}$  corresponds to the family intercept term and  $\epsilon_{ij}$  is the error term.

The family intercept term  $\beta_{0j}$  is composed by a fixed component  $\beta_{00}$  and a random component  $a_{0j}$ , according to Eq.(2).  $a_{0j}$  captures the permanent component of an individual's status that is shared among siblings in the same family.

$$\beta_{0j} = \beta_{00} + a_{0j} \tag{2}$$

By substituting  $\beta_{0j}$  from Eq.(2) in Eq.(1), we obtain Eq.(3), which combines the fixed and random components. As in Eq.(3), it is commonly assumed that the residuals,  $a_{0j}$  and  $\epsilon_{ij}$ , are normally distributed and independent of each other.<sup>4</sup>

<sup>&</sup>lt;sup>4</sup>This assumption allows for the conceptual separation of the permanent component into two parts: one that is perfectly correlated among siblings and another that is perfectly uncorrelated among siblings (Mazumder, 2008).

$$y_{ij} = \beta_{00} + a_{0j} + \epsilon_{ij} \tag{3}$$

Analyzing the variances in Eq.(3), while considering the constancy of the grand mean  $\beta_{00}$ , yields Eq.(4). It demonstrates the variance of  $y_{ij}$  as the sum of the variance of the random family-specific component  $a_{0j}$  and the variance of the individual error term  $\epsilon_{ij}$ .

$$Var(y_{ij}) = Var(a_{0j}) + Var(\epsilon_{ij}) = \tau_{00family} + \sigma^2$$
(4)

Thus, the total variance in income corresponds to the sum of the variance between families ( $\tau_{00family}$ ) and the variance within families ( $\sigma^2$ ). As a result, the sibling correlation is derived as follows:

$$\rho = \frac{\tau_{00family}}{\tau_{00family} + \sigma^2} \tag{5}$$

The intraclass correlation coefficient (ICC) in Eq.(5) shows the proportion of variation in the siblings' income that can be attributed to family components as a share of the total variance in siblings' income (see Appendix B for a formal derivation of  $\rho$ ). A lower sibling correlation indicates less impact of family background in shaping individuals' long-run incomes.

In order to determine the sibling correlation in long-run income,  $\rho$ , we employ a linear mixed-effects model<sup>5</sup> that allows for the inclusion of control variables. Specifically, we estimate the following 2-Level random intercept model:

$$y_{ij} = \beta \mathbf{X}_{ij} + a_{0j} + \epsilon_{ij}, \tag{6}$$

<sup>&</sup>lt;sup>5</sup>We use linear mixed models using restricted maximum likelihood (REML) estimates by applying the lmer function in the lme4 Package to identify the parameters. (Bates et al., 2015). p-values are provided in summary tables via Satterthwaite's degrees of freedom method (Kuznetsova et al., 2017).

where the vector  $\mathbf{X}_{ij}$  captures all fixed effects, whereas  $a_{0j}$  again corresponds to the random family component and  $\epsilon_{ij}$  to the error term according to Eq.(3)<sup>6</sup>. First, we estimate an "empty model" without including any control variables in the X-vector. This allows us to capture the baseline sibling correlation and the inherent within-family and betweenfamily variation.

Second, we introduce specific fixed effects to the vector  $\mathbf{X}_{ij}$ , such as sex, age, or parental income, individually, to assess their effect on the reduction of the variance components. This allows us to determine the respective explanatory power of different individual and family-related factors to the similarity of siblings.

In addition to the subsequent inclusion of control variables, we thirdly analyze models that include multiple covariates simultaneously. This approach allows us to examine the combined effect of multiple factors on the variance components and sibling correlation.

## 3.2 Data - Swiss Household Panel (SHP)

The Swiss Household Panel (SHP)<sup>7</sup> serves as the foundation for the data in this study. It is a comprehensive longitudinal panel study that has been conducted annually since 1999, employing random sampling to survey private households and individuals (SHP Group, 2022). The SHP is a large-scale study with a nationally representative scope<sup>8</sup> (Tillmann et al., 2022). Its primary focus is to observe social change, particularly the dynamics of changing living conditions and social representations within the Swiss population (Voorpostel et al., 2021).

The SHP is designed as an indefinite life panel study<sup>9</sup>. This design ensures that the

<sup>&</sup>lt;sup>6</sup>The model does not contain a transitory error component as we use long-run income directly in model (6). In line with Björklund et al. (2010), we will not use annual income and, therefore, not include the transitory error component in the models.

 $<sup>^{7}</sup>n = 9'828$  households and n = 15'882 persons interviewed in 2020

<sup>&</sup>lt;sup>8</sup>For more information, see Appendix A1 and Appendix A2

<sup>&</sup>lt;sup>9</sup>The design and content of the SHP draw upon insights from social science research and experiences from panel surveys in Europe (such as GSOEP) and North America (PSID)

same households and individuals are interviewed on an annual basis. The study consists of four samples<sup>10</sup>, which were drawn by the Swiss Federal Statistical Office (Voorpostel et al., 2021).

The SHP covers a wide range of fields and topics, including demographic, political, economic, psycho-social, life course, and health characteristics and perceptions. This breadth of coverage makes the SHP a valuable resource for studies and enables cross-domain analyses.(Voorpostel et al., 2021). For all estimates in this paper, we exclusively use the data set "Living in Switzerland Waves 1-22 + Beta version wave 23 + Covid 19 data" provided by the Swiss Centre of Expertise in the Social Sciences (SHP Group, 2022). Our analyses focus on survey waves spanning the years 1999 to 2021, incorporating a comprehensive range of variables from various domains.

### 3.3 Variable Selection

We restrict our sample to individuals aged at least 26 years old and at most 41 years in each survey wave. Inflation-adjusted long-run income, in 2021 prices (CHF), is the main outcome variable. To adjust prices, we employ the R package priceR<sup>11</sup> developed by Condylios (2022). The income variable<sup>12</sup> must be recorded at least once in the entire timespan. Following Mazumder (2008), we set a minimum income threshold of 500 USD in 1979 prices. We convert this threshold value into 2021 USD and further express it in CHF,

 $<sup>^{10}</sup>$ SHPI includes 5,074 households and 7,799 individuals interviewed for the first time in 1999; SHPII includes 2,538 households and 3,654 individuals interviewed for the first time in 2004; SHPIII includes 3,989 households and 6,090 individuals interviewed for the first time in 2013; SHPIV includes 4,380 households and 7,557 individuals interviewed for the first time in 2020

<sup>&</sup>lt;sup>11</sup>Inflation adjustment calculations in the package are based on the theoretical framework presented in Principles of Macroeconomics by Gregory Mankiw et al (2014), as referenced by (Condylios, 2022)

<sup>&</sup>lt;sup>12</sup>Gross yearly total personal income, constructed with working income (from employment or from selfemployment), old age and disability pensions, public transfer income (income from institutions), income from private persons, other sources (capital income, income from rents, inheritance, 3rd pillar) Kuhn (2021).

resulting in approximately 1,660 CHF<sup>13</sup>. The identification number used in the panel study serves to identify the individual persons across different waves. We utilize this variable to identify siblings primarily through their mother and, if unavailable, through their father, thereby increasing the number of observations. No distinction is made on whether the children are adopted, biological siblings or half-siblings.<sup>14</sup>

We only include explanatory variables<sup>15</sup> with a low share of missing values, to keep the number of imputations as low as possible. We used the "mice" package in R to perform multiple imputations, enabling us to impute missing values for both continuous numerical and categorical data.<sup>16</sup> The mice package implements multiple imputations by chained equations, which is a widely used approach for handling missing data. It generates imputations by iteratively imputing each variable using the specified imputation method while taking into account the relationships with other variables. This process is repeated until convergence is reached (van Buuren and Groothuis-Oudshoorn, 2011).

We construct long-run parental income as the average from the mother's and father's incomes from 1999-2021. We set an upper age limit for parents to the regular pension age of 65 for men and 64 for women in Switzerland. Common literature such as Chetty et al. (2014), measures the children's income when they are about 30 years old and the parent's income approximately 15 years prior. Applying this approach with the setup of

<sup>&</sup>lt;sup>13</sup>By applying this income threshold, the observation number decreases by n=16, including the individuals with income above the income limit but having siblings with income <1660 CHF. They are also excluded from the analyses, as they would otherwise mistakenly be considered as singletons

<sup>&</sup>lt;sup>14</sup>Analysis of parental identification numbers revealed that the data set consists of a total of 373 distinct families. Among these families, we find that both parental information is available and identical for all siblings in  $n_1 = 326$  families. Additionally, there are  $n_2 = 43$  families where only information about the mother is available, and  $n_3 = 5$  families where information about only the father is available. The summation is compromised by inaccuracies due to the availability of complete parental information for one sibling, while for their other siblings, only one parent's information is known.

<sup>&</sup>lt;sup>15</sup>For the variables used in the analysis, we employ the rounded annual average of political position on a scale from 0-10 (where higher values indicate a more right-wing position). To determine nationality, we take the first variable going backward from 2021 by years. Similarly, for working status, we go backward from 2019 due to the unique circumstances on the labor market caused by the Covid-19 pandemic.

<sup>&</sup>lt;sup>16</sup>After testing different imputation methods (pmm, polyreg, logreg), we found that the random forest imputation method provided the most accurate results, based on an analysis of descriptive statistics before and after imputations.

the SHP, which has only been collected since 1999, results in an analysis of only children's long-run income from approx. 2014 to 2021 and following parents' long-run income from approx. 1999-2006. Because of data limitations, the number of observations would have fallen sharply with this approach. Therefore, we use restrictions in which data collection of parental and child income overlap. However, not being able to keep the common gap of 15 years ends up with the analysis containing many parents whose incomes are recorded at a time being much older than the children when incomes are measured.<sup>17</sup>

#### **3.4** Descriptive Statistics

These descriptive statistics provide an overview of the final sample, highlighting key demographic and socioeconomic characteristics of the participants. In the final sample (1), a total of 818 siblings are identified. Of these, 414 are men (50.6%) and 404 are women (49.4%). Among the sibling relationships, 262 consist of only males, and 246 consist of only females. In the mixed sexes sample, the mean long-run income from 1999-2021 is 65,310 CHF. Brothers have a higher income (mean = 75,590 CHF) compared to sisters (mean = 58,660 CHF). Regarding other characteristics of the participants, the average age in 2021 is 35.32 years, with an interquartile range (IQR) of 31.00-39.00. When examining only the brothers and sisters separately, brothers have a slightly higher mean age of 36.43 years, while sisters have a slightly lower mean age of 34.60 years. The average years of education for the entire sample is 15.45 years. Brothers have a slightly lower mean of 15.19 years, while sisters have a slightly higher mean of 15.84 years. The individuals in the sample report an average of 6.39 monthly contacts with relatives. Regarding working status, the majority of participants are occupied (92.2% overall). A small percentage of the individuals are unemployed (1.3% overall), with slightly higher percentages among broth-

<sup>&</sup>lt;sup>17</sup>Administrative personal income data suggest men aged 50-65 earn almost 18% more in gross median income than men aged 30-39 and nearly 48% more than men aged 20-29. (Federal Statistical Office, 2022).

ers (2.1%) compared to sisters (1.2%). Additionally, 6.5% of the individuals in the sample are not in the labor force. In terms of nationality, the vast majority of our study sample (95.2% overall) is Swiss. When it comes to the political position, the largest proportion of the individuals identify as center (54.4% overall). A significant proportion identify as left (27.3% overall), with slightly higher percentages among sisters compared to brothers. A smaller proportion identify as right (18.3% overall), with higher percentages among brothers (26.0%). Regarding the presence of own children, 29.6% of the analyzed individuals do have children themselves.

	Full S	Sibling Sample	_	Brothers	_	Sisters
Long-run Income, mean (IQR)	65.31	(47.25-81.98)	72.59	(58.52 - 86.92)	58.66	(41.72 - 74.03)
Sex. n(%)	818	(100.0)	262	(100.0)	246	(100.0)
Male	414	(50.6)		()		()
Female	404	(49.4)				
Age, mean (IQR)	35.32	(31.00-39.00)	36.43	(32.00-41.00)	34.60	(30 - 38)
Height in cm, mean (IQR)	172.9	(166.0-179.9)	178.1	(174.0-183.0)	167.0	(163.0-171.0)
Years of Education, mean (IQR)	15.45	(12.00-19.00)	15.19	(12.00 - 19.00)	15.84	(12.00-19.00)
Contact Relatives, mean (IQR)	6.39	(3.00-8.00)	6.16	(2.00-8.00)	7.41	(3.00-10.00)
Working Status, n(%)						
Occupied	754	(92.2)	246	(93.8)	224	(91.1)
Unemployed	11	(1.3)	5	(2.1)	3	(1.2)
Not in Labor Force	53	(6.5)	11	(4.1)	19	(7.7)
Nationality, $n(\%)$						
Swiss	779	(95.2)	249	(95.0)	227	(92.3.0)
Non-Swiss	39	(4.8)	13	(5.0)	19	(7.7)
Political Position, $n(\%)$						
Left	223	(27.3)	46	(17.6)	105	(42.7)
Center	445	(54.4)	148	(56.4)	124	(50.4)
Right	150	(18.3)	68	(26.0)	17	(6.9)
Own Children, n(%)						
Yes	242	(29.6)	80	(29.5)	161	(65.5)
No	576	(70.4)	182	(69.5)	85	(34.6)

Table 1: Descriptive statistics: full sibling sample, brother sample, sister sample

*Notes*: IQR; interquartile range, Long-run Income; expressed in 2021 CHF (in 1,000), Contact Relatives; Times per month; Age measured in 2021

## 4 Results

# 4.1 Mixed Sexes Sibling Correlation, Brother Correlation, Sister Correlation

Table 2 shows the sibling correlation and variance component estimates using linear mixed models for long-run income. The full sibling sample model (column 1) yields a value for the intra-class correlation coefficient of 0.14 (n=818, 373 families). This indicates that, on average, the family background explains 14% of the total variance in income.

As the comparison between the within-family component ( $\sigma^2$ ) and the between-family component ( $\tau_{00family}$ ) reveals, the variance within families is more than 6 times bigger than the variance between families. In columns two and three of Table 2, we limit the analysis to brothers or sisters only, respectively. The influence of family background on the later success of brothers (0.32) is more important than that for sibling relationships consisting exclusively of women (0.22). This is in line with previous findings (see Section 2 for the Literature Review).

For comparison, the same estimates for height (in cm) are presented in Table A1 in Appendix A to provide a context for the initial estimates. It is widely acknowledged that genetics play a significant role in determining body size. Therefore, analyzing the influence of family imprinting on height variation could help in interpreting income results. Table A1 demonstrates that the family origin of siblings accounts for around one-third of the variance in height (0.26, 0.38, 0.30 for all siblings, brothers only, and sisters only, respectively), which is equivalent to the family background shared by siblings. The remaining two-thirds can be attributed to other environmental factors or different genetics. Family background has a more substantial impact on purely physical attributes such as height compared to economic outcomes like income. However, the difference in brother correlations between income and

	Full Sibling Sample	Brothers	Sisters
Long-run Income	65.44***	72.88***	58.70***
	(1.08)	(2.05)	(1.75)
$\sigma^2$	702.40	560.68	476.24
$ au_{00family}$	112.56	262.66	132.19
ICC	0.14	0.32	0.22
Obs.	818	262	246
Nb families	373	127	119

height is only approximately six percentage points.

Table 2: Sibling correlations in long-run income (CHF)

Notes: Significance Codes: '\*\*\*' 0.01 '\*\*' 0.05 '\*' 0.1

The standard errors (SE) are shown in parentheses.

Long-run income; Inflation adjusted in 2021 prices (in 1,000 CHF) of an individual from a randomly chosen family, randomly selected from the whole sample.

 $\sigma^2$  represents the estimated residual variance, which measures the within-family variation in long-run income.

 $\tau_{00family}$  represents the estimated variance component at the family level, capturing the between-family variation in long-run income..

ICC stands for intraclass correlation coefficient and indicates the proportion of total variation in long-run income that can be attributed to differences between families.

The table provides information on the number of observations (Obs.) and the number of families (Nb. families) included in the analysis for each sibling group.

The model was estimated using the linear mixed-effects modeling approach (lmer function) from the lme4 package in R. See Bates et al. (2015)

### 4.2 Drivers of sibling correlation

Consistent with prior studies, we further examine the impact of parental, family-specific characteristics on the outcome variable. As depicted in the fifth model, parental income has a positive effect, explaining approximately 4.3% of the intra-class correlation and about 3.5% of the variance between families. Thus, the explanatory power of parental income

is significantly lower in Switzerland than in other countries (see Section 2). This aligns with the finding that the association between parent-child income in Switzerland is notably lower when compared internationally (Chuard and Grassi, 2020).

Our analysis further enables us to determine the explanatory strength of additional family-specific factors. Table 3 demonstrates that variables such as nationality or parental marital status have no statistically significant impact on the outcome variable.

	baseline	parental income	nationality	parental civil status
FE Intercept	$65.44^{***}$	60.92***	60.47***	65.41***
	(1.08)	(2.37)	(4.83)	(1.13)
FE Estimate	-	0.06**	5.22	0.41
	-	(0.03)	(4.94)	(3.79)
$\sigma^2$	702.40	702.19	700.93	702.22
$ au_{00family}$	112.56	108.66	114.21	113.94
ICC	0.14	0.13	0.14	0.14
Obs.	818	818	818	818
Nb families	373	373	373	373
Comparison				
$\%\Delta$ ICC	-	-4.28	0.08	-0.28
$\%\Delta au_{00family}$	-	-3.46	1.47	1.23

Table 3: Drivers of sibling correlations

Notes: Significance Codes: '\*\*\*' 0.01 '\*\*' 0.05 '\*' 0.1

The standard errors (SE) are shown in parentheses.

FE Intercept represents the fixed effects estimates.

FE Estimate shows the fixed effects estimates for each variable.

 $\%\Delta$  represents the percentage change in the ICC compared to the baseline model for each subsequent model.

 $\%\Delta\tau_{00family}$  represents the percentage change in the between-family variance component compared to the baseline model for each subsequent model.

#### 4.3 Explanatory power of family-specific effects

The within-family variance ( $\sigma^2$ ), as listed in Table 3, serves as a useful indicator for identifying within-family differences in the outcome variable. If the inclusion of a fixed effect affects primarily the between-family variance ( $\tau_{00family}$ ) while leaving the within-family variance ( $\sigma^2$ ) relatively stable, it suggests that siblings share this particular aspect. This observation is applicable not only to parental income but also to parental marital status and the nationality of the siblings, as demonstrated in Table 3. Simultaneously considering these effects and comparing them with the baseline model, allows us to assess the joint explanatory power of the family-specific effects. Table 4 presents the corresponding results.

Once again, the statistical analysis reveals that only parental income has a significant impact, whereas nationality and parental marital status do not exhibit any significant influence on income.

Consequently, when considering these specific fixed effects, the combined explanatory power of family-specific factors amounts to less than 2% of the overall variance in income.<sup>18</sup>

<sup>&</sup>lt;sup>18</sup>The statistical insignificance of the respective regression coefficients might contribute to these findings.

	baseline	family-specific aspects
FE Intercept	$65.44^{***}$	56.06***
	(1.08)	(5.30)
FE Parental Income		0.06**
		(0.03)
FE Nationality		5.06
		(4.97)
FE Parental Civil Status		1.08
		(3.81)
$\sigma^2$	702.40	700.56
$ au_{00family}$	112.56	111.72
ICC	0.14	0.14
Obs.	818	818
Nb. families	373	373
Comparison		
$\%\Delta$ ICC	-	-1.76
$\%\Delta \  au_{00family}$	-	-0.75

Table 4: Explanatory power of family-specific aspects

*Notes*: Significance Codes: '\*\*\*' 0.01 '\*\*' 0.05 '\*' 0.1 The standard errors (SE) are shown in parentheses.

FE Intercept represents the fixed effects estimates.

FE Estimate shows the fixed effects estimates for each variable.

 $\%\Delta$  represents the percentage change in the ICC compared to the model with fammily-specific aspects.

 $\%\Delta\tau_{00family}$  represents the percentage change in the between-family variance component compared to the model with family-specific aspects.

When considering the three fixed effects of parental income, nationality, and parental marital status, their explanatory power in terms of the between-family variance component is limited, explaining only a small fraction of the overall variation. This suggests that factors other than these specific family characteristics play a more substantial role in accounting for income differences between families. In addition to that, it is noteworthy that the variance in income within families is six times higher than the variance between families. This suggests that variations between siblings within the same family contribute significantly more to the overall variation in income than differences between different families.

The analysis additionally suggests that income is not significantly dependent on nationality or parental marital status (p > 0.05), which is indicative of a positive sign in terms of equality of opportunity. This initial observation is further supported by Figure 1. The left graph displays the income distributions of all Swiss individuals in the sample alongside those of non-Swiss individuals. Similarly, the right graph compares the income distributions of siblings with divorced parents to those with non-divorced parents. In both graphs, a substantial overlap between the two groups is evident, figuratively representing the absence of statistically significant influence.



Figure 1: Histogram of income in dependence of nationality and parental civil status

*Notes*: The left graph displays the long-run income distributions (in 1,000 CHF) of all Swiss individuals in the sample alongside those of non-Swiss individuals. Similarly, the right graph compares the income distributions of siblings with divorced parents to those with non-divorced parents.

#### 4.4 Sensitivity Analysis

We test the sensitivity of our estimates using four alternative model specifications. Table 5 presents the respective results. In column (1) we investigate the familial influence on educational attainment. Instead of the long-run income, we analyze the effects on years of education. As the results reveal, the ICC is substantially higher for educational attainment than for long-run income (0.26 vs. 0.14). This is in line with Chuard and Grassi (2020) that show that the intergenerational persistence in education is substantially higher than persistence in income.

In column (2) we restrict the maximum age difference among siblings to eight years. As the results show, this limitation of age gaps does barely affect the variance estimates.

Finally, we run a sensitivity test with another age filter for the analysis in Table 5, column (3). Raising the age filter and including only siblings aged  $\geq 30$  decreases the number of observations by more than half compared to the baseline model. This disparity in the number of observations significantly reduces comparability. However, column (3) of the analysis indicates that when considering only older siblings, the family effect, as measured by the ICC, tends to increase.

	baseline	(1) Education	(2) 8 Yrs Age Diff.	(3) Age Filter $\geq 30$
Intercept	65.44***	$15.45^{***}$	65.35***	75.15***
	(1.08)	(0.14)	(1.12)	(2.01)
$\sigma^2$	702.40	8.68	684.4	925.30
$ au_{00family}$	112.56	3.11	126.8	271.04
ICC	0.14	0.26	0.16	0.23
Obs.	818	818	767	353
Nb. families	373	373	354	168

Table 5: Sensitivity analysis

*Notes*: Significance Codes: '\*\*\*' 0.01 '\*\*' 0.05 '\*' 0.1

The standard errors (SE) are shown in parentheses.

(1): Baseline model to estimate educational status for the full siblings sample. Dependant variable measured in max. years of education.

(2): Baseline model to estimate long-run income (in CHF 1,000) with an 8-year age difference within families for the full siblings sample.

(3): Baseline model to estimate long-run income (in CHF 1,000) for individuals with age  $\geq$  30 years for the full siblings sample.

## 5 Discussion and Conclusion

In this study, we examine the influence of family background on long-run income of siblings in Switzerland. Using a variance decomposition approach and employing a two-level linear mixed model, we unravel the drivers of sibling correlation and shed light on the role of the familial background in shaping individual outcomes.

The baseline estimates show that family-belonging explains 14% of the total variation in long-run income. This suggests that factors beyond the family, such as individual characteristics and external influences, contribute significantly to income differences among siblings. Moreover, the sibling correlation varies depending on the gender of the siblings, with brothers exhibiting a higher sibling correlation coefficient (32%) compared to sisters (22%). Additionally, we observe a higher correlation (ICC of 0.26) among siblings regarding the number of years of education, indicating that familial background has a stronger impact on educational attainment than on income outcomes.

Our findings contribute to the existing literature by providing empirical evidence on sibling correlation in long-run income and its dependence on various factors. We extend the literature by estimating baseline models and subsequently examining the influence of individual and family-specific factors on the sibling correlation. Notably, parental income emerges as a significant driver, aligning with the findings of Björklund et al. (2010) and Mazumder (2008). Moreover, we highlight the limited explanatory power of simultaneously tested family-specific effects, including parental income, nationality, and parental marital status. This indicates that other factors play a more substantial role in explaining the sibling correlation.

Furthermore, our findings hold implications for equality of opportunity. The low ICC of 14% suggests that Switzerland exhibits a higher degree of income mobility and social permeability, akin to Scandinavian countries and surpassing Germany or the United States,

as Schnitzlein (2014) demonstrate. This underscores the minor role of family background in explaining long-run incomes in Switzerland, aligning closely with the principle of equality of opportunity.

In light of Switzerland's context, one wonders about the Kennedy brothers — would their glorious achievements have been replicated had they grown up amidst the Swiss Alps?

Looking ahead, our study contributes to unraveling the mystery of the resemblance in long-run income among siblings, as previously emphasized by Solon (1999). Through our research, we have made strides in shedding light on this mystery and advancing our understanding of the underlying factors. Our study deepens the understanding of sibling similarities in long-run income and provides insights into the role of family background. While family-specific factors demonstrate some explanatory power, the majority of income variation in Switzerland is attributed to factors beyond the family. Exploring withinfamily differences, such as birth order and differences in upbringing, can provide valuable insights as they are factors that siblings do not share but are also part of the familial background. Future research could further investigate additional drivers that contribute to income disparities among siblings.

## References

- D. Bates, M. Mächler, B. Bolker, and S. Walker. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1):1-48, 2015. doi: 10.18637/jss.v067.i01. URL https://www.jstatsoft.org/index.php/jss/article/view/v067i01.
- P. Bauer. The Intergenerational Transmission of Income in Switzerland A Comparison between Natives and Immigrants. WWZ Discussion Paper, 2006.
- A. Björklund and M. Jäntti. Intergenerational mobility, intergenerational effects, sibling correlations, and equality of opportunity: A comparison of four approaches. *Research in Social Stratification and Mobility*, 70:100455, 2020. ISSN 0276-5624. doi: https://doi.org/10.1016/j.rssm.2019.100455. URL https://www.sciencedirect.com/ science/article/pii/S0276562419301544.
- A. Björklund and K. G. Salvanes. Education and Family Background: Mechanisms and Policies. Handbook of the Economics of Education, 3:201–247, 1 2011. ISSN 1574-0692. doi: 10.1016/B978-0-444-53429-3.00003-X.
- A. Björklund, L. Lindahl, and M. J. Lindquist. What More Than Parental Income, Education and Occupation? An Exploration of What Swedish Siblings Get from Their Parents. *The B.E. Journal of Economic Analysis & Policy*, 10(1), 2010. doi: doi:10.2202/1935-1682.2449. URL https://doi.org/10.2202/1935-1682.2449.
- S. Black and P. Devereux. Recent Developments in Intergenerational Mobility. Institute for the Study of Labor (IZA), IZA Discussion Papers, 2010. doi: 10.2139/ssrn.1587559.
- S. E. Black and P. J. Devereux. Recent Developments in Intergenerational Mobility, volume 4 Part B. CEPR, 2011.

- J. Bredtmann and N. Smith. Inequalities in Educational Outcomes: How Important Is the Family? Oxford Bulletin of Economics and Statistics, 80(6):1117-1144, 2018. ISSN 0305-9049. doi: https://doi.org/10.1111/obes.12258. URL https://onlinelibrary. wiley.com/doi/abs/10.1111/obes.12258.
- E. Bügelmayer and D. D. Schnitzlein. Is it the family or the neighborhood? Evidence from sibling and neighbor correlations in youth education and health. *The Journal of Economic Inequality*, 16(3):369–388, 2018. ISSN 1573-8701. doi: 10.1007/s10888-017-9364-8. URL https://doi.org/10.1007/s10888-017-9364-8.
- R. Chetty, N. Hendren, P. Kline, and E. Saez. Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States. *The Quarterly Journal of Economics*, 129(4):1553–1623, 2014.
- P. Chuard and V. Grassi. Switzer-Land of Opportunity: Intergenerational Income Mobility in the Land of Vocational Education. *Economics Working Paper Series*, 7 2020. URL https://ideas.repec.org/p/usg/econwp/202011.html.
- S. Condylios. priceR: Economics and Pricing Tools, 9 2022.
- M. Corak. Income Inequality, Equality of Opportunity, and Intergenerational Mobility. Journal of Economic Perspectives, 27(3):79-102, 2013. doi: 10.1257/jep.27.3.79. URL https://www.aeaweb.org/articles?id=10.1257/jep.27.3.79.
- M. Corak, M. J. Lindquist, and B. Mazumder. A comparison of upward and downward intergenerational mobility in Canada, Sweden and the United States. *Labour Economics*, 30:185-200, 2014. ISSN 0927-5371. doi: https://doi.org/10.1016/j.labeco.2014.03.013. URL https://www.sciencedirect.com/science/article/pii/S0927537114000530.
- Federal Statistical Office. Bundesamt für Statistik Schweizerische Lohnstrukturerhebung (LSE), 2022. URL https://www.bfs.admin.ch/bfs/de/home/statistiken/

arbeit-erwerb/loehne-erwerbseinkommen-arbeitskosten/lohnniveau-schweiz/ personenbezogene-merkmale.html.

- M. Häner and C. A. Schaltegger. The name says it all. Multigenerational social mobility in Switzerland, 1550-2019, 2021a. URL http://hdl.handle.net/10419/242355.
- M. Häner and C. A. Schaltegger. Fällt der Apfel weit vom Stamm?: Ein Überblick über den Forschungsstand zur intergenerationellen sozialen Mobilität. *Perspektiven der Wirtschaftspolitik*, 22(2):103–120, 2021b. doi: doi:10.1515/pwp-2020-0052. URL https://doi.org/10.1515/pwp-2020-0052.
- T. Hertz, T. Jayasundera, P. Piraino, S. Selcuk, N. Smith, and A. Verashchagina. The Inheritance of Educational Inequality: International Comparisons and Fifty-Year Trends. Advances in Economic Analysis & Policy, 7:1775, 2008. doi: 10.2202/1935-1682.1775.
- M. Jäntti and S. P. Jenkins. Income Mobility. In A. B. Atkinson and F. Bourguignon, editors, *Handbook of Income Distribution*, volume 2, chapter 10, pages 807–935. Elsevier, 2015. ISBN 1574-0056. doi: https://doi.org/10.1016/B978-0-444-59428-0.00011-4. URL https://www.sciencedirect.com/science/article/pii/B9780444594280000114.
- U. Kuhn. Collection, construction and checks of income data in the Swiss Household Panel, 2021.
- A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen. ImerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82(13):1 - 26, 2017. doi: 10.18637/jss.v082.i13. URL https://www.jstatsoft.org/index.php/jss/article/ view/v082i13.
- B. Mazumder. Sibling Similarities and Economic Inequality in the US. Journal of Population Economics, 21(3):685-701, 2008. ISSN 09331433, 14321475. URL http: //www.jstor.org/stable/40344699.

- E. Osterbacka. Family Background and Economic Status in Finland. The Scandinavian Journal of Economics, 103(3):467-484, 2001. ISSN 03470520, 14679442. URL http: //www.jstor.org/stable/3441083.
- D. D. Schnitzlein. How important is the family? Evidence from sibling correlations in permanent earnings in the USA, Germany, and Denmark. *Journal of Population Economics*, 27(1):69–89, 2014. ISSN 1432-1475. doi: 10.1007/s00148-013-0468-6. URL https://doi.org/10.1007/s00148-013-0468-6.
- SHP Group. Living in Switzerland Waves 1-22 + Beta version wave 23 + Covid 19 data [Dataset], 2022.
- G. Solon. Intergenerational mobility in the labor market. In O. Ashenfelter and D. Card, editors, *Handbook of Labor Economics*, volume 3, Part A, chapter 29, pages 1761–1800. Elsevier, 1999. URL https://EconPapers.repec.org/RePEc:eee:labchp:3-29.
- G. Solon. What Do We Know So Far about Multigenerational Mobility? The Economic Journal, 128(612):F340-F352, 2018. ISSN 0013-0133. doi: https://doi.org/10.1111/ecoj. 12495. URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ecoj.12495.
- G. Solon, M. Corcoran, R. Gordon, and D. Laren. A Longitudinal Analysis of Sibling Correlations in Economic Status. *The Journal of Human Resources*, 26(3):509-534, 1991.
  ISSN 0022166X. doi: 10.2307/146023. URL http://www.jstor.org/stable/146023.
- R. Tillmann, M. Voorpostel, E. Antal, N. Dasoki, H. Klaas, U. Kuhn, F. Lebert, G.-A. Monsch, and V.-A. Ryser. The Swiss Household Panel (SHP). Jahrbücher für Nationalökonomie und Statistik, 242(3):403–420, 2022. doi: doi:10.1515/jbnst-2021-0039. URL https://doi.org/10.1515/jbnst-2021-0039.
- S. van Buuren and K. Groothuis-Oudshoorn. mice: Multivariate Imputation by Chained Equations in R. Journal of Statistical Software, 45(3):1 – 67, 2011. doi:

10.18637/jss.v045.i03. URL https://www.jstatsoft.org/index.php/jss/article/ view/v045i03.

M. Voorpostel, R. Tillmann, F. Lebert, U. Kuhn, O. Lipps, V.-A. Ryser, E. Antal, G.-A. Monsch, N. Dasoki, H. Klaas, and B. Wernli. Swiss Household Panel Userguide (1999-2020), Wave 22. 2021.

## Appendix

## A Additional Figures and Tables



Figure A1: SHP participation numbers at the household level

Notes: Source: (Tillmann et al., 2022)



Figure A2: SHP participation number at the individual level

Notes: Source: (Tillmann et al., 2022)

	Full Sibling Sample	Brothers	Sisters
Height (cm)	172.89***	178.03***	167.04***
SE	(0.35)	(0.51)	(0.48)
$\sigma^2$	56.48	31.71	30.41
$ au_{00family}$	19.76	17.87	13.06
ICC	0.26	0.38	0.30
Obs.	818	262	246
Nb. families	373	127	119

Table A1: Sibling correlations in height (cm)

*Notes*: Significance Codes: '\*\*\*' 0.01 '\*\*' 0.05 '\*' 0.1 The standard errors (SE) are shown in parentheses.

Mean Hight (cm); expected mean height of an individual from a randomly chosen family, randomly selected from the whole sample.

The model was estimated using the linear mixed-effects modeling approach (lmer function) from the lme4 package in R. See Bates et al. (2015)

### B Relationship between the ICC and a "Correlation"

Let's consider the correlation in long-run income between two individuals within the same family. We denote them as  $y_{ij}$  and  $y_{i'j}$ , where *i* represents the individuals and *j* represents the family:

$$\operatorname{Corr}(y_{ij}, y_{i'j}) = \operatorname{Corr}(a_{0j} + \epsilon_{ij}, a_{0j} + \epsilon_{i'j})$$
$$= \frac{\operatorname{Cov}(a_{0j} + \epsilon_{ij}, a_{0j} + \epsilon_{i'j})}{\sqrt{\operatorname{Var}(a_{0j} + \epsilon_{ij}) \cdot \operatorname{Var}(a_{0j} + \epsilon_{i'j})}}$$
$$= \frac{\operatorname{Var}(a_{0j})}{\operatorname{Var}(a_{0j} + \epsilon_{ij})}$$
$$= \frac{\tau_{00family}}{\tau_{00family} + \sigma^2} = \rho.$$

In this derivation of  $\rho$ , we assume that the random effects a and  $\epsilon$  are normally distributed and independent of each other. We also consider a scenario where there are only two siblings' scores within each family.