

An Empirical Investigation of Intergenerational Mobility in Korea

Seunghee Lee*

January 2022

Click [here](#) for the most recent version

Abstract

This paper studies intergenerational income mobility in Korea using a novel approach to understanding how parental socioeconomic status predicts that of children. The conventional approach to measuring mobility is based on the correlation between parents' permanent income and children's permanent income. In contrast, we study how the trajectory of parental incomes across childhood, adolescence, and early adulthood collectively determine a child's future income. This allows one to explore how the timing of income matters beyond its long-run average as well as to understand the years where parental income is relevant. To this end, we employ a functional approach to uncover this relationship. Using KLIPS data, we find that parental income after the mid-teens is crucial for children's economic success. In contrast, parental incomes at other ages are especially powerful predictors of factors that help determine income. Specifically, teen years are essential for college attendance, while the early 20s play a crucial role in offspring's occupation. In addition, living in a single-parent family during the mid-teens tends to harm children's economic and educational success.

KEY WORDS: intergenerational mobility, functional data analysis

JEL CLASSIFICATION: C13, J62

*Department of Economics, Indiana University, Email: lee2057@iu.edu.

1 Introduction

Intergenerational mobility, while studied by sociologists and economists since the origins of their respective disciplines, has experienced a renaissance of research over the last two decades, largely driven by new data sets. See [Jäntti and Jenkins \(2015\)](#) for a survey. While much new empirical evidence has emerged, it is also clear that a new social science of mobility is still developing from the perspectives of econometric and theoretical models ([Cholli and Durlauf \(2021\)](#)).

This paper contributes to this literature by and developing a new way of measuring intergenerational income mobility. Conventional empirical approaches to mobility measure relationships between scalar measures of socioeconomic status for parents and children. In economics, the dominant form of analysis is based on linear regressions. Specifically, the regression coefficient in a linear model linking a child's permanent income and parents' permanent income is treated as the metric of intergenerational persistence, so that smaller coefficients are interpreted as evidence of greater mobility. The canonical intergenerational regression may be written as

$$y_{c,i} = \alpha + \beta y_{p,i} + \varepsilon_i \quad (1)$$

where $(y_{c,i})$ is log of child i 's permanent income, $(y_{p,i})$ is log of i 's parents' permanent income, and β is the intergenerational elasticity (IGE) of income.

This perspective on mobility can be theoretically justified via overlapping generations models of the type pioneered by [Becker and Tomes \(1979\)](#) and [Loury \(1981\)](#). The linear specification has been derived as an equilibrium law of motion by [Durlauf \(1996\)](#) and [Solon \(2004\)](#).

The general specification underlying standard intergenerational mobility analyses has rigorous microfoundations only because of the assumption that childhood and adulthood are undifferentiated objects in the sense that neither is envisioned as comprising many time periods. In other words, the microfoundations of (1) are predicated on the time scale of the two-period overlapping generations model. Put differently, the model (1) assumes the actual process by which children are exposed to parental income as they grow up. As such, the conventional approach does not account for the trajectories of parental incomes that are experienced by a child at different ages. In contrast, this trajectories approach is the basis of the modern skills literature ([Cunha and Heckman \(2007\)](#)). This paper takes this trajectories perspective as the foundation for measuring mobility.

To this end, we consider a trajectory of parental income instead of parents' permanent income and study how parental income trajectory maps onto children's permanent income.

Our approach is based on a functional regression model.

$$y_i = \alpha + \int_p^q \beta(r) f_i(r) dr + \varepsilon_i \quad (2)$$

where (y_i) is an average log income of child i , (f_i) is a parental log income trajectory. This model enables us to study mapping of a trajectory of parental income from child's age p to q onto children's permanent income, and as a byproduct, we obtain $\beta(r)$ which measures the association between a trajectory regressor and a scalar regressand. As we discuss in detail in the next section, approximating the parental income trajectory using parental incomes at either too few or too many ages of offspring will result in serious bias or multicollinearity, depending on whether one uses a few specific age-specific income levels or incomes at all ages. This is the main motivation for developing a novel, yet easily implementable, functional approach incorporating the entire parental income trajectory.

Our trajectories approach also allows one to expand the conception of family exposures to include variables beyond income, variables which provide a more comprehensive characterization of how family influences map into adult outcomes. A second empirical contribution of the paper is to explore how adult incomes are affected by family structure. As emphasized in the research program on early childhood investment, family structure, among other things, proxies for time constraints parents face. Further, family structure, of course, can change over the life course. Our functional approach provides intuition about the relationship between the timing of living in a single-parent family and the children's economic outcomes and whether parental divorce deteriorates children's human capital development.

Finally, we move beyond looking at how family income affects children to explore the effects of family income on some of the mechanisms that determine incomes: college attendance and occupation. This is important since the income to income relationship masks the reasons why income affects children. Our approach uncovers critical periods ([Cunha and Heckman \(2007\)](#)) for forming these outcomes and how these outcomes are coupled with intergenerational income mobility.

Our empirical focus is intergenerational mobility in Korea. After the Korean War (1950-1953), Korea was one of the poorest countries in the world. In 1960, GDP per capita was \$60. Since 1960, the Korean economy grew at a rapid pace, and in 2007, GDP per capita reached \$20,000 ([Chun \(2010\)](#)). In 2020, Korea ranked as the 10th largest country by nominal GDP. Korea therefore presents an especially interesting case because of its dramatic growth since 1975 and its associated democratization.

Our empirical results demonstrate that the predictive value of parental income on their children's success varies across by child's age. Hence, parental income effects cannot be

summarized by a scalar permanent income measure. For income, parental income in the mid-teens and 20s have stronger effects than other ages, and so these appear to be critical periods. We find that this pattern is linked to the predictive power of parental income trajectories on college attendance and occupation, which suggests the mechanisms by which parental income matters.

Our work is close in spirit to several others [Heckman and Mosso \(2014\)](#) documents the importance of timing of income and the role of credit constraint for human capital formation and provides a wide-ranging discussion of links between skills formation and mobility. [Carneiro et al. \(2021\)](#) conduct a detailed analysis of mobility in Norway, dividing parental income into early, middle, and late childhood and use a semiparametric specification of the effects of these income measures to uncover evidence on the importance of timing. They find that early and late childhood have stronger effects than income in the middle years. The paper finds a similar relationship between parental income and years of schooling.

Two other papers have also used functional data methods to study intergenerational mobility: [Chaussé et al. \(2021\)](#) and [Durlauf et al. \(2021\)](#). [Chaussé et al. \(2021\)](#) focus on integrating functional age and functional income dependence into mobility equation. [Durlauf et al. \(2021\)](#) use spline methods to evaluate trajectories and explore trajectory heterogeneities by parental characteristics such as education and structure. Our paper uses functional principal components methods to estimate trajectories, one well suited to address collinearity in family income at different dates. We appear to be the first to consider how family income trajectories affect education and occupation as well as to provide an integrated approach to the interplay of family structure and mobility.

The paper is organized as follows. Section 2 describes our functional regression approach and compares it with conventional econometric methods used in the intergenerational mobility literature. Section 3 presents the basic model of intergenerational mobility and provides empirical analysis, while Section 4 studies educational mobility and occupational mobility as mechanisms underlying intergenerational mobility. Section 5 concludes, and Appendix provides details of Korean mobility data, econometric methodologies used for our empirical analyses, and various robustness checks.

2 Econometric Methodology

2.1 Functional Regression

Throughout we use the functional regression model given by

$$y_i = \alpha + \int_p^q \beta(r)x_i(r)dr + \varepsilon_i \quad (2)$$

for $i = 1, \dots, n$, where (y_i) is the regressand, (x_i) is a functional regressor with functional regression coefficient β , and (ε_i) is the regression error. Various trajectories such as parental income profile, family structure profile and father's job profile are used as covariates in the regressions we use in our empirical investigations. The argument r in the functional covariate (x_i) signifies child age, and p and q are the beginning and end of period of parental influence, respectively. In our functional regression, the effect on (y_i) of parental income, family structure or father's job defined as (x_i) over an interval $[r - \delta, r + \delta]$ around child age r is approximately given by $\int_{r-\delta}^{r+\delta} \beta(s)ds$ for small $\delta > 0$, and $\beta(r)$ may be interpreted as the instantaneous effect on (y_i) of (x_i) at child age r . We assume that (ε_i) satisfies the usual conditions required for regression errors. Following the standard approach in functional data analysis, we regard (x_i) as functions defined in all $r \in [p, q]$, although they are not observed continuously in r . In fact, we estimate them using annual data in our empirical studies.

To relate our results from the functional regression in (2) more directly with those from the conventional regression model in the existing literature of intergenerational mobility, we let

$$[p, q) = \bigcup_{j=1}^m [p_j, q_j)$$

where $([p_j, q_j))$ is a partition of $[p, q)$ for $j = 1, \dots, m$. It follows that

$$y_i \approx \alpha + \sum_{j=1}^m \int_{p_j}^{q_j} \beta(r) \left(\frac{1}{q_j - p_j} \int_{p_j}^{q_j} x_i(r)dr \right) dr + \varepsilon_i,$$

and therefore, if we set

$$\beta_j = \int_{p_j}^{q_j} \beta(r)dr \quad \text{and} \quad x_{ij} = \frac{1}{q_j - p_j} \int_{p_j}^{q_j} x_i(r)dr \quad (3)$$

for $i = 1, \dots, n$ and $j = 1, \dots, m$, the functional regression reduces to the standard regression

model

$$y_i \approx \alpha + \sum_{j=1}^m \beta_j x_{ij} + \varepsilon_i \quad (4)$$

approximately, for $i = 1, \dots, n$. This is indeed the conventional regression used to study the intergenerational mobility.

If m is set to be too small, the approximation of $x_i(\cdot)$ by $(1/(q_j - p_j)) \int_{p_j}^{q_j} x_i(r) dr$ on the sub-interval $[p_j, q_j]$ for $j = 1, \dots, m$ becomes poor and the regression (4) does not yield a precise estimate for the functional regression (2). If we set m large to make $\max_{1 \leq j \leq m} (q_j - p_j)$ sufficiently small so that the partition of the interval $[p, q]$ given by $([p_j, q_j])$ becomes fine enough, then the regression model (4) approximates the functional regression (2) well. The number of regressors, however, increases as m increases, which will necessarily cause a serious multicollinearity problem in the regression (4). The problem of multicollinearity is particularly critical in the type of regressions we use to investigate the intergenerational mobility, since the parents' incomes at different ages of their children are strongly correlated. In sum, the regression coefficient estimates will be seriously biased if m is small, whereas they will be highly unstable if m is large. This is an unavoidable dilemma of the traditional approach based on the regression model (4). This dilemma will be analyzed and demonstrated in more detail in the next section. For this reason, we will use our methodology that has been developed specifically for the functional regression (2).

For our subsequent analysis, we assume that the functions $\beta(\cdot)$ and $x_i(\cdot)$, $i = 1, \dots, n$, are all square integrable, and regard them as elements in a Hilbert space, say, H . Then we let

$$\langle \beta, x_i \rangle = \int_p^q \beta(r) x_i(r) dr,$$

where $\langle \cdot, \cdot \rangle$ is formally defined as the inner product in H . By the Riesz representation theorem, any linear functional from H can be represented as a functional given by an inner product. Therefore, for any linear functional $\ell : H \rightarrow R$ given, we can find β such that

$$\ell(x) = \langle \beta, x \rangle$$

for all $x \in H$, which implies that the regression function in (2) accommodates any linear effect of the functional covariate (x_i) on (y_i) .

Finally, we note that it is possible to generalize our functional regression to allow for interactions between incomes at different ages. For the data under study, we did not find evidence that such nonlinearities matter. Please see appendix for details.

2.2 Estimation

In this section, we explain how to estimate our functional regression (2), which requires the projection of (y_i) onto the functional observations (x_i) . This is possible only when we replace the infinite-dimensional (x_i) with their finite-dimensional approximations. Our approach is to convert our functional regression to a standard regression and estimate it in the standard regression framework. This is not only necessary to contrast our functional regression more directly with the conventional regression, but also useful to estimate our functional regression more conveniently and efficiently. In what follows, we let (v_j) be an orthonormal basis of the Hilbert space H , and represent any $w \in H$ as

$$w = \sum_{j=1}^{\infty} \langle v_j, w \rangle v_j,$$

which is approximated as

$$w \approx \sum_{j=1}^m \langle v_j, w \rangle v_j$$

for an appropriately chosen m . The standard approach used in the statistical functional data analysis, the details of which we refer to [Ramsay and Silverman \(2005\)](#) and [Ramsay and Silverman \(2007\)](#), does not pay much attention to the choices of the basis (v_j) and the truncation number m . They are, however, crucial for the performance of the resulting estimator, as will be made clear in our subsequent discussions.

Our estimation procedure relies on the framework and the theory developed recently by [Chang et al. \(2021\)](#). We consider a mapping π on H defined as

$$\pi : w \rightarrow \begin{pmatrix} \langle v_1, w \rangle \\ \vdots \\ \langle v_m, w \rangle \end{pmatrix}$$

for any $w \in H$. If we let H_m be the subspace of H spanned by the sub-basis $(v_j)_{j=1}^m$, the mapping π defines an isometry between H_m and R^m . For any $w \in H_m$, there exists one and only one m -dimensional vector $\pi(w)$. Moreover, for any $w = \sum_{j=1}^m c_j v_j$ given as a linear combination of $(v_j)_{j=1}^m$, we have $\pi(w) = (c_1, \dots, c_m)'$, and therefore,

$$\|w\|^2 = \sum_{j=1}^m c_j^2 = \|\pi(w)\|^2,$$

where we use the notation $\|\cdot\|$ to denote the Hilbert space norm in H for w , $\|w\|^2 = \langle w, w \rangle$, on

the left-hand side and the usual Euclidean norm in R^m for $\pi(w)$, $\|\pi(w)\|^2 = \|(c_1, \dots, c_m)'\|^2$, on the right-hand side.

For any $w \in H$, we may write $w = \Pi_m w + (1 - \Pi_m)w$, where Π_m is the projection on H_m and $1 - \Pi_m$ is the projection on the orthogonal complement of H_m , i.e., the subspace of H spanned by $(v_j)_{j=m+1}^\infty$. In our approach, we approximate w by $\Pi_m w$, and analyze $\Pi_m w$ by converting it to an m -dimensional vector using the isometry π . Throughout the paper, we write $\pi(w) = [w]$ for notational brevity. Note that we may readily recover $\Pi_m w$ from $[w]$ by

$$\Pi_m w = \sum_{j=1}^m \langle v_j, w \rangle v_j = \sum_{j=1}^m [w]_j v_j,$$

where $[w] = ((w)_1, \dots, (w)_m)'$. Furthermore, for our subsequent analysis, we redefine y_i and x_i as $y_i - (1/n) \sum_{i=1}^n y_i$ and $x_i - (1/n) \sum_{i=1}^n x_i$, respectively, for $i = 1, \dots, n$ and assume that (x_i) is already demeaned, so that we may let $\alpha = 0$ without loss of generality. This is to focus on the estimation of the functional regression coefficient $\beta(\cdot)$.

Using the convention and notation introduced above, we may deduce from (2) that

$$y_i = \langle \beta, x_i \rangle + \varepsilon_i = \langle \beta, \Pi_m x_i \rangle + \langle \beta, (1 - \Pi_m) x_i \rangle + \varepsilon_i, \quad (5)$$

where

$$\langle \beta, \Pi_m x_i \rangle = [\beta]'(x_i)$$

is the main term. The approximation error term $\langle \beta, (1 - \Pi_m) x_i \rangle$ disappears in the limit as $m \rightarrow \infty$, since

$$\sum_{i=1}^n \|(1 - \Pi_m) x_i\|^2 \rightarrow_p 0$$

for an appropriately chosen basis (v_j) . We estimate (β) by its least squares estimator $(\widehat{\beta})$ in the regression

$$y_i \approx [\beta]'(x_i) + \varepsilon_i, \quad (6)$$

from which we may easily obtain an estimate for $\beta(\cdot)$ as explained above. Therefore, our estimation procedure is extremely simple, once we fix an orthonormal basis (v_j) . We use the same approach also for the estimation of discrete choice models with functional covariates.

For our empirical analysis, we use the functional principal component basis, following Bosq (2000), Park and Qian (2012) and Chang et al. (2021), among others. Let

$$Q = \sum_{i=1}^n (x_i \otimes x_i), \quad (7)$$

where “ \otimes ” signifies the tensor product defined in the Hilbert space H which reduces to the outer product when H is finite dimensional. From the practical point of view, we may indeed identify two functions z and w as long vectors consisting of their values in the ordinate corresponding to a fine enough grid of values in the abscissa, and interpret the tensor product $z \otimes w$ of two functions z and w as their outer product zw' , analogously as we may interpret the inner product $\langle z, w \rangle$ as $z'w$ if we scale r so that $dr = 1$. We define (λ_j^*, v_j^*) to be the set of pairs of eigenvalue and eigenfunction, which are ordered so that $\lambda_1^* > \lambda_2^* > \dots$. The eigenfunctions (v_j^*) thus defined are indeed normalized functional principal components.

The functional principal component analysis (FPCA) generalizes the principal component analysis (PCA). If (x_i) is finite dimensional, Q reduces to $\sum_{i=1}^n x_i x_i'$. The principal components of (x_i) in this case are given by the eigenvectors of Q associated with its leading eigenvalues, and they represent the linear combinations of (x_i) whose Euclidean norms are maximized. For functional (x_i) , the functional principal components are defined analogously as the eigenfunctions of Q in (7) associated with its leading eigenvalues, and they represent the linear combinations of (x_i) whose Hilbert space norms are maximized.

The choice of a basis is crucially important, and using the functional principal component basis has some critical advantages. Let Π_m^* be the projection on the subspace spanned by the m -leading principal components $(v_j^*)_{j=1}^m$. Then, as shown in [Chang et al. \(2021\)](#), we have

$$\Pi_m^* = \operatorname{argmin}_{\Pi_m} \sum_{i=1}^n \| (1 - \Pi_m) x_i \|^2, \quad (8)$$

where Π_m is used as a generic notation to signify a projection on any m -dimensional subspace of H . This means that the principal component basis approximates (x_i) by $(\Pi_m^* x_i)$ most effectively by the squared error sense. The squared error becomes larger when any m elements from any other basis is used. Second, when the functional principal basis is used, we have

$$\sum_{i=1}^n (\Pi_m^* x_i \otimes (1 - \Pi_m^*) x_i) = \Pi_m^* Q (1 - \Pi_m^*) = 0. \quad (9)$$

Note that $\langle v_j, Q v_k \rangle = 0$ for all $j \neq k$. This implies that the two functional regressors in (5) are orthogonal, we may omit the second term $\langle \beta, (1 - \Pi_m^*) x_i \rangle$ when we estimate β in the first term using $(\Pi_m^* x_i)$. The regression (6) is then essentially defined without any omitted term even in finite samples. No other basis provides this pleasant property. These two advantages play an important role in the general sample and bootstrap consistency proof of our procedure as shown in [Chang et al. \(2021\)](#). For choices of other bases, no such general consistency has been established to the best of our knowledge.

2.3 Functional vs. Conventional Regression

To facilitate the comparison with our functional regression, we reinterpret the conventional regression introduced in (4) as a functional regression with a particular basis. For any m given, we define an orthonormal basis

$$v_j = \frac{1}{\sqrt{q_j - p_j}} \mathbf{1}\{p_j \leq r < q_j\}, \quad (10)$$

$j = 1, \dots, m$, where $\mathbf{1}\{\cdot\}$ denotes the indicator function. Then we have $\sqrt{q_j - p_j}x_{ij} = \langle v_j, x_i \rangle$ for $i = 1, \dots, n$ and $j = 1, \dots, m$ and $(1/\sqrt{q_j - p_j})\beta_j = \langle v_j, \beta \rangle$ for $j = 1, \dots, m$, from which it follows that

$$\sum_{j=1}^m \beta_j x_{ij} = \sum_{j=1}^m \langle v_j, \beta \rangle \langle v_j, x_i \rangle = \left\langle \beta, \sum_{j=1}^m \langle v_j, x_i \rangle v_j \right\rangle = \langle \beta, \Pi_m x_i \rangle,$$

where Π_m is the orthogonal projection on the subspace spanned by $(v_j)_{j=1}^m$ defined in (10). Consequently, the conventional regression (4) may be viewed as the functional regression (5), which is estimated by (6). Throughout this section, we let $(v_j)_{j=1}^m$ be the basis defined in (10) and Π_m be the orthogonal projection on the span of $(v_j)_{j=1}^m$ without any further reference.

The conventional regression is thus a functional regression specified and estimated implicitly by the basis given by (10), which compares with our functional regression using the principal component basis. Unfortunately, the functional regression relying on the basis (10) does not have the optimality properties of our functional regression in (8) and (9). For the basis used in the conventional regression, we have

$$\sum_{i=1}^n \|(1 - \Pi_m)x_i\|^2 > \sum_{i=1}^n \|(1 - \Pi_m^*)x_i\|^2$$

and

$$\sum_{i=1}^n (\Pi_m x_i \otimes (1 - \Pi_m)x_i) \neq 0.$$

Therefore, the conventional regression has a larger approximation error than our functional regression, and has an omitted variable problem which is non-existing in our functional regression. This will be explained in more detail below.

There are three important reasons why we believe the use of our functional regression is much more effective in investigating intergenerational mobility. First, our functional regression estimates the infinite-dimensional functional coefficient β by an estimate of its finite-dimensional approximation $\Pi_m^*\beta$. From the practical point of view, this is the most

effective approximation. Constructing from the principal components, $\Pi_m^*x_i$ provides the m -dimensional components of $x_i(\cdot)$ having the largest variations over $i = 1, \dots, n$. Clearly, for any $v \in H$, $\langle v, \beta \rangle$ is identified only when $\langle v, x_i \rangle$ has enough variation over $i = 1, \dots, n$. Second, the regressors in the conventional regression are not orthogonal to the approximation errors, and therefore, the estimator of $\Pi_m\beta$ is biased. In contrast, the regressors in our functional regression are orthogonal to the approximation errors, and our estimator of $\Pi_m^*\beta$ is unbiased. Third, in all cases we consider in this paper and our other related work, the integrated variance of $\hat{\beta}$ is always smaller than that of $\tilde{\beta}$. This will be discussed further below.

For the rest of this section, we will use our main model for intergenerational mobility, which defines the children's income as (y_i) and parents' income profiles as (x_i) , and empirically evaluate and compare our novel functional regression (2) and the conventional regression (4). We consider various choices of m , which denotes the number of functional principal components used to approximate parents' income profiles for our functional regression, and also the number of age sub-intervals made in the partition of the total range of ages in the parents' income profiles for the conventional regression. Note that m denotes the number of regressors in both of our functional regression and the conventional regression, although they are set to be different in our subsequent discussions. In our data set, we have 20 years of parents' income profiles, and therefore, we consider the conventional regression with $m = 1, 2, 4, 5, 10$ and 20, which correspond to the partitions of 20 years interval by sub-intervals of lengths 20, 10, 5, 4, 2 and 1, respectively. For our functional regression, we consider $m = 1, \dots, 20$.

We now compare the actual estimates of $\beta(\cdot)$ obtained from our functional regression and the conventional regression, which are denoted as $\hat{\beta}(\cdot)$ and $\tilde{\beta}(\cdot)$, respectively. If we let $(\tilde{\beta}_j)$ be an estimate of (β_j) for $j = 1, \dots, m$ in the conventional regression (4), then $\tilde{\beta}(\cdot)$ is defined as

$$\tilde{\beta}(r) = \sum_{j=1}^m \frac{\tilde{\beta}_j}{q_j - p_j} 1\{p_j \leq r < q_j\}, \quad (11)$$

where $1\{\cdot\}$ signifies the indicator function as before. To compare $\tilde{\beta}(\cdot)$ with $\hat{\beta}(\cdot)$, we compute

$$\kappa = \frac{\|\tilde{\beta} - \hat{\beta}\|}{\|\hat{\beta}\|}, \quad (12)$$

where $\|\cdot\|$ denotes the Hilbert space norm given by $\|w\|^2 = \int w(r)^2 dr$. Note that κ measures how distinct $\tilde{\beta}$ is from $\hat{\beta}$. The computed values of κ is given in Table 1.

Table 1 shows that the two estimates $\hat{\beta}(\cdot)$ and $\tilde{\beta}(\cdot)$ for the response of the children's

Table 1: Distinctiveness of $\tilde{\beta}(\cdot)$ from $\hat{\beta}(\cdot)$

m	1	2	4	5	10	20
κ	0.68	0.92	5.23	2.92	6.38	11.15

Notes: Presented are the distinctiveness measures defined in (12), where $\tilde{\beta}(\cdot)$ is obtained as in (11) from the conventional regression with various choices of the number m of regressors.

income to the parents' income trajectory obtained from our functional regression and the conventional regression are quite distinctive of each other. We obtain $\tilde{\beta}(\cdot)$ with various choices of m , while we set $m = 2$ and use two leading principal components for $\hat{\beta}(\cdot)$ as in our empirical analysis reported in the next section. The actual estimates of the functional coefficients $\tilde{\beta}(\cdot)$ and $\hat{\beta}(\cdot)$ are presented Appendix B. Interestingly, $\tilde{\beta}(\cdot)$ becomes more distinctive of $\hat{\beta}(\cdot)$ as m increases. For $m = 1$, the magnitude of the difference between the two estimates $\tilde{\beta}(\cdot)$ and $\hat{\beta}(\cdot)$ is already more than half of the magnitude of $\hat{\beta}(\cdot)$. The potential bias problem in the conventional regression seems non-negligible. As m increases, the bias is expected to decrease. However, the relative distance between $\tilde{\beta}(\cdot)$ and $\hat{\beta}(\cdot)$ further increases and it gets quickly larger. We believe that this is because as m increases the problem of multicollinearity becomes severe, although the approximation error shrinks. In fact, $\tilde{\beta}(\cdot)$ looks totally nonsensical for $m = 10$ and 20. It appears that the problem of multicollinearity gets worse faster than the rate of approximation error disappearing as m increases.

Let \mathbb{E} denote the conditional expectation given (x_i) . Also, we let $\bar{\beta} = \Pi_m \beta$ and $\bar{\beta}^* = \Pi_m^* \beta$. To compare $\hat{\beta}(\cdot)$ and $\tilde{\beta}(\cdot)$ more formally, we write

$$\begin{aligned}\hat{\beta} - \bar{\beta}^* &= (\hat{\beta} - \mathbb{E}\hat{\beta}) \\ \tilde{\beta} - \bar{\beta} &= (\tilde{\beta} - \mathbb{E}\tilde{\beta}) + (\mathbb{E}\tilde{\beta} - \bar{\beta})\end{aligned}$$

Note that $\tilde{\beta}(\cdot)$ has an additional bias term, whereas $\hat{\beta}(\cdot)$ does not. We consider the integrated mean squared error (IMSE) defined by

$$\begin{aligned}\mathbb{E}\|\hat{\beta} - \bar{\beta}^*\|^2 &= \mathbb{E} \int_p^q (\hat{\beta}(r) - \bar{\beta}^*(r))^2 dr \\ \mathbb{E}\|\tilde{\beta} - \bar{\beta}\|^2 &= \mathbb{E} \int_p^q (\tilde{\beta}(r) - \bar{\beta}(r))^2 dr.\end{aligned}$$

For the estimator $\hat{\beta}$ based on the functional principal component basis (v_j^*) , we may

Table 2: Comparisons of $\hat{\beta}$ and $\tilde{\beta}$

	m	1	2	4	5	10	20
$\hat{\beta}$	FR2	0.67	0.76	0.85	0.88	0.95	1
	IVAR ($\times 10^{-5}$)	0.38	5.41	113.45	98.17	709.27	5879.81
$\tilde{\beta}$	FR2	0.54	0.63	0.71	0.75	0.85	1
	IVAR ($\times 10^{-5}$)	0.39	6.20	148.02	111.61	859.01	5879.81
	IBS ($\times 10^{-5}$)	0.13	3.14	0.92	0.41	0.17	0

Notes: This table presents the FR2 and IVAR of $\hat{\beta}$ based on our functional regression, and the FR2, IVR and IBS of $\tilde{\beta}$ obtained from the conventional regression, both with various choices of the number m of regressors.

readily deduce that $\mathbb{E}\hat{\beta} = \bar{\beta}^*$ and

$$\mathbb{E}\|\hat{\beta} - \bar{\beta}^*\|^2 = \mathbb{E}\|\hat{\beta} - \mathbb{E}\hat{\beta}\|^2 = \sigma^2 \text{trace} \left(\sum_{i=1}^n (x_i)^*(x_i)^* \right)^{-1} \quad (13)$$

Similarly, for the estimator $\tilde{\beta}$, we have

$$\mathbb{E}\|\tilde{\beta} - \bar{\beta}\|^2 = \mathbb{E}\|\tilde{\beta} - \mathbb{E}\tilde{\beta}\|^2 + \|\mathbb{E}\tilde{\beta} - \bar{\beta}\|^2 = \sigma^2 \text{trace} \left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} + \|\text{bias}(\tilde{\beta})\|^2 \quad (14)$$

where

$$\text{bias}(\tilde{\beta}) = \left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} \left(\sum_{i=1}^n (x_i) \langle \beta, x_i \rangle \right) - (\beta).$$

The IMSE of $\tilde{\beta}$ is therefore defined as the sum of the integrated variance (IVAR) and the integrated bias squared (IBS) of $\hat{\beta}$. The technical details of the econometric results here are provided in Appendix C.1.

We also define

$$\rho^2 = \frac{\sum_{i=1}^n \|\Pi_m x_i\|^2}{\sum_{i=1}^n \|x_i\|^2} = \frac{\text{trace}(\Pi_m Q \Pi_m)}{\text{trace } Q}, \quad (15)$$

and denote by ρ_*^2 the corresponding value obtained with Π_m^* in place of Π_m . For each m , ρ^2 and ρ_*^2 represent the proportions of the total variation of (x_i) explained by $(\Pi_m x_i)$ and $(\Pi_m^* x_i)$ in the conventional regression and our functional regression, respectively, which will be referred to as the functional R-square (FR2) of $\hat{\beta}$ and $\hat{\beta}^*$. The IVAR, IBS and FR2 computed for $\tilde{\beta}(\cdot)$ and $\hat{\beta}(\cdot)$ obtained with various choices of m are reported in Table 2.

We set $m = 2$ as selected by the cross-validation in our empirical study reported in

the next section, and with this choice of $m = 2$, the FR2 of our functional regression is 76%. This implies, in our functional regression, $m = 2$ already explains 76% of the total variation of (x_i) . The conventional regression reaches this level if we set $m = 5$. However, as shown in Table 2, $\tilde{\beta}$ in the conventional regression has IVAR which is about twenty times larger than that of $\hat{\beta}$ in our functional regression. Our functional regression is therefore by far more efficient than the conventional regression. In addition, as m increases, multicollinearity is necessarily aggravated, which causes the so-called ill-posed inverse problem that is well known to be a serious problem especially in the functional data analysis. We thus cannot get reliable estimates for the coefficients of individual regressors, even when the set of regressors as a whole explains a good portion of the total variation in the data. In fact, the condition number, which is defined as the ratio of the maximum and minimum eigenvalues of the design matrix and often used as a measure of the degree of multicollinearity, is given by 1, 6.27, 23.17, 22.13, 43.13, 82.71 for the conventional regression with $m = 1, 2, 4, 5, 10, 20$, respectively. We believe that this explains why the estimates from the conventional regression become unstable and uninterpretable as we increase m up to 10 and 20. The condition number of our functional regression is $\lambda_1^*/\lambda_2^* = 6.98$, which is about the same as that of the conventional regression with $m = 2$. Lastly, as noted, $\hat{\beta}$ is unbiased and it has no bias term. The IBS of $\tilde{\beta}$ depends on the unknown functional parameter β and we use as a proxy our functional regression estimator $\hat{\beta}$ with $m = 2$. The magnitude of the IBS of $\hat{\beta}$ appears to be substantial when m is set to be too small.

3 Empirical Analysis of Intergenerational Economic Mobility

In this section, we use our methodology to systematically investigate intergenerational mobility. In the model, we employ the regressand (y_i) as a time average of child's log income and a functional covariate (f_i) as a trajectory of parents' log income over a long period, and estimate the functional regression coefficient β using our functional estimation method described in Section 2.2.

3.1 Data

We view (f_i) , which is a trajectory of parental log income, as functional data with discrete observations. For example, consider a collection of annual income for 20 years. From our viewpoint, this collection of income data is functional data, recorded annually, with 20 observations.

To view (f_i) as functional data with discrete observations, we need to track individuals to collect data comprising several observations. Obviously, a greater number of observations is better. To this end, we employ longitudinal data and track individuals over a long period. As the number of observations depends on the length of the period and frequency, the best scenario is to track each individual at a high frequency over a long time. However, as the frequency of most longitudinal data is at most annual, to track parental income, we consider longitudinal data that span over 20 years. In most countries, particularly in Korea, the majority of parents have children aged 19 years or less. Therefore, we focus on parental income during the time children are aged between 0 and 19 years, which covers the entire period when children were minors. This allows us to study the relationship between parental income when their children were minor and their offspring's success as an adult.

The Korean Labor & Income Panel Study (KLIPS) data set satisfies the above-mentioned condition. KLIPS is a longitudinal survey of the labor market, income activities of households, and individuals residing in urban areas in Korea and is available from 1998 to 2019 at an annual frequency (in total, 22 years).

As noted earlier, our objective is to investigate the association between parental income when the children were minors and their success as an adult using regression analysis (2),

$$y_i = \alpha + \int_p^q \beta(r)f_i(r)dr + \varepsilon_i. \quad (2)$$

For empirical regression analysis (2), the regressand (y_i) and the regressor (f_i) have to be defined first. The permanent log income of children (y_i) is approximated as the average log income. We define it as the average log income in the early 30s because an adult tends to begin their career before 30, and data on income during this period is generally available. However, how we define (y_i) raises an issue. As the entire data span of KLIPS is 22 years, obtaining parental income when their children were minors (aged 0–19 years) and their children's income in the early 30s simultaneously is not possible. To this end, as the second-best, we examine the association between parental income from a child's age 10 to 29 years and the proxy of a child's permanent income. This approach allows us to explore the association between parental income when their children were in their teens and 20s and their offspring's income in the early 30s. Analyzing the association during this period can be interesting as most parents in Korea support and live with their children until they get married. Commonly, an adult in Korea lives with their parents even in their late 20s. Thus, we track parental income from a child's age 10 to 29 years and consider this collection of income as the functional data.

We focus on the birth cohorts between 1987 and 1989, track their parental income for 20

years and obtain their income in their early 30s. The total number of observations is 190¹. Specifically, we construct the data set for each birth cohort (from 1987 to 1989) as follows.

For each birth cohort (from 1987 to 1989),

1. Collect data on families who have a newborn baby (child) in that year.
2. For each child's parents, track their income for 20 years (from the child's age 10 to 29 years).
3. Track the child's income during their adulthood for three years (from the child's age 30 to 32 years).

A functional covariate (f_i), which indicates a collection of parental log income for 20 years, is obtained in step 2. In step 3, we obtain the regressand (y_i) by taking the average log income of children when they reach the early 30s. Both parental income and child income are defined following Chetty et al. (2014). Specifically, parental income implies “parents’ generation income,” which is the summation of father’s and mother’s income. Similarly, we define offspring’s income as “offspring generation income,” which is the summation of individual i ’s and their partner’s income. All monetary variables are measured in the 2015 Korean Won using the consumer price index in Korea.

We refer to previous studies to consider various control variables to evaluate the pure effect of parental log income on offspring’s log income. For example, we consider the number of offspring, average parents’ age at their child’s birth² and its square, and the education level of both parents and children.

Table 3 reports summary statistics of the control variables that describe the characteristic of parents of birth cohorts from 1987 - 1989. On average, the father is older than the mother at childbirth, and the number of offspring equals 2.15. Most parents’ education levels include high school degrees or below, and the father tends to be more educated than the mother.

3.2 Intergenerational Income Mobility

Our baseline model to analyze intergenerational income mobility is as the following.

$$y_i = \alpha + \int_{10}^{29} \beta(r) f_i(r) dr + \varepsilon_i \quad (16)$$

¹Any observations with missing value in parental income trajectory are excluded.

²As the sample comprises some single-parent families, we take the average of parents’ ages and employ it as a control variable.

Table 3: Summary Statistics

	Mean	Stdev	# Obs
<u>A. Characteristic of Family</u>			
Age of father at birth	29.77	3.67	174
Age of mother at birth	26.02	5.19	180
Number of offspring	2.15	0.67	190
<u>B. Father's Education Level</u>			
High school or below	0.70	0.46	174
2-year college	0.10	0.30	174
4-year college or above	0.20	0.40	174
<u>C. Mother's Education Level</u>			
High school or below	0.84	0.37	180
2-year college	0.05	0.22	180
4-year college or above	0.11	0.32	180

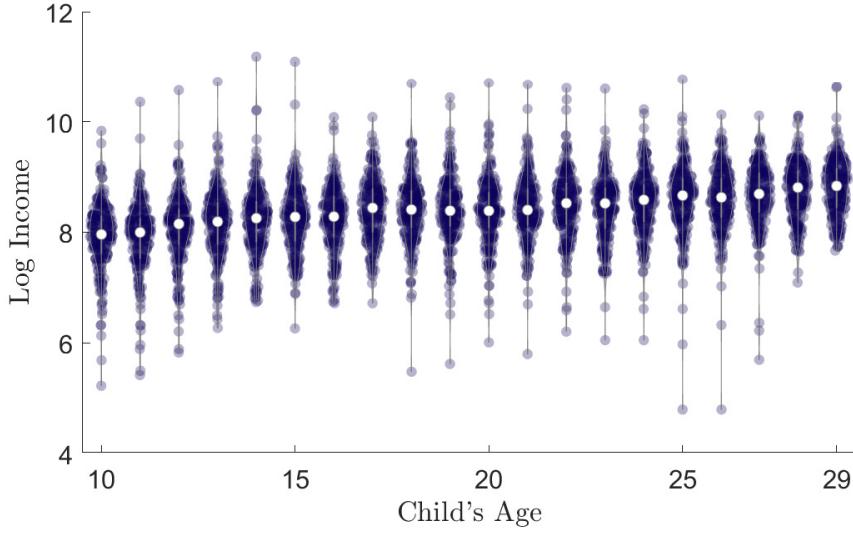
Notes: This table reports summary statistics of the control variables for birth cohorts 1987-1989. Panel A presents the variables related to the characteristics of the family. Panel B and C display the variables on parents' education levels. All variables in these panels are dummy variables. For example, 'High school or below' equals 1 if one's highest level of education is high school graduate or below and 0 otherwise. The other education dummy variables, '2-year college' and '4-year college or above', are defined similarly.

where (y_i) is the 3-year average (between 30 and 32 years old) log income of child i , (f_i) is the trajectory of i 's log parental income for 20 years (from a child i 's age 10 to 29 years), and $\beta(\cdot)$ is the intergenerational elasticity (IGE) of income function over the same period (i.e., from a child's age 10 to 29 years). To highlight its age heterogeneity, we specifically call $\beta(\cdot)$ the age-varying IGE of income. The model (16) maps the parental income trajectory onto the proxy of a child's permanent income, and $\beta(\cdot)$ in this model assesses the association between the trajectory of parental income from a child's age of 10 to 29 years and the child's permanent income.

Figure 1 presents the violin plots of parental log income of 190 parents of birth cohorts from 1987-1989 at each age of the child. The violin plot displays a shape of parental income density by illustrating the rotated kernel density plot at each age of the child and the shape of log income close to normal distribution. The white circle in the middle of the violin plots indicates the median parental income. It ranges from 7.96 (child's age 10 years) to 8.84 (child's age 29 years). As evident in Figure 1, parental log income tends to increase.

We estimate the coefficient function of interest $\beta(\cdot)$ in (16) using functional principal component analysis as introduced in Section 2.2. The dimension m is chosen as 2 by mini-

Figure 1: Violin Plots for Parental Income



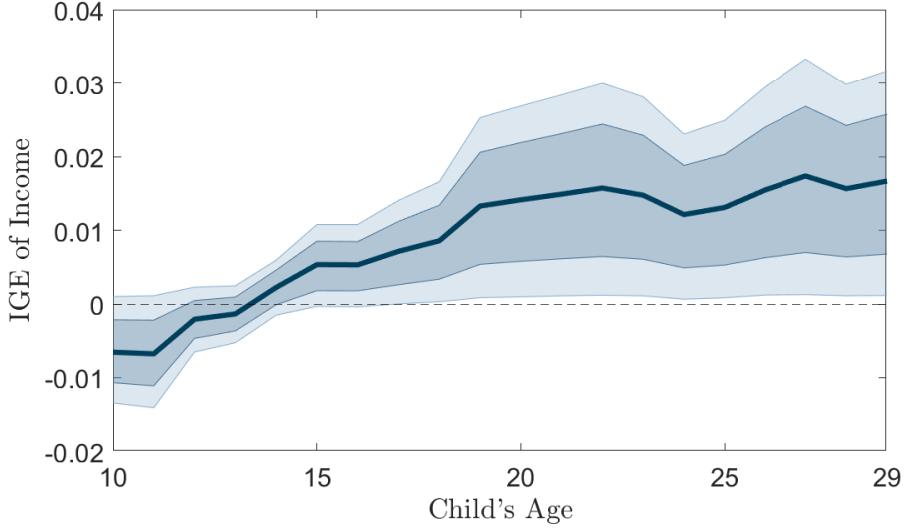
Notes: This figure presents the violin plots for parental income of the birth cohorts between 1987 and 1989. The violin plot displays a shape of parental income density by showing the rotated kernel density plot at each child's age. The white circle in the middle indicates the median parental log income.

mizing the leave-one-out cross-validation error (see Appendix D.1 for details). The estimated $\hat{\beta}(\cdot)$ in Figure 2 indicates how IGE of income evolves as a child grows up. First, the estimated age-varying IGE of income $\hat{\beta}(\cdot)$ exhibits strong evidence of heterogeneity of timing in intergenerational economic mobility, which is a supporting piece of evidence for applying a functional approach to explore intergenerational mobility. Specifically, the estimated $\hat{\beta}(\cdot)$ tends to increase monotonically and turns to positive and is statistically significant after a child's age of 15 years. These findings imply parental income after their offspring's age reaches 15 years is crucial to a child's permanent income. Thus, the period after the mid-teens is critical for intergenerational income mobility.

To compare our results with the standard empirical IGE measure of income (β in (1)), we integrate the estimated age-varying IGE of income function $\hat{\beta}(\cdot)$ over time. Specifically, we compare $\int_{10}^{29} \hat{\beta}(r) dr$ with the estimated $\hat{\beta}$ in (1). As discussed in Section 2.1, the functional regression is reduced to the standard regression model. Note that

$$y_i \approx \alpha + \sum_{j=1}^m \beta(r_j) \int_{p_j}^{q_j} x_i(r) dr + \varepsilon_i,$$

Figure 2: Age-Varying IGE of Income



Notes: This figure plots the estimated $\hat{\beta}(\cdot)$ in (16). $\hat{\beta}(\cdot)$ measures the percentage change of a child's income in response to parental log income from a child's age 10 to 29. The darker (lighter) shaded areas indicate 68% (90%) bootstrap confidence intervals.

and when $m = 1$, we have

$$y_i \approx \alpha + \beta(r_1) \int_{10}^{29} x_i(r) dr + \varepsilon_i,$$

where $r_1 \in [10, 29]$. If we set

$$\beta = 19\beta(r_1) \quad \text{and} \quad x_i = \frac{1}{19} \int_{10}^{29} x_i(r) dr$$

for $i = 1, \dots, n$, then the functional regression reduces to the standard regression model

$$y_i \approx \alpha + \beta x_i + \varepsilon_i.$$

The choice of r_1 is crucially important, and we choose r_1 that makes $\beta(r_1)$ equal to the average function value of $\beta(\cdot)$, i.e., $\beta(r_1) = \frac{1}{19} \int_{10}^{29} \beta(r) dr$. This relates the standard measure of IGE of income β , and our measure of the IGE of income trajectory, $\beta = \int_{10}^{29} \beta(r) dr$.

From the data, we obtain $\int_{10}^{29} \hat{\beta}(r) dr = 0.17$, which is similar to those in the existing studies that examine intergenerational income mobility in Korea. Using KLIPS data, Kim (2009) estimates IGE of income in Korea and demonstrates that it does not exceed 0.2.

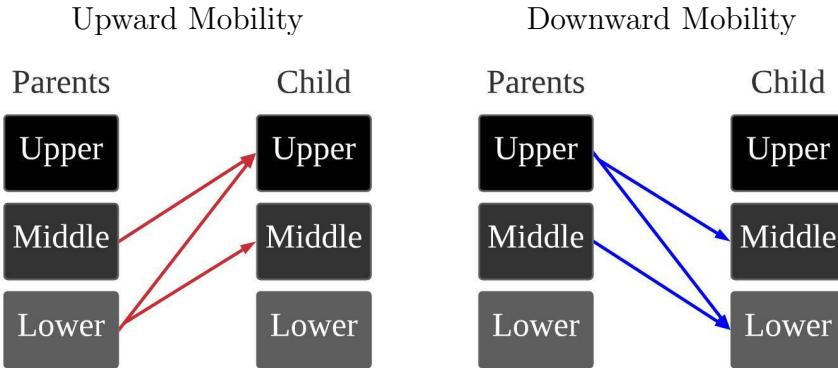
3.3 Upward and Downward Income Mobility

According to our results, parental income trajectory impacts offspring's economic success and failure. Moreover, as the effect of parental income on their children's future income varies depending on the children's age, the shape of the parental income trajectory can be more informative than the simple average of parental income when it comes to analyzing the success of the offspring. To this end, we characterize the shape of parental income that promotes or impedes offspring's economic success.

As the first step, we divide parents and children into three classes: lower-, middle-, and upper-income. Statistics Korea defines the middle-income class as individuals whose income is between 50% and 150% of median income. Following this definition, we distribute parents and children into these three income classes by average income, respectively.

We say there is upward income mobility when a child belongs to a higher-income group than their parents. Similarly, we say there is downward income mobility when the opposite is true, when a child belongs to a lower-income group than their parents. Figure 3 presents the corresponding cases of upward and downward mobility. The left-hand panel illustrates the cases of upward mobility. Upward mobility has three cases, from lower-income parents to upper- or middle-income children and from middle-income parents to upper-income children. By contrast, the right-hand panel shows the cases of downward mobility. The transitions from upper-income parents to middle- or lower-income children and from middle-income parents to lower-income children fall under this mobility.

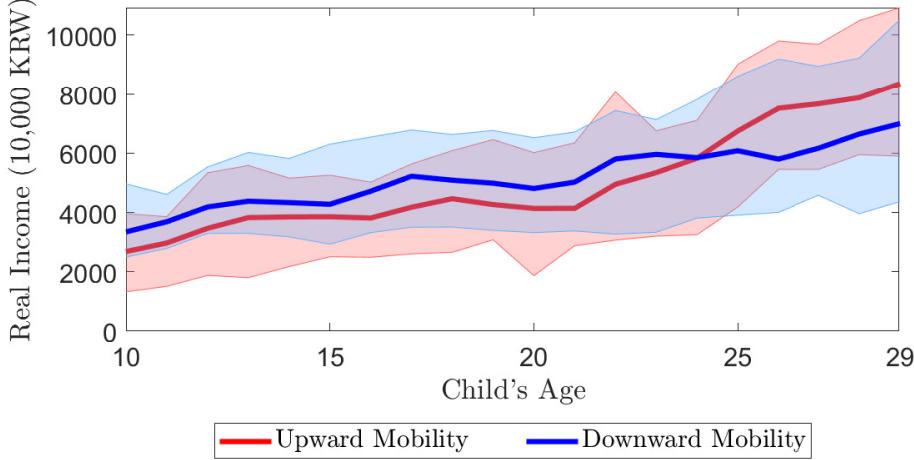
Figure 3: Upward and Downward Mobility



Notes: This figure presents how upward and downward mobility are defined. The left panel and right panel display the corresponding cases of upward and downward mobility, respectively.

For our analysis, we focus on middle-income parents and compare upward (from middle-income parents to upper-income children) and downward (from middle-income parents to lower-income children) mobility as both upward and downward mobility are possible only in

Figure 4: Parental Income Trajectories of Upward and Downward Mobility



Notes: This figure presents the parental income trajectories of upward and downward mobility. p is set at 50. The red line plots median income of set S_{50}^U , and the shaded area indicates a central 68% range (from 16th to 84th percentile) of S_{50}^U . Similarly, the blue line displays the median income of set S_{50}^D , and the shaded area is for the central 68%.

the middle-income class. To characterize parental income trajectory for each mobility case, we define two sets of parental income trajectory, S_p^U and S_p^D , for upward and downward mobility, respectively.

$$S_p^U (S_p^D) = \{ \text{Parental income trajectory such that family (parents and child) exhibits upward(downward) income mobility with } p\% \text{ probability} \}$$

To find these two sets, first, we define the indicator variable

$$z_i^U (z_i^D) = \begin{cases} 1 & \text{if family } i \text{ exhibits upward (downward) mobility} \\ 0 & \text{otherwise} \end{cases}$$

and fit the weighted functional logit model using parental income trajectory as a regressor. Once we fit the model, the predicted probability can be derived for each family i . If this predicted probability is greater than p , then the set includes the parental income trajectory of the child i . We set $p = 50$ for our analysis.

Figure 4 presents two sets of parental income trajectories, upward and downward mobility. The red line and the shaded area indicate the median and central 68% range (from 16th to 84th percentile) of upward mobility parental income trajectory set, S_{50}^U . Similarly, the blue line and the shaded area are for the median and central 68% range of S_{50}^D . We find an

interesting pattern here. The average income level of the upward mobility parental income profile is lower than that of the downward mobility parental income profile. The median (red line in Figure 4) average income of S_{50}^U equals 49,950,000 KRW while that (blue line in Figure 4) of S_{50}^D is 51,650,000 KRW. Although the upward mobility group has a lower average income, it has a more remarkable increasing trend in parental income trajectory, and after a child reaches the age of 25 years, this group tends to enjoy higher income than the downward mobility group. Our results illustrate an association between the increasing trend of parental income trajectory and upward mobility, which cannot be explained if we only focus on the average income level. We obtain similar results with different values of p (see Appendix E.1 for more details).

3.4 Family Structure Trajectory and Economic Mobility

Children are exposed to their parent's income trajectory throughout their life, although several other trajectories also affect the children. One example is the family structure. Many studies have examined the role of family structure on socioeconomic transmission from parents to children. For instance, [Duncan and Rodgers \(1991\)](#), and [Musick and Mare \(2004\)](#) investigate the link between poverty and family structure. [Bloome \(2017\)](#) explores the interaction between family structure and the persistence of intergenerational income mobility. As demonstrated in previous studies, if the family structure influences intergenerational mobility, then the timing of the family structure also impacts mobility. For example, the timing of parents being divorced or separated can be as important as whether they are divorced.

We consider two types of family structure, single-parent families and two-parent families. The transition from a two-parent family to a single-parent family is commonly caused by divorce or separation. Figure 5 presents the ratio of single-parent families in the sample. The ratio of single-parent families, the navy-blue line in Figure 5, ranges from 6.84% (at child's age 10 years) to 19.47% (at child's age 23 and 24 years), and its average, presented by the red dashed line marked at 14.66%.

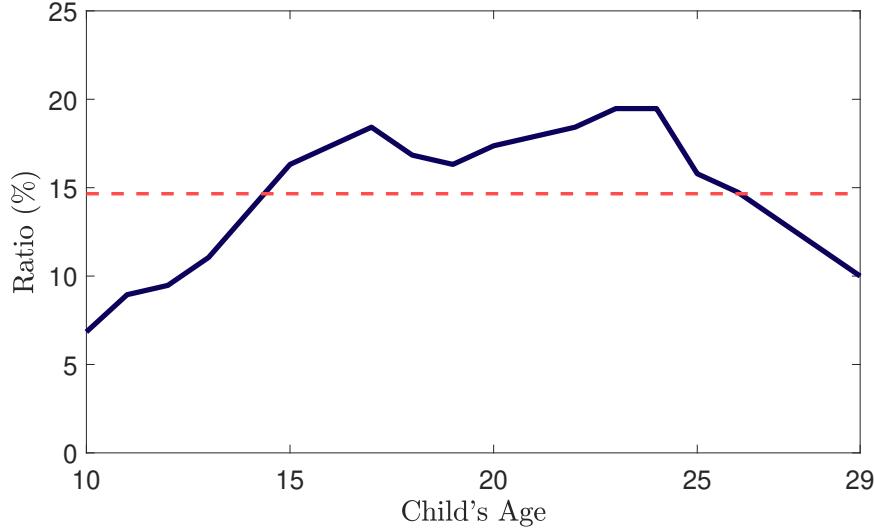
We extend our baseline model (16) to the multiple functional regression model by adding a functional regressor (g_i), family structure trajectory to the model³ and consider the following

$$y_i = \alpha + \int_{10}^{29} \beta(r)f_i(r)dr + \int_{10}^{29} \gamma(r)g_i(r)dr + \varepsilon_i \quad (17)$$

where (y_i) is the 3-year average log income of child i , (f_i) is the trajectory of i 's log parental

³Our regression model (17) can be easily extended to the one with an interaction term. See Appendix E.3 for details.

Figure 5: Ratio of Single-parent Families



Notes: The navy line plots the ratio of single-parent families at each child's age for birth cohort 1987-1989. The red-dashed line indicates the average ratio of single-parent families which is 14.66%.

income for 20 years, and (g_i) is the trajectory of family structure for 20 years. In this model, we employ additional function (g_i) as a functional regressor, while $\gamma(\cdot)$ reflects the association between the trajectory of family structure and children's income after controlling for income effect. We define (g_i) as

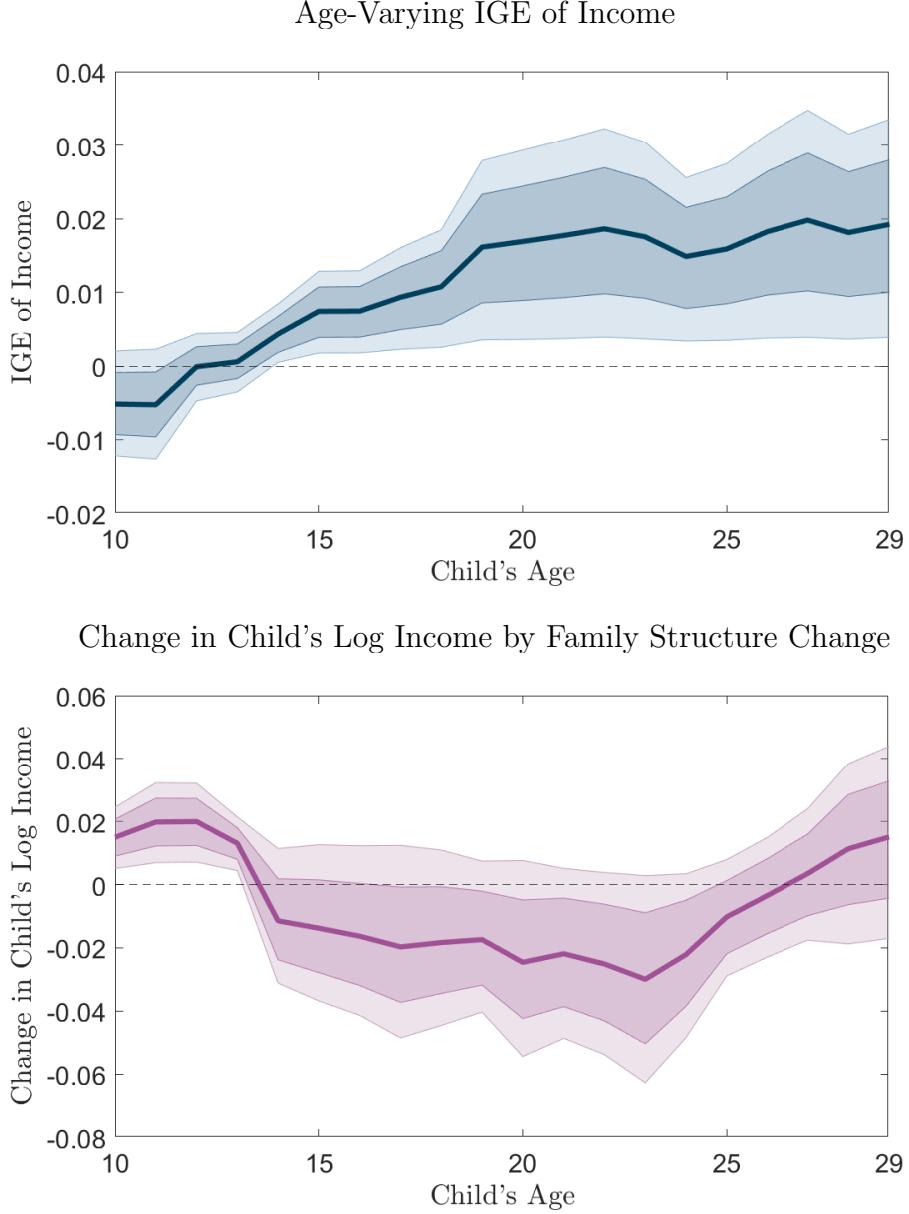
$$g_i(r) = \begin{cases} 1 & \text{if child } i \text{ lives in single-parent family at child's age } r \\ 0 & \text{if child } i \text{ lives in two-parent family at child's age } r \end{cases}$$

The interpretation of $\gamma(\cdot)$ is similar to that derived from the linear regression with a dummy variable. For example, $\gamma(r) < 0$ implies that a child in single-parent families at child's age r has a lower permanent income than that in two-parent families at the same age by $(100 \times \gamma(r))\%$.

We estimate $\hat{\beta}(\cdot)$ and $\hat{\gamma}(\cdot)$ in regression (17) using functional principal component analysis⁴ as introduced in Section 2.2. Figure 6 presents the estimated $\hat{\beta}(\cdot)$ and $\hat{\gamma}(\cdot)$ in regression (17). Note that Figure 2 (estimate in regression (16)) and the upper panel of Figure 6 (estimate in regression (17)) do not differ significantly. The reason for this is that the two regressors, (f_i) and (g_i) , are less correlated. We represent each functional regressor, (f_i) and (g_i) , as the linear combination of their respective principal components, and calculate the canonical correlation, which provides the maximum possible correlation by finding linear

⁴The number of functional principal component m is chosen as 2 and 3, by minimizing the leave-one-out cross-validation error.

Figure 6: Parental Income and Family Structure Trajectories and Offspring's Income



Notes: The panels in this figure present $\hat{\beta}(\cdot)$ and $\hat{\gamma}(\cdot)$ in (17), respectively. The upper panel presents the estimated $\hat{\beta}(\cdot)$, which measures age-varying IGE of income from a child's age 10 to 29. The lower panel plots $\hat{\gamma}(\cdot)$. This estimate measures the association between a trajectory of family structure (from a child's age 10 to 29) and a child's log income. The darker (lighter) shaded areas indicate 68% (90%) bootstrap confidence intervals.

combination of these functions. The maximum canonical correlation is 0.34, which is not high. Therefore, the estimate in simple functional linear regression does not differ significantly from that in multiple functional linear regression.

The lower panel illustrates the estimated $\hat{\gamma}(\cdot)$ in regression (17). The estimated $\hat{\gamma}(\cdot)$ exhibits age heterogeneity and turns to negative at the child's age of 15 years. After the age of 25, the estimate becomes positive, but it is not significant. Overall, this indicates that living in a single-parent family harms children's economic success, and this family structure has a severe adverse effect on children in their mid-teens and early 20s. Throughout our analysis, mid-teens and the early 20s are critical periods for intergenerational mobility. We will explore the possible mechanisms for this mobility in the following section.

4 Mechanisms Underlying Intergenerational Mobility

Our results in the previous section suggest that the mid-teens and early 20s are crucial periods for a child's economic success and intergenerational economic mobility. To explore why these periods are critical, we point out tertiary education and occupation as the key factors determining one's income and examine the association between the trajectory of parental income and two intermediate outcomes, educational attainment and occupational choices, to explain the offspring's economic success.

4.1 Educational Attainment

A college degree is a standard measure of educational attainment, and it is one of the most crucial factors that determine income level. Furthermore, it is often considered as a signal for a high level of human capital. For this reason, we investigate the relationship between parents' income and a child's college attendance in this section. In particular, we study the relationship between offspring's college attendance and various trajectories of family characteristics obtained using the KLIPS data. We focus on college attendance because the drop-out rate for undergraduate college in Korea is extremely low. According to Korean Educational Development Institute⁵, the drop-out rate of 4-year colleges was 4.6% in 2019. For this reason, one's college attendance is practically the same as earning a college degree in Korea. Using this variable as a regressand, we consider ordered logit models to explore the relationship between offspring's college attendance and trajectories such as parental income, family structure, and father's occupation.

⁵K-Indicator, Statistics Korea, Dropout Rate by Educational Institution, at https://index.go.kr/smart/mbl/chart_view.do?idx_cd=5072 (visited September 9, 2021).

Table 4: Summary Statistics

	Mean	Stdev	# Obs
<u>A. Characteristic of Family</u>			
Age of father at birth	31.39	4.79	240
Age of mother at birth	29.00	4.97	251
Number of offspring	2.02	0.61	256
<u>B. Father's Education Level</u>			
High school or below	0.55	0.50	240
2-year college	0.13	0.33	240
4-year college or above	0.33	0.47	240
<u>C. Mother's Education Level</u>			
High school or below	0.71	0.45	251
2-year college	0.12	0.33	251
4-year college or above	0.17	0.37	251

Notes: This table reports summary statistics of the control variables for the birth cohorts 1997-1999. Panel A presents the variables related to the characteristics of the family. Panel B and C display the variables on parents' education levels. All variables in these panels are dummy variables. For example, 'High school or below' equals 1 if one's highest level of education is high school graduate or below and 0 otherwise. The other education dummy variables, '2-year college' and '4-year college or above', are defined similarly.

4.1.1 Data

Note that a child's college attendance results can be observed in the early 20s. Thus, to answer our primary research question – the relationship between parental income when their children were minors and children's success – this section explores the association between parental income trajectory (from a child's age 0 to 19 years) and their children's college attendance. For this analysis, we focus on birth cohorts from 1997 to 1999. We have 256 observations for the analysis.

We consider the same control variables as in the previous section, and Table 4 reports summary statistics of the control variables for birth cohorts 1997, 1998, and 1999. Similar to Table 3, on average, the father is older than the mother at childbirth and the number of offspring equals 2.02. Most parents' education levels comprise high school degrees or below, and the father tends to be more educated than the mother. A comparison of these two tables reveals the overall trend in Korea. First, the parental age at childbirth is increasing. Second, the number of offspring is on a decline. Last, on average, parental education level is on the rise. These findings are consistent with the overall observed trends for Korea as reported, for example, in Phang and Kim (2003), and Shin et al. (2019).

One unique feature of the KLIPS data set is that it provides detailed information about college attendance, such as the school that a survey respondent attended and graduated from. Using this detailed data, we categorize children into three groups: children who attend 1) the top 30 universities⁶ or medical schools⁷, 2) one of the other 4-year colleges, and 3) 2-year colleges or end up with a high school degree.

4.1.2 Parental Income Trajectory and Offspring's Educational Attainment

To analyze the association between the trajectory of parental income and their child's college attendance, we consider the following functional ordered logit model

$$z_i = \begin{cases} 1 & \text{if } z_i^* \leq \tau_1 \\ 2 & \text{if } \tau_1 < z_i^* \leq \tau_2, \quad z_i^* = \int_0^{19} \mu(r) f_i(r) dr + \varepsilon_i, \\ 3 & \text{if } \tau_2 < z_i^* \end{cases} \quad (18)$$

where

$$z_i = \begin{cases} 1 & \text{if child } i \text{ attends 2-year colleges or below} \\ 2 & \text{if child } i \text{ attends one of the other 4-year colleges} \\ 3 & \text{if child } i \text{ attends the top 30 universities or medical schools} \end{cases}$$

and (f_i) is the trajectory of parental income for 20 years. The distribution of (ε_i) is a standard logistic distribution.⁸

Figure 7 plots the shape of parental log income of 256 families at each age of the child. The white circle in the middle of the violin plots indicates the median parental log income, ranging from 7.89 (child's age 0) to 8.77 (child's age 18). Similar to Figure 1, the parental log income tends to increase.

We estimate the response function $\mu(\cdot)$ in (18) using functional principal component analysis as introduced in Section 2.2. The number of functional principal component m is chosen as 3 by minimizing the leave-one-out cross-validation error (see Appendix D.2 for details).

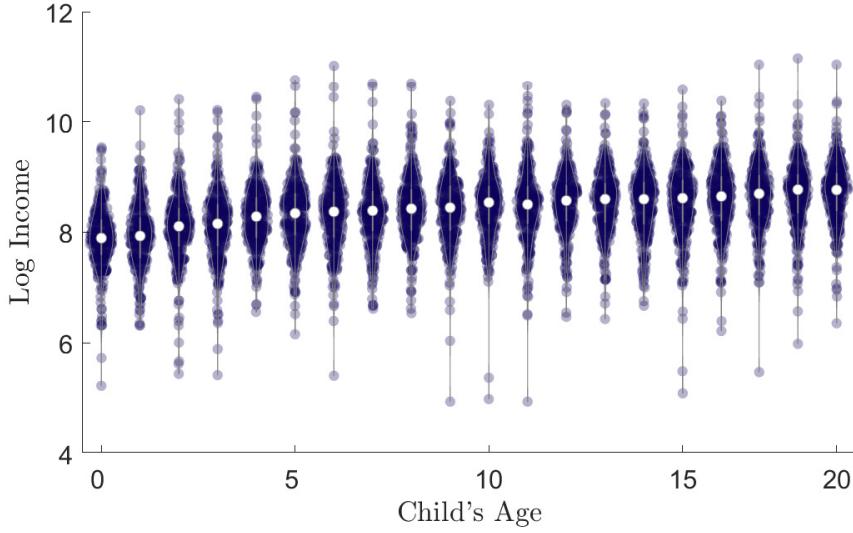
Note that as parental log income trajectory $f_i(r)$ is positive at all child's ages r , a positive $\mu(r)$ implies a positive relationship between the parental income at child's age r and their child attending a higher-level educational institution according to our classification. As illustrated in Figure 8, $\hat{\mu}(\cdot)$ is positive and statistically significant between a child's age 3 and

⁶We refer to the “Joongang Ilbo University Ranking” to define the top 30 universities group.

⁷In Korea, most medical schools offer a 6-year undergraduate program, which is different from the U.S.

⁸We also consider the functional multinomial probit model, and the estimation results are robust to specification of the distribution of the error term. See Appendix E.4 for details.

Figure 7: Violin Plots for Parental Income

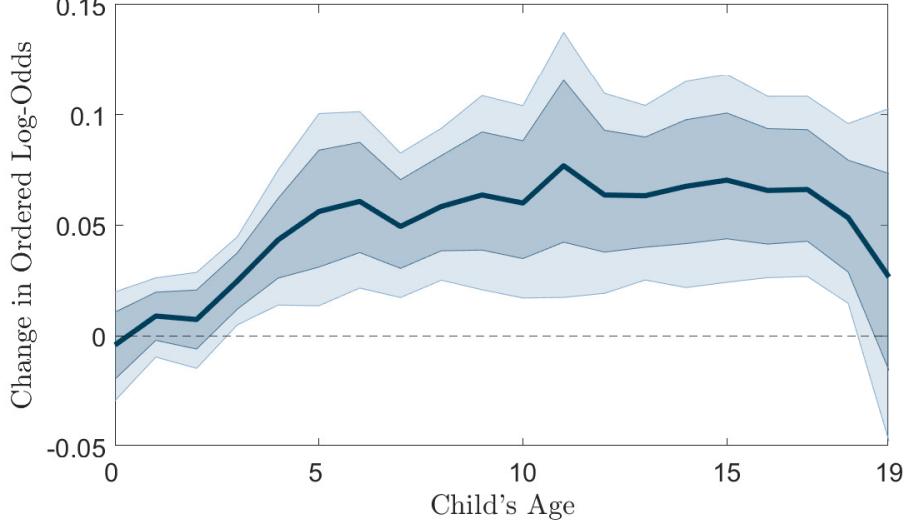


Notes: This figure presents the violin plots for parental income of the birth cohorts between 1997 and 1999. The violin plot displays a shape of parental income density by showing the rotated kernel density plot at each child's age. The white circles in the middle indicate the median parental log incomes.

18; thus, overall, parental income and attending a better institution for tertiary education are positively related. According to this result, parental income in the mid-teens appears more critical than other ages for a child's college attendance. Although we can determine the overall association between the parental income trajectory and college attendance using $\hat{\mu}(\cdot)$, quantifying the size of the effect of parental income trajectory on the college attendance is difficult from this estimate.

The exact interpretation of the estimated logit coefficient $\hat{\mu}(r)$ at age r is that when the functional regressor increases by one unit at age r , the ordered log-odds ratio also increases by the estimated level $\hat{\mu}(r)$. The interpretation is not as straightforward as the functional linear regression model, so we consider the marginal effect in place of the logit estimate. Precisely, we quantify the marginal effect of parental income at each child's age on college attendance. The marginal effect at a child's age s is derived by taking the derivative of predicted probability with respect to parental income at age s . We have the following

Figure 8: Change in Ordered Log-Odds by Change in Parental Income



Notes: This figure plots the estimate of the response of ordered log-odds to parental incomes across their child's ages from the functional ordered logit model, $\hat{\mu}(\cdot)$ in (18). This estimate measures the change of ordered log-odds ratio in response to the change in parental income trajectory. The darker (lighter) shaded areas indicate 68% (90%) bootstrap confidence intervals.

predictive probability of a child going to a level k educational institution for $k = 1, 2, 3$

$$\begin{aligned}\mathbb{P}(z_i = 1) &= F\left(\tau_1 - \int_0^{19} \mu(r)f_i(r)dr\right) \\ \mathbb{P}(z_i = 2) &= F\left(\tau_2 - \int_0^{19} \mu(r)f_i(r)dr\right) - F\left(\tau_1 - \int_0^{19} \mu(r)f_i(r)dr\right) \\ \mathbb{P}(z_i = 3) &= 1 - F\left(\tau_2 - \int_0^{19} \mu(r)f_i(r)dr\right)\end{aligned}$$

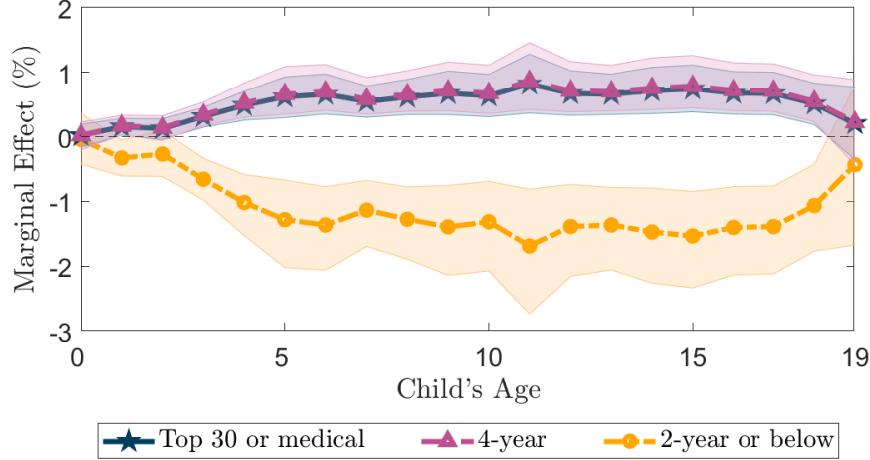
where F is the distribution function of the standard logistic distribution, that is, $F(x) = \exp(x)/(1 + \exp(x))$. We then compute the marginal effect of an increase in parental income at child's age s on the probability of their child joining to a level k institution for $k = 1, 2, 3$ as in the following three equations

$$\frac{\partial \mathbb{P}(z_i = 1)}{\partial f_i(s)} = -\mu(s)F'\left(\tau_1 - \int_0^{19} \mu(r)f_i(r)dr\right) \quad (19)$$

$$\frac{\partial \mathbb{P}(z_i = 2)}{\partial f_i(s)} = \mu(s)\left(F'\left(\tau_1 - \int_0^{19} \mu(r)f_i(r)dr\right) - F'\left(\tau_2 - \int_0^{19} \mu(r)f_i(r)dr\right)\right) \quad (20)$$

$$\frac{\partial \mathbb{P}(z_i = 3)}{\partial f_i(s)} = \mu(s)F'\left(\tau_2 - \int_0^{19} \mu(r)f_i(r)dr\right) \quad (21)$$

Figure 9: Age-Varying Marginal Effect of Parental Income on College Attendance



Notes: This figure presents the marginal effect of parental income trajectory on probability of their offspring attending to a level k institution. The navy-star line presents the marginal effect of parental income on attending top 30 universities or medical schools (21). The purple-triangle line and the yellow-circle line indicate the marginal effect of parental income on attending other 4-year colleges, (20), and attending 2-year colleges or below, (19), respectively. The shaded areas indicate 68% bootstrap confidence intervals.

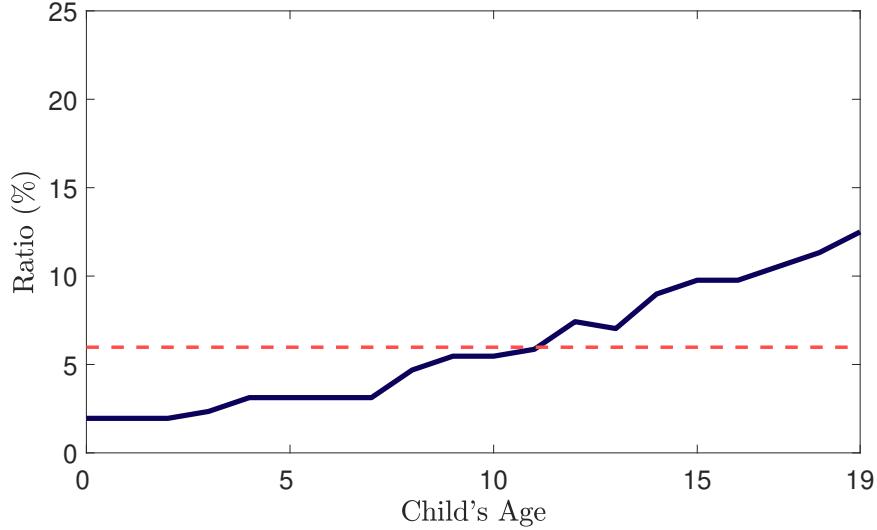
Next, we evaluate the marginal effects at the average parental income trajectory, $\bar{f} = (1/n) \sum_{i=1}^n f_i$.

Figure 9 plots the marginal effect of parental income at each age of the child on the probability of their offspring attending a level k institution. For all ages of the child, parental income tends to reduce a child's probability of attending 2-year colleges or ending up with a high school degree (see the yellow-circle line in Figure 9 plotting the marginal effect computed using equation (19)), and increases the probabilities of attending institutions 4-year colleges (see the purple-triangle line and equation (20)) and top 30 universities or medical schools (navy-blue line, (21)). Notably, parental income when children are in their teens has the greatest effect on their child attending a higher-level educational institution. Our findings therefore suggest that high parental income can reduce the likelihood of the children attending a 2-year college or below, and the teen years are crucial for such educational attainment.

4.1.3 Family Structure Trajectory and Offspring's Educational Attainment

Since a child has been exposed to parental income and family structure during their childhood years, it would also be meaningful to investigate the association between family structure trajectory and offspring's college attendance. To this end, we extend the functional multi-

Figure 10: Ratio of Single-Parent Families



Notes: The navy line plots the ratio of single-parent families at each child's age for the birth cohorts 1997-1999. The red-dashed line indicates the average ratio of single-parent families which is 5.98%.

nomial logit model by considering an additional functional regressor, a trajectory of family structure. Figure 10 presents the ratio of single-parent families in the sample families of the birth cohorts 1997–1999. The ratio of single-parent families, the navy-blue line in Figure 10, ranges from 1.95% (at child's age 0, 1, and 2) to 12.50% (at child's age 19), and its average, presented by the red dashed line, is 5.98%.

We consider the following functional multinomial logit model,⁹

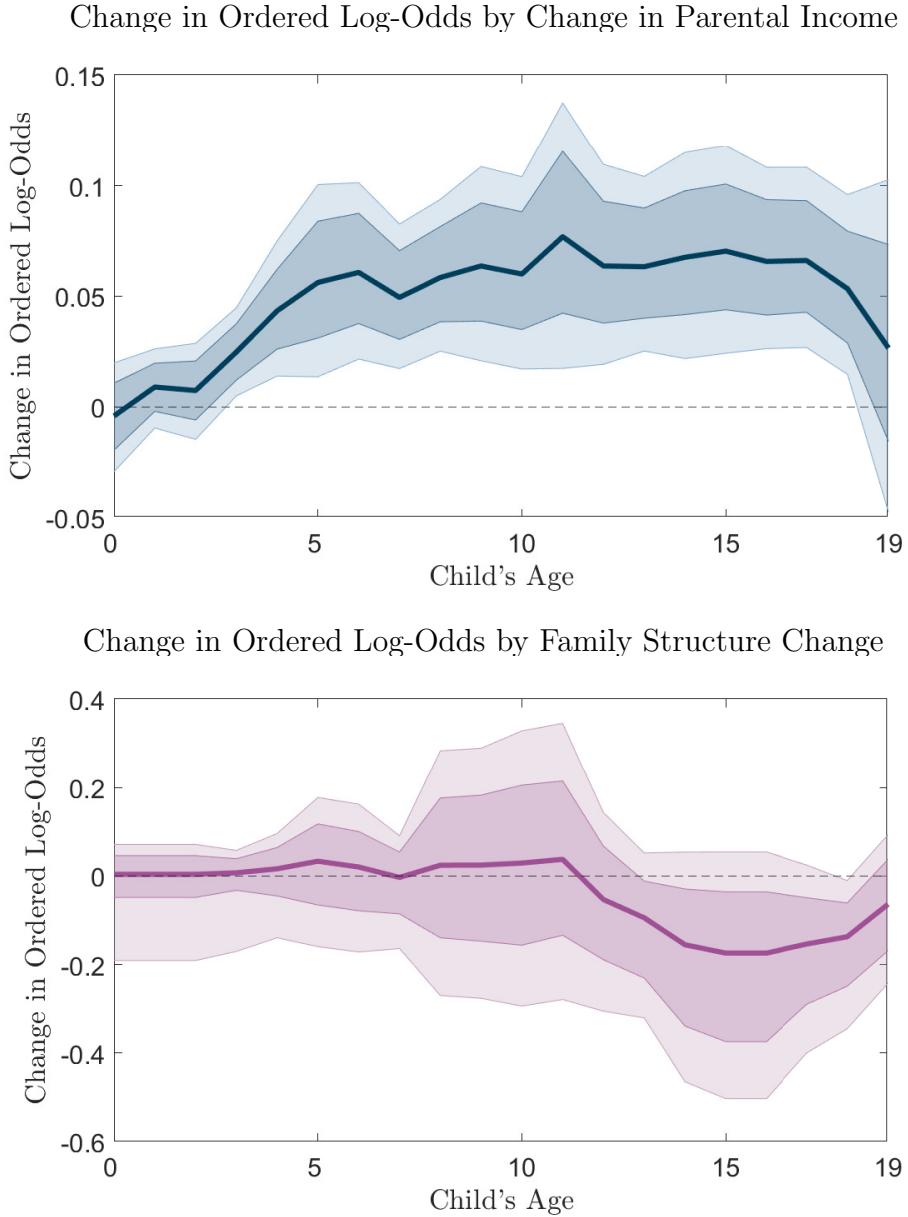
$$z_i = \begin{cases} 1 & \text{if } z_i^* \leq \tau_1 \\ 2 & \text{if } \tau_1 < z_i^* \leq \tau_2, \quad z_i^* = \int_0^{19} \mu(r)f_i(r)dr + \int_0^{19} \nu(r)g_i(r)dr + \varepsilon_i \\ 3 & \text{if } \tau_2 < z_i^* \end{cases} \quad (22)$$

where (f_i) and (g_i) are the trajectories of parental income and family structure, respectively, and (z_i) is defined similarly as in (18).

Figure 11 presents the estimated change in the log-odds ratio by parental income trajectory $\hat{\mu}(\cdot)$ in the upper panel, and the change in the log-odds ratio by the trajectory of family structure $\hat{\nu}(\cdot)$ in the lower panel. Note that even though we consider an additional functional regressor (g_i) to the model, there is no significant difference to estimate $\hat{\mu}(\cdot)$. Notice that $\hat{\mu}(\cdot)$ in (22) presented in Figure 8 is similar to $\hat{\mu}(\cdot)$ in (18) presented in the upper panel of

⁹Based on the leave-one-out cross-validation error, the number of functional principal components m are chosen as 3 and 4 for (f_i) and (g_i) , respectively.

Figure 11: Parental Income and Family Structure Trajectories and College Attendance



Notes: This figure presents the effect of the parental income trajectory and family structure trajectory on the ordered log-odds ratio ($\hat{\mu}(\cdot)$ and $\hat{\nu}(\cdot)$ in (18), respectively). The upper panel presents the estimated $\hat{\mu}(\cdot)$, which measures the parental income trajectory effect on ordered log-odds. The lower panel plots $\hat{\nu}(\cdot)$, which measures the family structure trajectory effect on ordered log-odds. The darker (lighter) shaded areas indicate the 68% (90%) bootstrap confidence intervals.

Figure 11. A plausible explanation is that these two functional regressors are not highly correlated. The maximum canonical correlation between the functional principal components of the functional covariates (f_i) and (g_i) is only 0.29, which is not substantial. The lower panel of Figure 11 presents the $\hat{\xi}(\cdot)$ estimated in (22). The effect of family structure on college attendance is not significant until the early teens, after which, it becomes significantly negative. These results imply that living in a single-parent family during teens tends to hamper children's access to better tertiary education institutions.

Next, we consider the marginal effect of family structure on the probability of their offspring attending a level k institution. Since the family structure variable is not continuous, we define the marginal effect of the family structure at a child's age s on the probability of a child going to a level k educational institution for $k = 1, 2, 3$ as follows.

$$\mathbb{P}(z_i = k | f_i = \bar{f}, g_i = \delta_s) - \mathbb{P}(z_i = k | f_i = \bar{f}, g_i = 0) \quad (23)$$

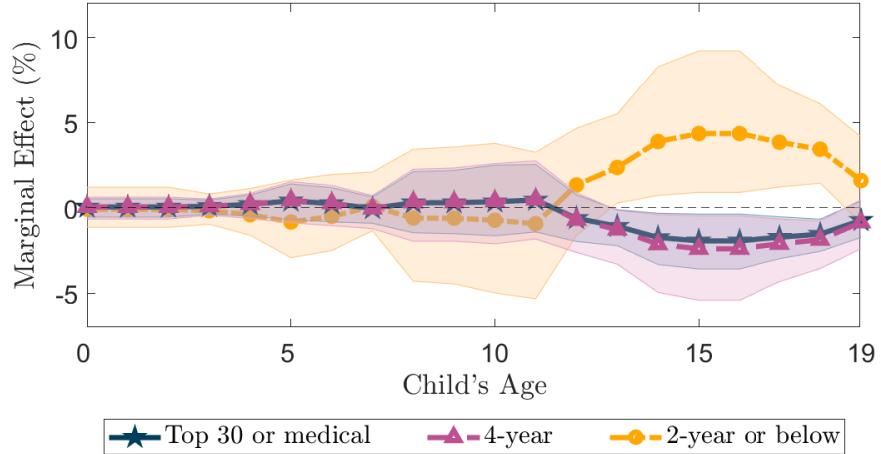
for ages $s = 0, \dots, 19$, and where $\delta_s(\cdot)$ is a hypothetical family structure trajectory, which is defined as

$$\delta_s(r) = \begin{cases} 1 & \text{if } r = s \\ 0 & \text{otherwise.} \end{cases}$$

In particular, $g_i(r) = \delta_s(r)$ describes the hypothetical case where child i lives in a single-parent family only at age s . By contrast, $g_i(r) = 0$ indicates the case that son i lived in a two-parent family and his parents were never divorced or separated. The equation (23) measures the difference between two predicted probabilities of a child attending a specific type of tertiary educational institution when parental income trajectory is fixed at its average. The first probability is obtained when the child lived in a single-parent family only at age s and the second probability is obtained assuming the child lived in a two-parent family for 20 years. A comparison of these two predicted probabilities enables measuring the marginal effect of family structure at each age of the child on the probability of the offspring attending a level k institution.

Figure 12 plots the marginal effect of family structure at each age of the child. According to these results, the effect of a change in the family structure on college attendance depends on when it happens. When children are young, living in a single-parent family does not seem to affect their educational attainment. However, in their teen years, absence of one parent strongly influences their college attendance. It decreases the probability of attending a 4-year college or above, while it increases the probability of attending a 2-year college or ending with a high school degree. Our results imply that the mid-teens are more critical than other periods for children's educational attainment. Therefore, we may say that

Figure 12: Age-Varying Marginal Effect of Family Structure on College Attendance



Notes: This figure presents the marginal effect of family structure on the probability of their offspring attending to a level k institution, (23). The navy-star line presents the marginal effect of parental income on attending top 30 universities or medical schools. The purple-triangular line and the yellow-circle line indicate the marginal effect of parental income on attending other 4-year colleges and attending 2-year colleges or below, respectively. The shaded areas indicate 68% bootstrap confidence intervals.

this period is important for creating a channel through which college attendance influences intergenerational economic mobility.

4.2 Occupational Choices

An individual's occupation is one of the decisive factors that determine one's income level. To find a possible explanation for why the early 20s are crucial for intergenerational mobility, we examine the association between parental influence and children's occupational choices. Parents can influence children's occupational choices by, for example, acting as role models or offering support as they find a desirable career.

This section investigates the relationship between offspring's occupational choices and trajectories such as parental income and family structure. More specifically, we explore how the trajectories of parental income and family structure affect children's occupation type and intergenerational occupational mobility. We first categorize offspring's job choices, and subsequently consider a multinomial logit model to explain how parental income and family structure are correlated with an offspring's job choices. In addition, we study the relationship between parents' occupation trajectory and their child's occupation choice to reveal the critical periods of intergenerational occupational mobility.

Table 5: Occupation Classes in Korea

Class 1	Legislators, Senior Officials and Managers, Professionals
Class 2	Technicians and Associate Professionals, Clerk, Craft and Related Trades Workers, Equipment, Machine Operating and Assembling Workers
Class 3	Service Workers, Sales Worker, Skilled Agricultural, Forestry and Fishery Workers, Elementary Workers

Notes: This table presents the occupation classes in Korea defined by [Lee and Min \(2016\)](#). A job in Class 1 requires a high level of education, similar to a white-collar job. Class 2 jobs require a certain amount of training, comparable with skilled workers. Relatively less training is needed for a Class 3 job, comparable with unskilled workers.

4.2.1 Data

Offspring's occupational choice and intergenerational occupational mobility are actively studied areas in intergenerational mobility ([Long and Ferrie \(2013\)](#), [Emran and Shilpi \(2011\)](#)). In our study of occupational mobility, we focus on the transition from a father's occupation to a son's occupation as in other previous studies. Historically in Korea, the employment rate of women was low and the majority of women were housewives. Therefore, exploring occupational mobility from mother to daughter is not feasible, which is why we consider only the transition from father to son. In the KLIPS data, we have 95 such father-son pairs.

[Lee and Min \(2016\)](#) investigates intergenerational occupational mobility in Korea, and defines three types of occupations based on the KLIPS data, which are presented in Table 5. We follow these definitions for occupation types. Professional and administration jobs belong to Class 1, similar to a white-collar job in the typical occupation classification used for jobs in the U.S. As can be seen from Table 6 which presents the summary statistics of income and education levels of fathers in each job class, fathers in Class 1 have a higher income and education level than other classes. Fathers whose jobs belong to the other two classes are comparable to skilled and unskilled workers, respectively. In contrast, fathers in Class 3 tend to earn less than those in the other two classes and most do not enter tertiary education.

Table 6: Summary Statistics

	Class 1	Class 2	Class 3
<u>A. Income</u>			
Average log income	8.73	8.45	8.15
<u>B. Education Level</u>			
High school or below	30%	69.35%	91.3%
2-year college	0%	14.52%	8.7%
4-year college or above	70%	16.13%	0%

Notes: This table reports summary statistics of the father in each job class. Panel A presents the average log income for each job class. Panel B displays the father's education level for each job class.

4.2.2 Parental Income Trajectory and Offspring's Occupational Choice

To investigate the association between the trajectory of parental income and a son's job choice, we consider the following functional multinomial logit model¹⁰

$$\log\left(\frac{\mathbb{P}\{z_i = k\}}{\mathbb{P}\{z_i = 3\}}\right) = \tau_k + \int_{10}^{29} \mu_k(r) f_i(r) dr \text{ for } k = 1, 2, 3, \quad (24)$$

where k is the occupation class a son's chosen job belongs to, and we have

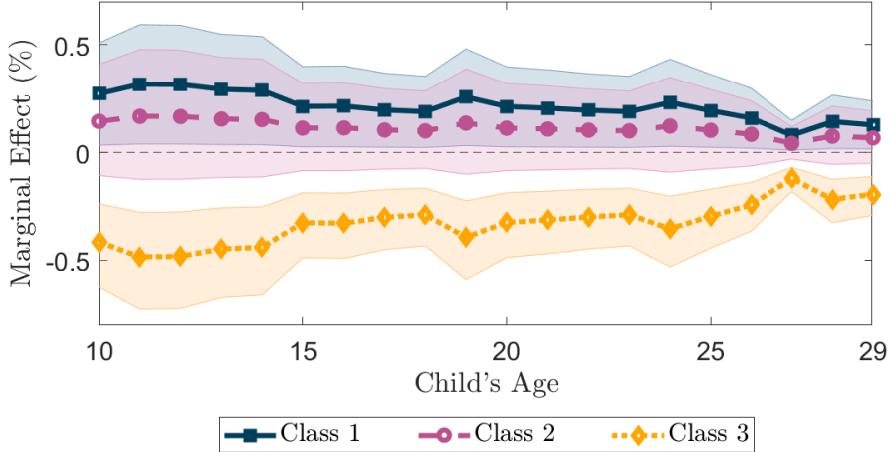
$$z_i = \begin{cases} 1 & \text{if child } i\text{'s job belongs to Class 1} \\ 2 & \text{if child } i\text{'s job belongs to Class 2} \\ 3 & \text{if child } i\text{'s job belongs to Class 3.} \end{cases}$$

Model (24) measures the association between parental income at each age and a son's specific job choice. The interpretation of the logit coefficient in the multinomial logit model is not straightforward, and for this reason, we analyze marginal effects instead. Using the fitted model, we obtain the marginal effect of parental income on a son's job choice probability. Note that

$$\mathbb{P}(z_i = k) = \frac{\phi_k}{\phi_1 + \phi_2 + \phi_3} \text{ for } k = 1, 2, 3$$

¹⁰To estimate this model, the number of functional principal components m is chosen to be 3 by minimizing leave-one-out cross-validation error.

Figure 13: Age-Varying Marginal Effect of Parental Income on Occupational Choices



Notes: This figure presents the marginal effect of family structure on the probability of a son's occupational choices, (25). The navy-square line presents the marginal effect of parental income on a son having a Class 1 job. The purple-circle line and the yellow-diamond line indicate the marginal effect on a son having a job in Class 2 and 3, respectively. The shaded areas indicate 68% bootstrap confidence intervals.

where

$$\phi_k = \begin{cases} \exp\left(\tau_k + \int_{10}^{29} \mu_k(r) f_i(r) dr\right) & \text{for } k = 1, 2 \\ 1 & \text{for } k = 3. \end{cases}$$

The marginal effect of the parental income at child's age s on the probability of their son having a job in class k can be derived by taking derivatives of $\mathbb{P}(z_i = k)$ with respect to $f_i(s)$ as follows

$$\frac{\partial \mathbb{P}(z_i = k)}{\partial f_i(s)} = \mathbb{P}(z_i = k) \left(\mu_k(s) - \sum_{j=1}^3 \mathbb{P}(z_i = j) \mu_j(s) \right) \quad \text{for } k = 1, 2, 3 \quad (25)$$

where $\mu_3 = 0$.

Figure 13 plots the marginal effects of parental income at each age on a son's job. In this figure, the marginal effect is similar for all ages, although the marginal effect at the teens is more substantial than the marginal effect during the late 20s. In addition, as evident in Figure 13, parental income and the probability of having a Class 1 occupation are positively related, and the effects are statistically significant. Similarly, parental income and the probability of the son having a Class 3 job are negatively related. According to our findings, a son from a wealthy family is more likely to get a job in Class 1 which includes mostly professional occupations. We obtain similar results from the ordered logit and probit

models. In these ordered models, Class 1 is considered the highest class, and Class 3 is the lowest class (see Appendix E.5 for details).

4.2.3 Family Structure Trajectory and Offspring's Occupational Choice

Since the family structure is also known to matter for a son's job choice (Biblarz et al. (1997)), we add a functional regressor (g_i) representing a trajectory of family structure to the model (24) and consider the following extended model¹¹

$$\log\left(\frac{\mathbb{P}\{z_i = k\}}{\mathbb{P}\{z_i = 3\}}\right) = \tau_k + \int_{10}^{29} \mu_k(r)f_i(r)dr + \int_{10}^{29} \nu_k(r)g_i(r)dr \text{ for } k = 1, 2 \quad (26)$$

Similarly, we focus on the marginal effects of family structure trajectory on a son's job, and derive the probability of the son's job choice belonging to each occupation class k as follows

$$\mathbb{P}(z_i = k) = \frac{\varphi_k}{\varphi_1 + \varphi_2 + \varphi_3} \text{ for } k = 1, 2, 3,$$

where

$$\varphi_k = \begin{cases} \exp\left(\tau_k + \int_{10}^{29} \mu_k(r)f_i(r)dr + \int_{10}^{29} \nu_k(r)g_i(r)dr\right) & \text{for } k = 1, 2 \\ 1 & \text{for } k = 3. \end{cases}$$

Since $g_i(r)$ is not continuous, we define the marginal effect of the family structure at child's age s on the probability of their son having a job in class k as

$$\mathbb{P}\{z_i = k | f_i = \bar{f}, g_i = \delta_s\} - \mathbb{P}\{z_i = k | f_i = \bar{f}, g_i = 0\} \quad (27)$$

for each job class $k = 1, 2, 3$, and age $s = 10, \dots, 29$, where δ_s and the 0 function are defined as in the previous section. $g_i(r) = \delta_s(r)$ is the case where child i lived in a single-parent family at age s , whereas $g_i(r) = 0$ indicates the case where child i 's parents had never been divorced or separated.

In addition, we compare the following predicted probability of son's occupational choice.

$$\mathbb{P}\{z_i = k | f_i = \bar{f}, g_i = 1_{r \leq s}\} \quad (28)$$

¹¹Based on the leave-one-out cross-validation error of this model, the number of functional principal components are chosen to be 3 and 1 for (f_i) and (g_i) , respectively.

for $k = 1, 2, 3$ and $s = 10, \dots, 29$, where an arbitrary function $1_{r \leq s}(\cdot)$ is defined as

$$1_{r \leq s}(r) = \begin{cases} 1 & \text{if } r \leq s \\ 0 & \text{otherwise} \end{cases}$$

To be specific, $g_i(r) = 1_{r \leq s}(r)$ illustrates the case where a son i had lived in a single-parent family until the age s . For example, $g_i = 1_{r \leq 15}$ implies son i had lived in a single-parent family until he was 15 years old, and after then, he lived with two parents. By changing the value of s in $1_{r \leq s}(\cdot)$, we can measure how the probability of a son's occupation choice evolves as he lived longer under a single-parent family during his childhood years.

The upper panel of Figure 14 plots the marginal effects of living in a single-parent family on a son's job at each age of the child. According to our results from Figure 14, living with only one parent and having a Class 3 job are positively related. As a worker in Class 3 tends to earn less than those in the other classes, living in a single-parent family can be an obstacle to an offspring's economic success. In addition, the marginal effect before the age of 20 is slightly more significant than the effect after the age of 20 years, reflecting the tendency that parental divorce or separation during the teenage affects children more.

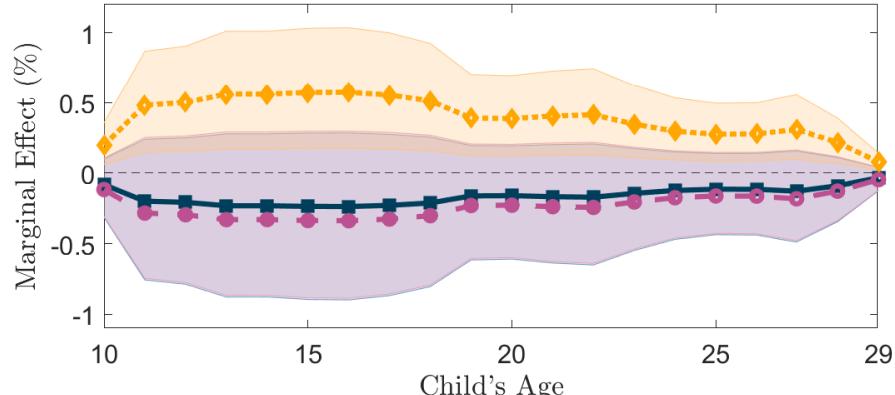
Similarly, the lower panel of Figure 14 plots how the predicted probability of the son's job changes by family structure. For example, if the parents had never divorced or separated, the predicted probability of the son having a Class 3 job is 21.55%. However, the probability of a son having a Class 3 job increases as parental divorce or separation lasts longer. For example, for a son who lived in a single-parent family for 20 years (from age 10 to 29 years), the predicted probability equals 31.91%, approximately increased by 10 percentage points compared to a son who lived in a two-parent family for the same 20 year period.

4.2.4 Intergenerational Occupational Mobility

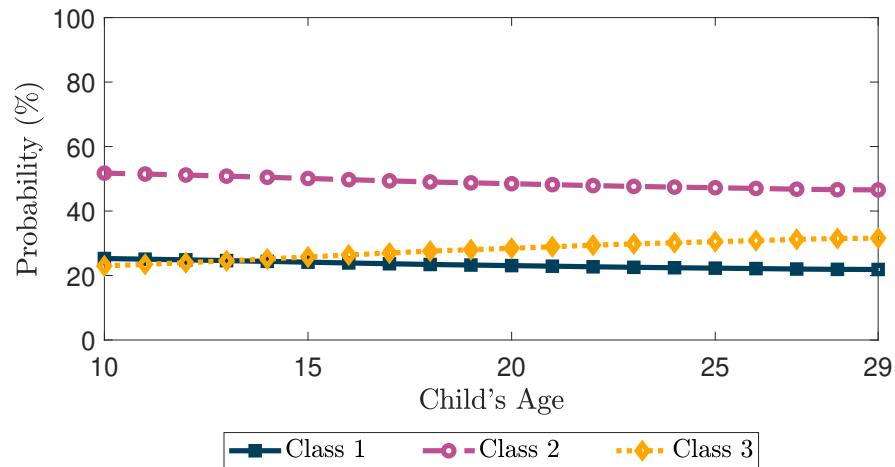
We explore intergenerational occupational mobility using a functional multinomial logit model. The standard approach of intergenerational occupational mobility is to find an occupation transition matrix from father to son. Table 7, which presents the occupational transition matrix from father to son, demonstrates strong evidence of intergenerational occupational mobility. For example, when a father had a Class 1 job, then the probability of his son having the same class job equals 40%. If the father had a different job, the son's probability of having a job in Class 1 is much lower, 21% (30%) if the father had a job in Class 2 (Class 3). Specifically, the diagonal values of the transition matrix (Table 7) implies intergenerational occupational mobility, and these values are greatest in their respective columns. This finding indicates that sons tend to have the same job class as their fathers.

Figure 14: Family Structure Trajectory and Offspring’s Occupational Choices

Age-Varying Marginal Effect of Family Structure on Occupational Choices



Changes in Probability of Offspring’s Occupational Choice by Duration of Single-parent Family



Notes: The upper panel depicts the marginal effect of family structure at each child’s age on the probability of the offspring’s occupational choices (27). The navy-square line presents the marginal effect of family structure trajectory on the probability of a son having a Class 1 job. The purple-circle line and the yellow-diamond line indicate the marginal effect on the probability of a son having a job in Class 2 and Class 3, respectively. The shaded areas indicate 68% bootstrap confidence intervals. The lower panel shows how the predicted probability of the son’s occupational choices (28) evolves as the son lives in a single-parent family for longer. The navy-square line presents the predicted probability of a son having a Class 1 job. The purple-circle line and the yellow-diamond line indicate the probability of having a job in Class 2 and Class 3, respectively.

Table 7: Occupational Transition Matrix

		To Son	Class 1	Class 2	Class 3
From Father \					
		Class 1	40%	50%	10%
		Class 2	21%	56%	23%
		Class 3	30%	40%	30%

Notes: This table presents the transition matrix from a father's job class to a son's job class. Both job classes are measured as the mode of their job class over a given period.

However, this standard approach based on the occupational transition matrix can only explain whether the transition from the father's job to the son's job frequently occurs or not. In contrast, our approach can explore a critical period for intergenerational occupational mobility and help reveal the mechanism through which this transmission occurs. To this end, we study the association between the father's occupational class trajectory, from the child's age 10 to 29 years, and the son's job type in his early 30s. Our functional approach allows us to study when a child's occupational choice is most critically influenced by parental income and family structure. To quantify this influence explicitly, we fit the following functional multinomial logit model¹² for each job type j of the father

$$\log\left(\frac{\mathbb{P}\{z_i = k\}}{\mathbb{P}\{z_i = 3\}}\right) = \tau_k^j + \int_{10}^{29} \mu_k^j(r)f_i(r)dr + \int_{10}^{29} \nu_k^j(r)g_i^j(r)dr \text{ for } k = 1, 2, \quad (29)$$

where j and k denote, respectively, a father and son's occupational types,

$$z_i = \begin{cases} 1 & \text{if child } i\text{'s job belongs to Class 1} \\ 2 & \text{if child } i\text{'s job belongs to Class 2} \\ 3 & \text{if child } i\text{'s job belongs to Class 3,} \end{cases}$$

(f_i) is the trajectory of parental income, and (g_i^j) is the trajectory indicating whether the father's job belongs to Class j at each age of his son. Specifically, we define

$$g_i^j(r) = \begin{cases} 1 & \text{if occupation of } i\text{'s father belongs to class } j \text{ at child's age } r \\ 0 & \text{otherwise} \end{cases}$$

¹²The number m of the functional principal components used for all our analyses are chosen by minimizing the leave-one-out cross-validation error. Chosen are $m = 3$ for (f_i) , $m = 2$ for (g_i^1) , $m = 2$ for (g_i^2) , and $m = 1$ for (g_i^3) .

for $r \in [10, 29]$.

The intuition for using (g_i^j) s in place of a father's occupational type trajectory can be viewed similar to using a dummy variable as a regressor for the qualitative variable. As the occupational type has three categories, we have three dummy-like functional regressors (g_i^j) for $j = 1, 2, 3$, then we estimate the functional multinomial logit model for each occupational type of the father. Once the model is estimated, we derive the marginal effect of the father's job class on the probability of his son's occupational choice and how the father's job type trajectory can change the probability of the son's occupation choice. Note that

$$\mathbb{P}(z_i = k) = \frac{\psi_k}{\psi_1 + \psi_2 + \psi_3} \text{ for } k = 1, 2, 3$$

where

$$\psi_k = \begin{cases} \exp\left(\tau_k^j + \int_{10}^{29} \mu_k^j(r) f_i(r) dr + \int_{10}^{29} \nu_k^j(r) g_i^j(r) dr\right) & \text{for } k = 1, 2 \\ 1 & \text{for } k = 3 \end{cases}$$

As $g_i^j(r)$ is a dummy at child's age r , we define the marginal effect of the father's occupation type j at child's age s on the probability of their son having a job in class k as

$$\mathbb{P}\{z_i = k | f_i = \bar{f}, g_i^j = \delta_s\} - \mathbb{P}\{z_i = k | f_i = \bar{f}, g_i^j = 0\} \quad (30)$$

for each father and son's job class $j, k = 1, 2, 3$ and age $s = 10, \dots, 29$, where δ_s and the 0 function are defined as in the previous section. Note that $g_i^j(r) = \delta_s(r)$ indicates the case where i 's father's has a job in Class j at child's age s , whereas $g_i^j(r) = 0$ describes the case where i 's father never had a job in Class j . A comparison of the predicted probability between these two cases enables quantifying the marginal effect of the father's job class on the probability of son's occupation choice at each age of child.

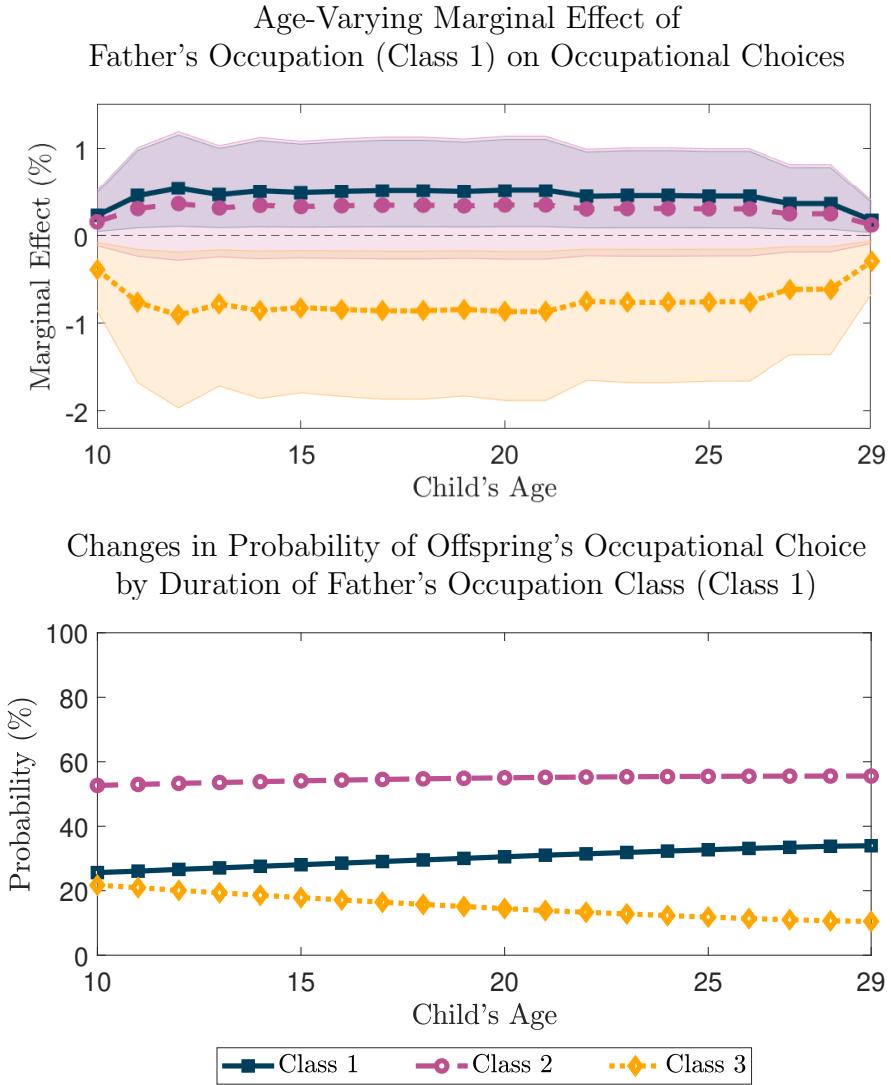
We also consider the following predicted probability of the son's occupational choice.

$$\mathbb{P}\{z_i = k | f_i = \bar{f}, g_i^j = 1_{r \leq s}\} \quad (31)$$

for each father and son's job class $j, k = 1, 2, 3$ and age $s = 10, \dots, 29$, where $1_{r \leq s}$ is defined as in the previous section. Note $g_i^j(r) = 1_{r \leq s}(r)$ implies that i 's father had worked in Class j until child's age s . For example, $g_i^1(r) = 1_{r \leq 20}$ indicates the father had kept a job in Class 1 until the child's age of 20 years, after which he takes up a different job.

The upper panel of Figure 15 presents the marginal effect of a father having a job in Class 1 on the probability of son's occupational choice at each child's age. The navy-square line indicates the marginal effect on the probability of a son having a job in Class 1. For all

Figure 15: Father's Occupation Trajectory and Offspring's Occupational Choice - Class 1



Notes: The upper panel depicts the marginal effect of the father's occupation (in Class 1) on the probability of the offspring's occupational choices (30) for $j = 1$. The navy-square line presents the marginal effect of the father having a job in Class 1 on the probability of the son having a Class 1 job. The purple-circle line and the yellow-diamond line indicate the marginal effect on the probability of the son having a job in Class 2 and Class 3, respectively. The shaded areas indicate 68% bootstrap confidence intervals. The lower panel shows how the predicted probability of the son's occupational choices (31) evolves as the father retains the Class 1 job for longer. The navy-square line presents the predicted probability of a son having a Class 1 job. The purple-circle line and the yellow-diamond line indicate the probability of a son having a job in Class 2 and Class 3, respectively.

ages of the child, the marginal effect is positive, and the effect is especially more prominent from the mid-teens to the mid-20s. By contrast, other marginal effects on the probability of having a Class 2 or Class 3 job, tends to decrease during this period.

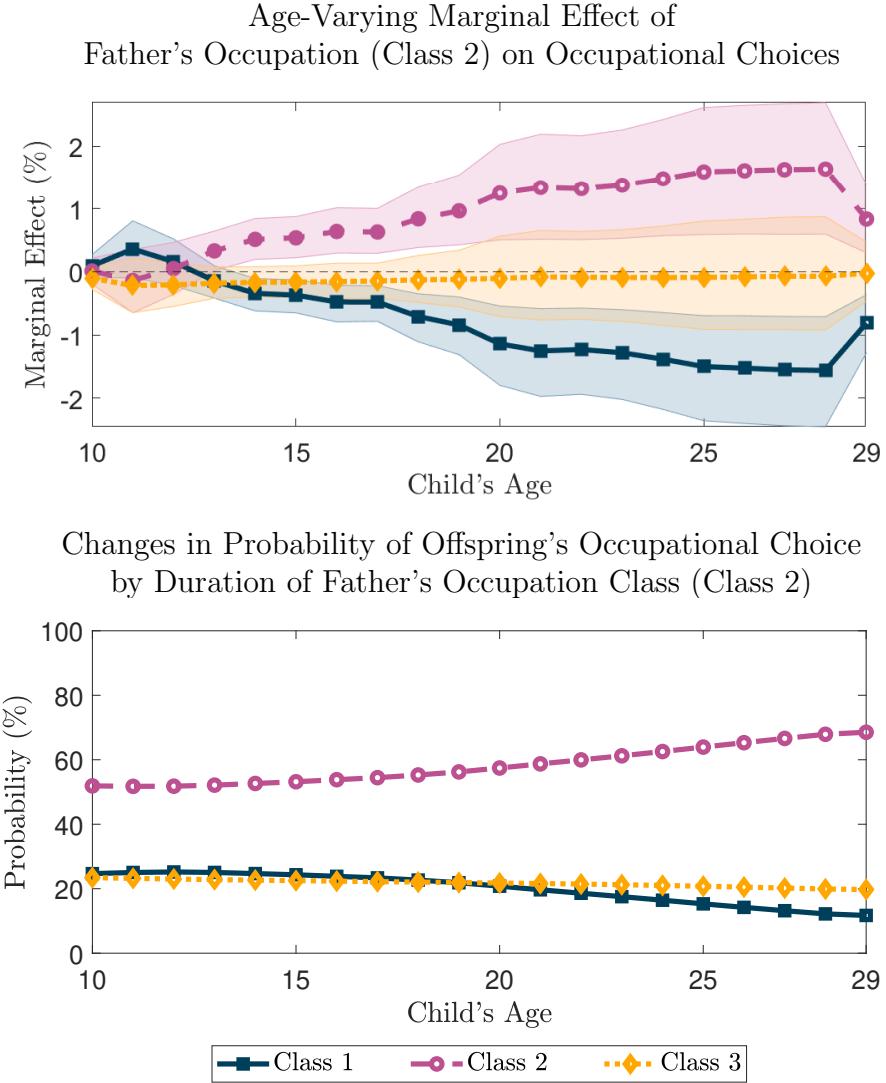
The lower panel of Figure 15 presents how the predicted probability of a son’s job evolves as the father retains a job in Class 1 for longer. For example, if the father never had a Class 1 job, the predicted probability of the son having a Class 1 job is 25.10%. However, if the father held a Class 1 job during the entire 20-year period (from child’s age of 10 to 29 years), the probability of the son having a Class 1 job increases from 25.10% to 38.36%. Thus, the probability of the son having a Class 1 job increases by 13.26 percentage points if the father has a Class 1 job for 20 years. As evident in the upper panel of Figure 15, the transition probability increases faster after the son’s age 20.

A similar finding is evident in Figure 16. The upper panel plots the marginal effect of a father having a job in Class 2 on the probability of his son’s occupational choice at each age of the son. The purple-circle line indicates the marginal effect on the probability that the son has a job in Class 2. The marginal effect is positive overall, and it is larger, especially when the son is in the 20s. Our findings from the upper panels of 15 and 16 imply that the 20s are crucial for a son to have a job in Class 1 and Class 2, and a high level of education or training is required to have a job in these classes.

The lower panel of Figure 16 presents how the predicted probability of a son’s job evolves as the father retains a job in Class 2 longer. If the father has never had a Class 2 job, the probability of the son getting a Class 2 job is 53.45%. However, if the father has a 20-year career in a Class 2 job, then the probability of the son having a Class 2 job rises to 70.68%. The transition probability thus increases by 17.23 percentage points in this case. Similar to the navy-square line in the lower panel of Figure 15, the transition probability tends to increase sharply when the son is in his 20s.

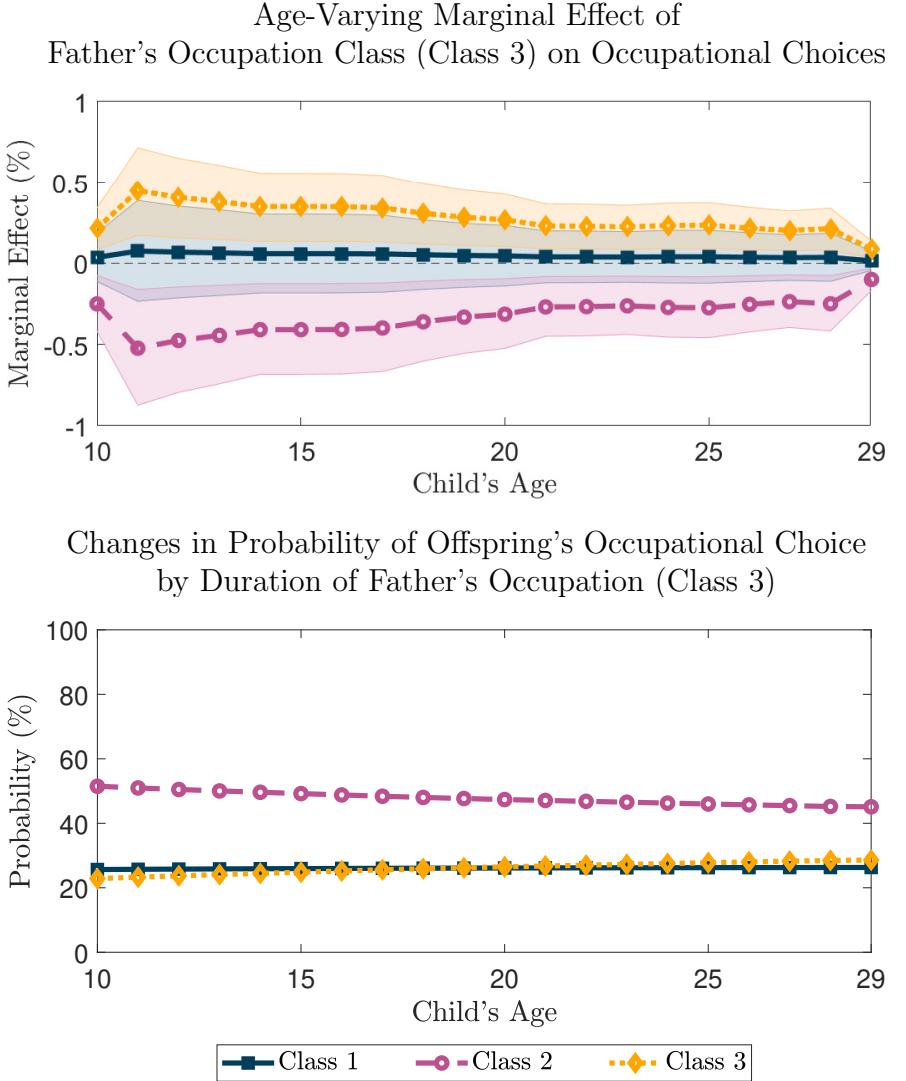
The upper panel of Figure 17 illustrates the marginal effect of a father having a job in Class 3 on the probability of his son’s occupational choice at each age. The marginal effect on the probability of the son having a job in Class 3, the yellow-diamond line, is positive, and the effect is almost equal for all ages of the child. In the lower panel of Figure 17, which presents how the predicted probability of a son’s job evolves as the father retains a job in Class 3 longer, no drastic change is observed. However, the probability of having the same job type as the father increases slightly.

Figure 16: Father's Occupation Trajectory and Offspring's Occupational Choice - Class 2



Notes: The upper panel depicts the marginal effect of the father's occupation (in Class 2) on the probability of the offspring's occupational choices (30) for $j = 2$. The navy-square line presents the marginal effect of the father having a job in Class 2 on the probability of the son having a Class 1 job. The purple-circle line and the yellow-diamond line indicate the marginal effect on the probability of the son having a job in Class 2 and Class 3, respectively. The shaded areas indicate 68% bootstrap confidence intervals. The lower panel shows how the predicted probability of the son's occupational choices (31) evolves as the father retains the Class 2 job for longer. The navy-square line presents the predicted probability of a son having a Class 1 job. The purple-circle line and the yellow-diamond line indicate the probability of a son having a job in Class 2 and Class 3, respectively.

Figure 17: Father's Occupation Trajectory and Offspring's Occupational Choice - Class 3



Notes: The upper panel depicts the marginal effect of the father's occupation (in Class 3) on the probability of offspring's occupational choices (30) for $j = 3$. The navy-square line presents the marginal effect of the father having a job in Class 3 on the probability of the son having a Class 1 job. The purple-circle line and the yellow-diamond line indicate the marginal effect on the probability of son having a job in Class 2 and Class 3, respectively. The shaded areas indicate 68% bootstrap confidence intervals. The lower panel shows how the predicted probability of the son's occupational choices (31) evolve as the father retains the Class 3 job longer. The navy-square line presents the predicted probability of a son having a Class 1 job. The purple-circle line and the yellow-diamond line indicate the probability of a son having a job in Class 2 and Class 3, respectively.

5 Conclusion

In this paper, we study intergenerational mobility in Korea using a functional regression approach. There is a marked contrast between the standard approach (1) and the functional approach (16). The estimated IGE $\widehat{\beta}(\cdot)$ in (1), using KLIPS data, equals to 0.12. That is the only information that the standard approach can provide; the size of the correlation between parents and their children's permanent income. However, we can enjoy a variety of benefits from the functional approach.

By employing a functional approach, we find numerous interesting patterns of intergenerational mobility in Korea. In intergenerational income mobility, first, we find the association between parental income and their children's income varies with children's age. In particular, parental income after child's age 15 turns out to be more productive on forming their children's human capital. Second, we also investigate how parental income is linked to offspring's college attendance and occupation choice as a possible mechanism. It turns out that parental income matters for these two outcomes, so these findings hint at a possible channel of intergenerational income mobility since these intermediate outcomes are crucial to determining one's income. In addition, we also explore the relations between family structure and children's outcomes and intergenerational occupational mobility and find that these effects also vary with a child's age.

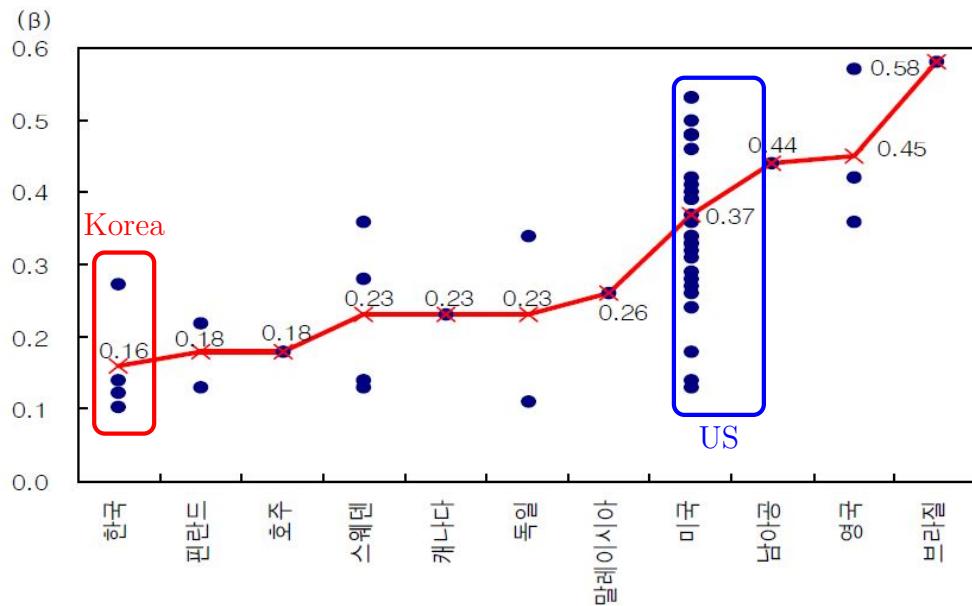
In future work, we will investigate another possible mechanism that can explain the pattern of intergenerational mobility in Korea. In particular, the Korean context provides a unique way to understand how parental tutoring investments affect mobility since the spending on private tutoring is a direct measure of parents' choice and a realization of parents' education fever, a unique characteristic of Korea.

Appendix

A Characteristics of Korea

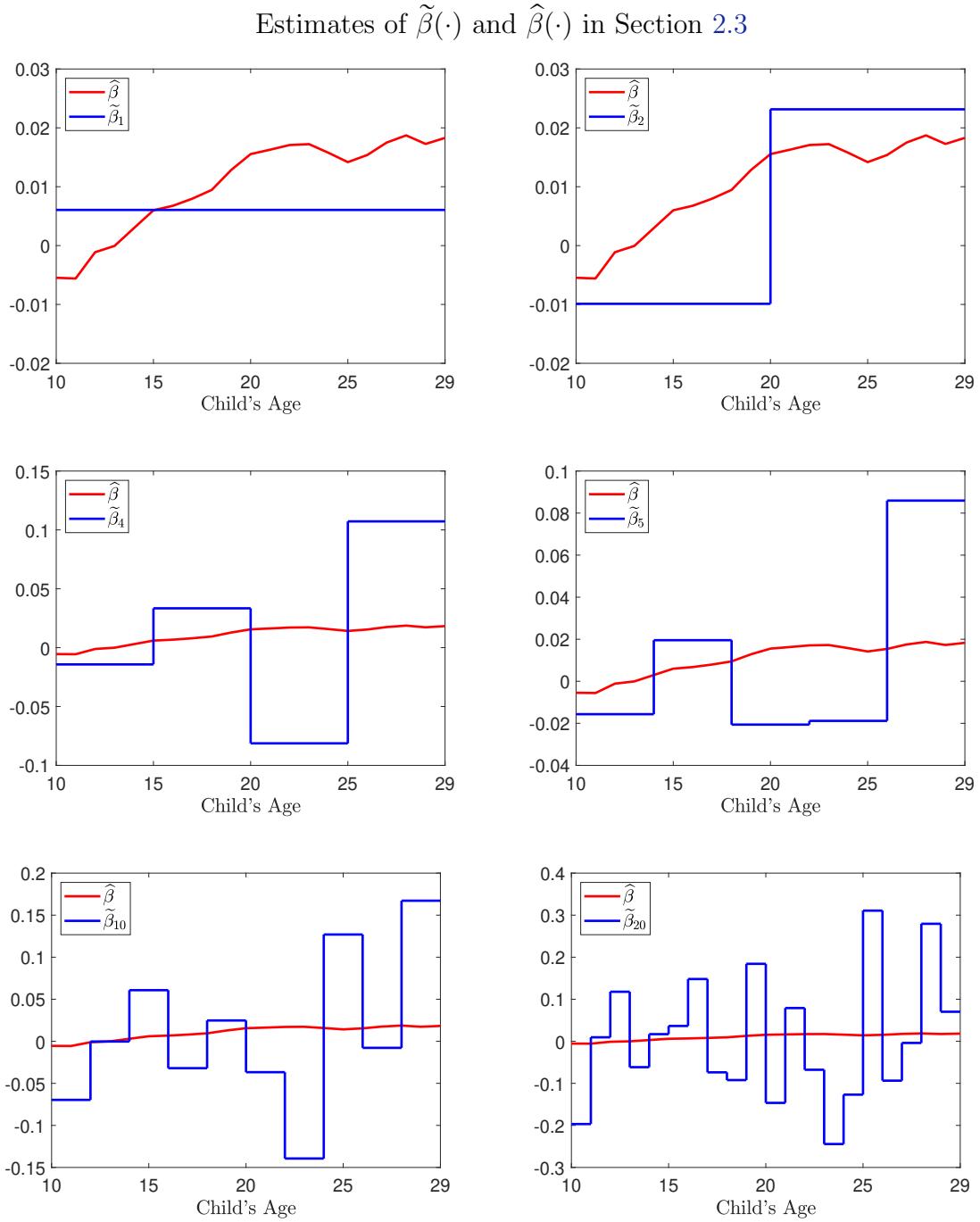
A.1 IGE of Income Estimate

Kim (2009) for Korea, Solon (1999), Solon (2002) for US and other countries



Notes: Each point indicates the estimated IGE of income (β in the model (1)) in the previous literature. The median is marked as \times . Source: [Kim \(2009\)](#)

B Functional vs. Conventional Regression Estimates



Notes: This figure compares functional regression coefficient $\hat{\beta}(\cdot)$ (red line) and $\tilde{\beta}(\cdot)$ (blue line) in (11). $\hat{\beta}(\cdot)$ is estimated using two leading principal components ($m = 2$), while $\tilde{\beta}(\cdot)$ is obtained with various choice of m . Each panel of this figure compares $\hat{\beta}$ and $\tilde{\beta}_m$ with a choice of $m = 1$ (upper left), 2, 4, 5, 10, 20(lower right).

C Technical Details

C.1 Derivations of Econometric Results in Section 2.3

Note that

$$\begin{aligned} (\tilde{\beta}) &= \left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} \left(\sum_{i=1}^n (x_i)y_i \right) \\ &= (\beta) + \left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} \left(\sum_{i=1}^n (x_i)\langle \beta, (1 - \Pi_m)x_i \rangle \right) + \left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} \left(\sum_{i=1}^n (x_i)\varepsilon_i \right) \end{aligned}$$

and

$$(\hat{\beta}) = \left(\sum_{i=1}^n (x_i)^*(x_i)^{*'} \right)^{-1} \left(\sum_{i=1}^n (x_i)^*y_i \right) = (\beta)^* + \left(\sum_{i=1}^n (x_i)^*(x_i)^{*'} \right)^{-1} \left(\sum_{i=1}^n (x_i)^*\varepsilon_i \right),$$

since

$$y_i = \langle \beta, x_i \rangle + \varepsilon_i = \langle \beta, \Pi_m x_i \rangle + \langle \beta, (1 - \Pi_m)x_i \rangle + \varepsilon_i = (\beta)'(x_i) + \langle \beta, (1 - \Pi_m)x_i \rangle + \varepsilon_i$$

We define

$$\tilde{\beta} = \pi^{-1}((\tilde{\beta})) \quad \text{and} \quad \hat{\beta} = \pi^{*-1}((\hat{\beta})).$$

where π and π^* are the isometry, $\pi : H_m \rightarrow R^m$ and $\pi^* : H_m^* \rightarrow R^m$, respectively. Therefore, we have

$$\begin{aligned} \|\tilde{\beta} - \mathbb{E}\tilde{\beta}\| &= \|(\tilde{\beta}) - \mathbb{E}(\tilde{\beta})\| = \left\| \left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} \left(\sum_{i=1}^n (x_i)\varepsilon_i \right) \right\| \\ \|\mathbb{E}\tilde{\beta} - \bar{\beta}\| &= \|\mathbb{E}(\tilde{\beta}) - (\beta)\| = \left\| \left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} \left(\sum_{i=1}^n (x_i)\langle \beta, (1 - \Pi_m)x_i \rangle \right) \right\|, \end{aligned}$$

and

$$\|\hat{\beta} - \bar{\beta}^*\| = \|(\hat{\beta})^* - (\beta)^*\| = \left\| \left(\sum_{i=1}^n (x_i)^*(x_i)^{*'} \right)^{-1} \left(\sum_{i=1}^n (x_i)^*\varepsilon_i \right) \right\|.$$

By assuming the standard assumptions on the classical regression hold for our functional regressions with $\mathbb{E}(\varepsilon_i^2) = \sigma^2$, we have

$$\begin{aligned}\mathbb{E} \left\| \left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} \left(\sum_{i=1}^n (x_i)\varepsilon_i \right) \right\|^2 &= \sigma^2 \operatorname{trace} \left(\left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} \right) \\ \mathbb{E} \left\| \left(\sum_{i=1}^n (x_i)^*(x_i)^{*'} \right)^{-1} \left(\sum_{i=1}^n (x_i)^*\varepsilon_i \right) \right\|^2 &= \sigma^2 \operatorname{trace} \left(\left(\sum_{i=1}^n (x_i)^*(x_i)^{*'} \right)^{-1} \right)\end{aligned}$$

Moreover, we have

$$\left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} \left(\sum_{i=1}^n (x_i) \langle \beta, (1 - \Pi_m)x_i \rangle \right) = \left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} \left(\sum_{i=1}^n (x_i) \langle \beta, x_i \rangle \right) - (\beta)$$

Therefore,

$$\mathbb{E} \|\hat{\beta} - \bar{\beta}^*\|^2 = \mathbb{E} \|\hat{\beta} - \mathbb{E}\hat{\beta}\|^2 = \sigma^2 \operatorname{trace} \left(\left(\sum_{i=1}^n (x_i)^*(x_i)^{*'} \right)^{-1} \right) \quad (13)$$

and

$$\mathbb{E} \|\tilde{\beta} - \bar{\beta}\|^2 = \mathbb{E} \|\tilde{\beta} - \mathbb{E}\tilde{\beta}\|^2 + \|\mathbb{E}\tilde{\beta} - \bar{\beta}\|^2 = \sigma^2 \operatorname{trace} \left(\left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} \right) + \|\operatorname{bias}(\tilde{\beta})\|^2 \quad (14)$$

where

$$\operatorname{bias}(\tilde{\beta}) = \left(\sum_{i=1}^n (x_i)(x_i)' \right)^{-1} \left(\sum_{i=1}^n (x_i) \langle \beta, x_i \rangle \right) - (\beta).$$

D Leave-One-Out Cross-Validation

The number of functional principal components m is chosen by minimizing leave-one-out cross-validation (LOOCV) error. We consider two types of model in this paper, linear regression and discrete choice model with functional covariates, and the LOOCV error is defined differently on the type of model. This section illustrates how to conduct LOOCV for regression with one functional covariate. We use the same approach for the regression with multiple functional covariates.

D.1 Linear Regression with Functional Covariates

Consider a functional linear regression (16).

$$y_i = \alpha + \int_{10}^{29} \beta(r) f_i(r) dr + \varepsilon_i \quad (16)$$

where (y_i) is average log income of child (between 30 and 32 years old), (f_i) is a trajectory of log parental income for 20 years (from a child's age 10 to 29). We calculate LOOCV error for each number of functional principal component $m = 1, \dots, M$ and find m via minimization of LOOCV error. The LOOCV error is calculated as follows.

For $i = 1, \dots, n$, we do the followings:

1. Fit the model (16) without i -th observation, and obtain estimates, $\hat{\alpha}_{-i}$ and $\hat{\beta}_{-i}$.
2. Calculate the fitted value of i -th observation, \hat{y}_i ,

$$\hat{y}_i = \hat{\alpha}_i + \int_{10}^{29} \hat{\beta}_{-i}(r) f_i(r) dr$$

3. Measure the squared difference between actual value y_i and fitted value \hat{y}_i

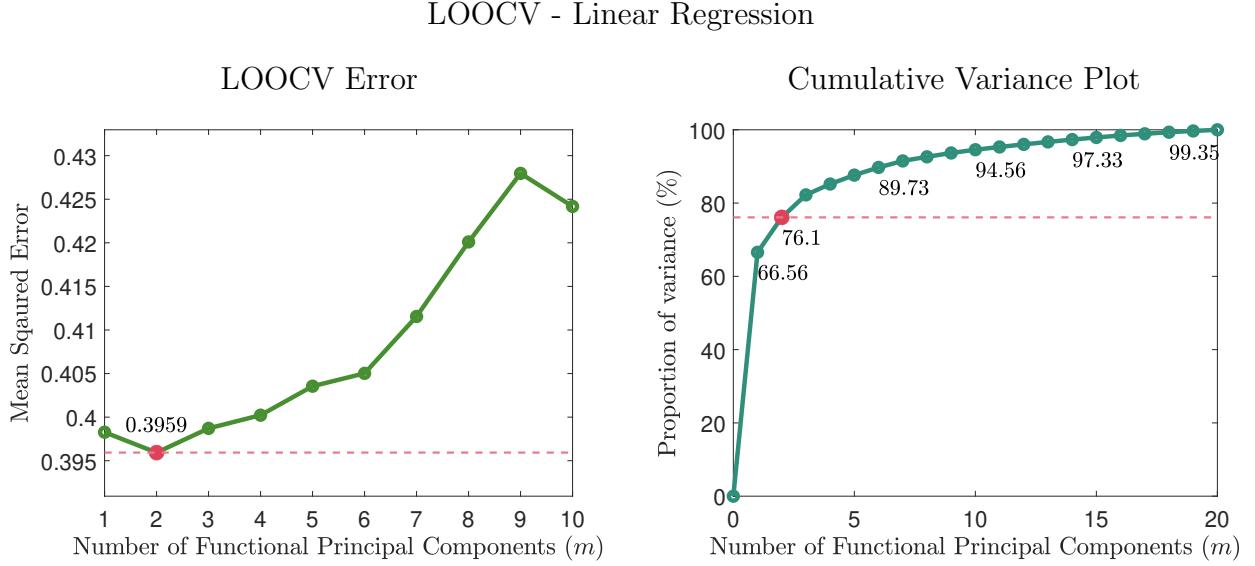
$$L(y_i, \hat{y}_i) = (y_i - \hat{y}_i)^2$$

Lastly, obtain the average of $L(y_i, \hat{y}_i)$.

$$L_m = \frac{1}{n} \sum_{i=1}^n L(y_i, \hat{y}_i)$$

In sum, we calculate the average LOOCV mean squared error for each m and find the minimum. For example of regression (18), we calculate LOOCV error for $m = 1, \dots, 10$, and

set $m = 2$ where the LOOCV error attains the minimum. The first two functional principal components explain 76.10% of the total variation of the functional covariate.



Notes: The left panel presents the LOOCV error. The minimum is attained at $m = 2$. The right panel shows the cumulative scree plot. This panel plots the total percentage of variation explained by the number of functional principal components. When $m = 2$, 76.10% of total variation is explained by the first two components.

D.2 Discrete Choice Model with Functional Covariates

Consider a logit model with the ordered dependent variable (18)¹³.

$$z_i = \begin{cases} 1 & \text{if } z_i^* \leq \tau_1 \\ 2 & \text{if } \tau_1 < z_i^* \leq \tau_2, \quad z_i^* = \int_0^{19} \mu(r) f_i(r) dr + \varepsilon_i \\ 3 & \text{if } \tau_2 < z_i^* \end{cases} \quad (18)$$

where

$$z_i = \begin{cases} 1 & \text{if child } i \text{ attends a 2-year college or below} \\ 2 & \text{if child } i \text{ attends one of other 4-year colleges} \\ 3 & \text{if child } i \text{ attends top 30 universities or medical school} \end{cases}$$

and (f_i) is a trajectory of log parental income for 20 years (from a child's age 0 to 19). Similarly, we calculate LOOCV error for each number of functional principal component $m = 1, \dots, M$ and find m via minimization of LOOCV error. The LOOCV error is calcu-

¹³It can be easily applied to multinomial logit model.

lated as follows.

For $i = 1, \dots, n$, we do the followings:

1. Fit the model (18) without i -th observation, and obtain estimates, $\hat{\tau}_{1,-i}$, $\hat{\tau}_{2,-i}$ and $\hat{\mu}_{-i}$.
2. Calculate the fitted value of i -th observation, \hat{z}_i ,

$$\hat{z}_i = \begin{cases} 1 & \text{if } z_i^* \leq \tau_{1,-i} \\ 2 & \text{if } \tau_{1,-i} < z_i^* \leq \tau_{2,-i}, \quad z_i^* = \int_0^{19} \hat{\mu}_{-i}(r) f_i(r) dr \\ 3 & \text{if } \tau_{2,-i} < z_i^* \end{cases}$$

3. Measure the difference between actual value z_i and fitted value \hat{z}_i

$$L(z_i, \hat{z}_i) = \begin{cases} 1 & \text{if } z_i \neq \hat{z}_i \\ 0 & \text{if } z_i = \hat{z}_i \end{cases}$$

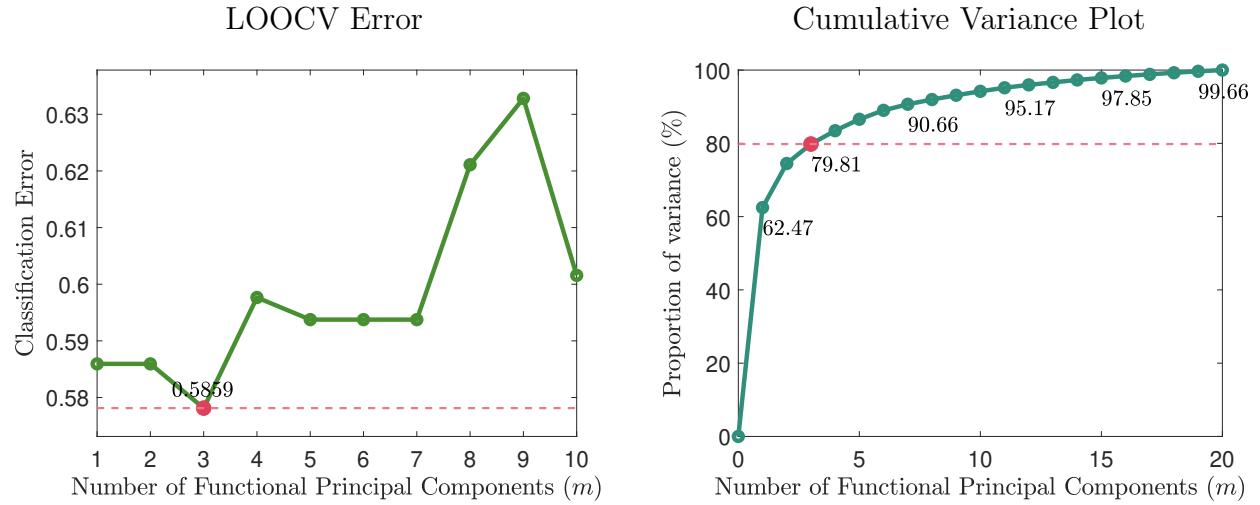
Lastly, obtain the average of $L(z_i, \hat{z}_i)$.

$$L_m = \frac{1}{n} \sum_{i=1}^n L(z_i, \hat{z}_i)$$

In short, we calculate the LOOCV classification error for each m and find the minimum¹⁴. For example of regression (18), we calculate LOOCV error for $m = 1, \dots, 10$, and set $m = 3$ where the LOOCV error attain the minimum. The first three functional principal components explain 79.81% of the total variation of the functional covariate.

¹⁴In other words, we find m that maximizes the accuracy of LOOCV.

LOOCV - Discrete Choice

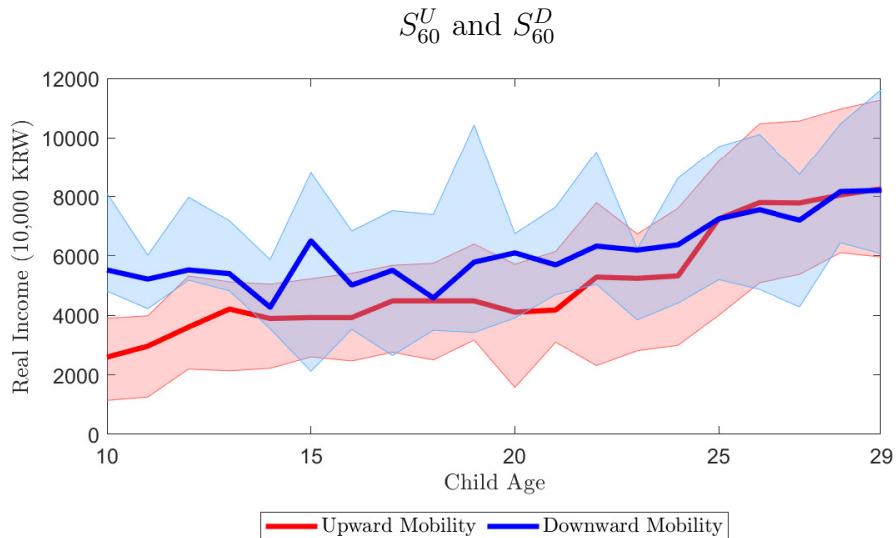
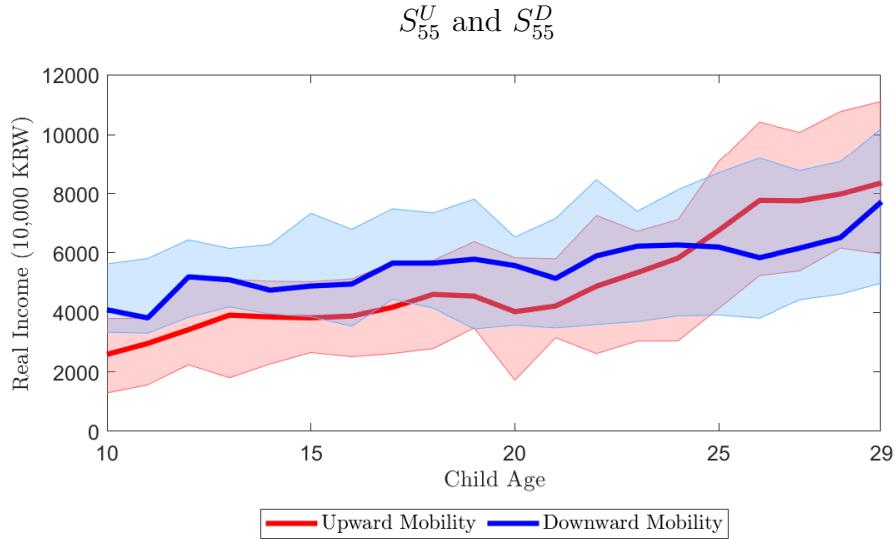


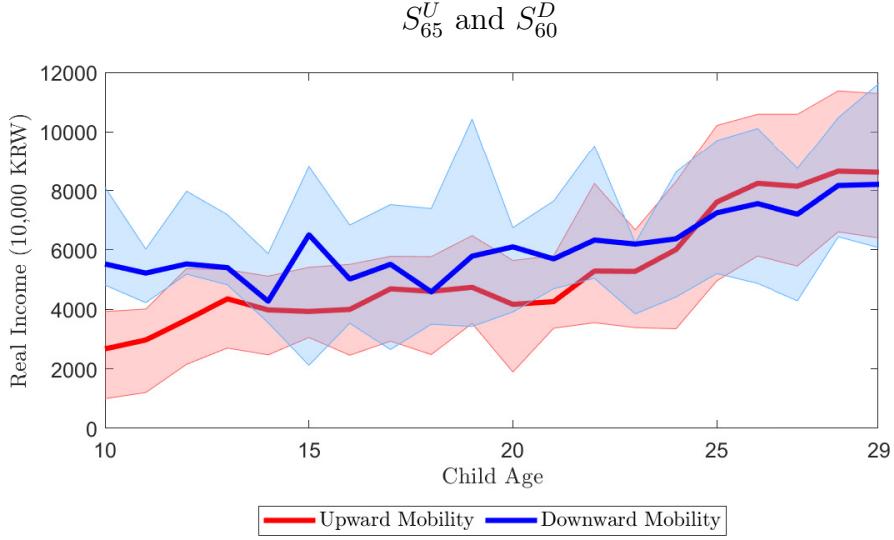
Notes: The left panel presents the LOOCV error. The minimum is attained at $m = 3$. The right panel shows the cumulative scree plot. This panel plots the total percentage of variation explained by the number of functional principal components. When $m = 3$, 79.81% of total variation is explained by the first three components.

E Robustness Check

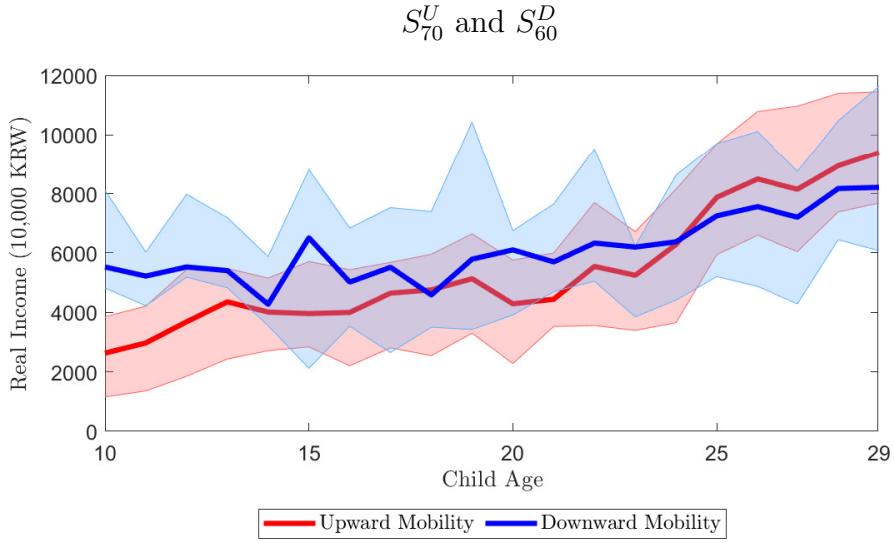
E.1 Upward and Downward Mobility

In all probability levels, we observe a definite upward trend for the upward mobility group. The red line plots the median income of the set S_p^U , and the shaded area indicates a central 68% range of S_p^U . Similarly, the blue line displays the median income of set S_p^D , and the shaded areas are for central 68% (from 16th and 84th percentile) of sets S_p^U and S_p^D .





We have only few observations for the set S^D if $p > 60$. Thus, we set $p = 60$ as the upper bound.



We have only few observations for the set S^D if $p > 60$. Thus, we set $p = 60$ as the upper bound.

E.2 Quadratic Effect of Parental Income Trajectory

Our approach is also open to the quadratic form of functional regression (Yao and Müller (2010)), such as

$$y_i = \alpha + \int_{10}^{29} \beta(r) f_i(r) dr + \int_{10}^{29} \int_{10}^{29} f_i(r) f_i(s) \Gamma(r, s) ds dr + \varepsilon_i \quad (32)$$

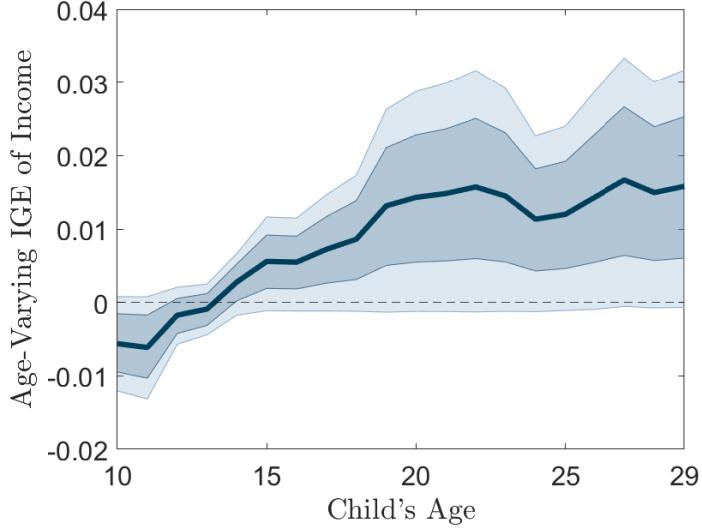
Our exercise here is important since the absence of quadratic effects is a challenge to literature because interactions across points in the trajectories are suggested by dynamic complementarities (Cunha and Heckman (2007) and Cunha et al. (2010)). We explicitly measure quadratic effects and evaluate the existence of evidence of dynamic complementarities.

Specifically, as parental log income trajectory $f_i(\cdot) > 0$ at all ages, a positive $\Gamma(r, s)$ implies the presence of dynamic complementarities between parental income at child's age r and s . To this end, we represent the estimated bivariate function (3D surface) in Figure 18. We also plot 20 numbers vertical slices of the surface at each child's age in Figure 19.

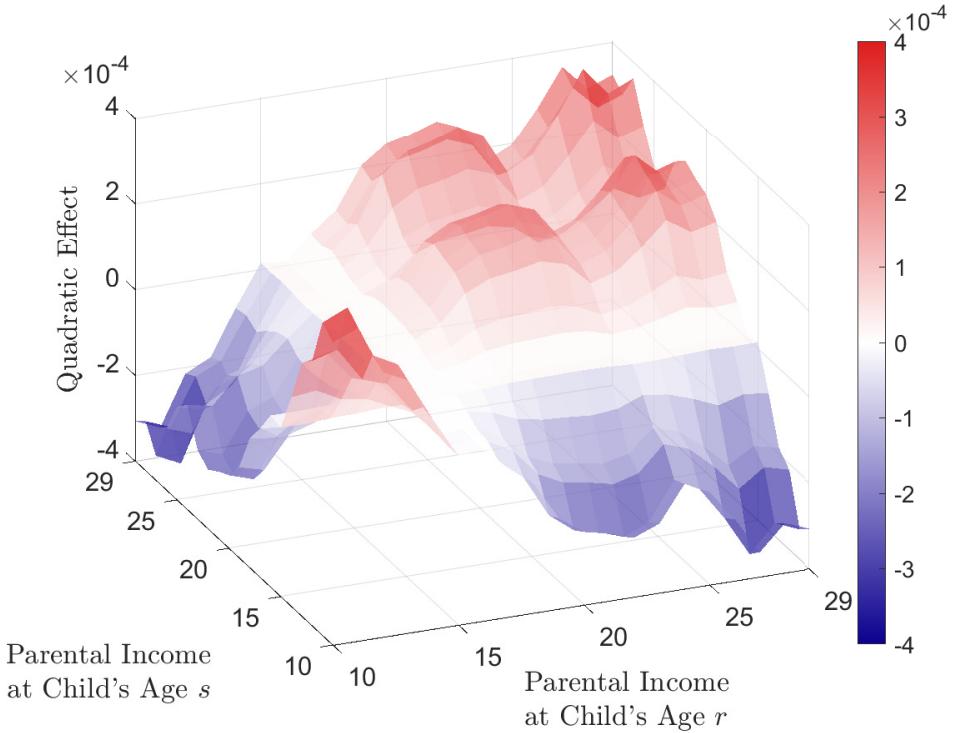
The upper panel of Figure 18 presents the estimated $\hat{\beta}$ in regression (32). When we compare this panel with Figure 2, the estimate under quadratic effect is statistically less significant. The lower panel of Figure 18 shows the estimated bivariate function $\hat{\Gamma}(r, s)$. As evident in Figure 19, the estimated function is not statistically significant at all r and s except early teens. During the early teens, $\hat{\Gamma}(r, s) > 0$, which implies the presence of dynamic complementarities in this period. However, the effects are not evident.

Figure 18: Quadratic Effect on Offspring's Income

Age-Varying IGE of Income

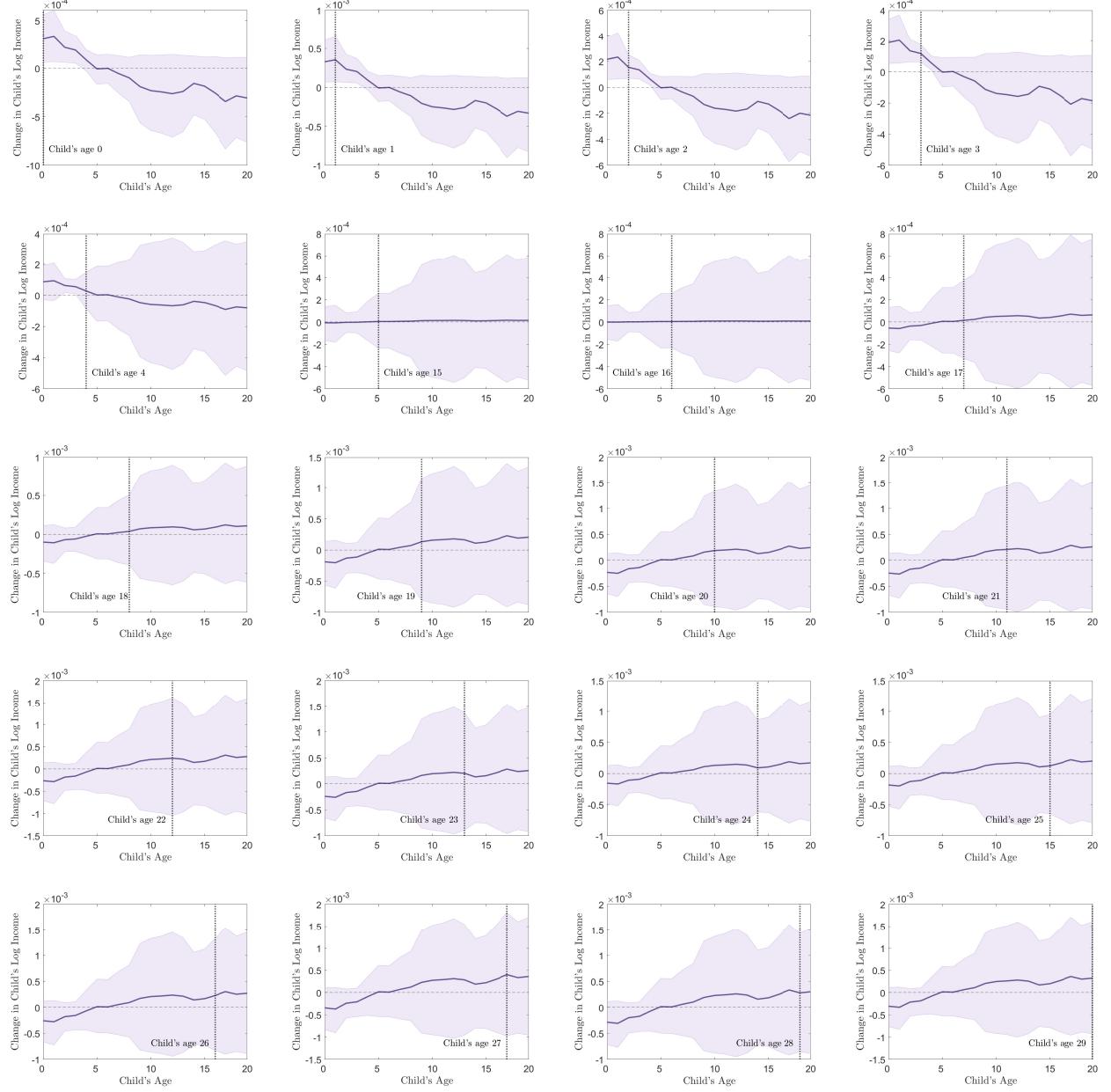


Quadratic Effect of Parental Income on Child's Income



Notes: The upper panel shows the estimated age-varying IGE of income in the model (32). The lower panel depicts the estimated bivariate coefficient function $\hat{\Gamma}(r, s)$ for $r, s \in [10, 29]$. The red-colored area indicates the value of the estimated coefficient function is positive, while the blue-colored one denotes the negative. The interpretation of $\Gamma(r, s)$ in the model (32) is, it quantifies the effect of covariation of parental income at age r and s on the child's permanent income.

Figure 19: Slices of Quadratic Coefficient Function, $\hat{\Gamma}(r, \cdot)$ for $r = 0, \dots, 19$



Notes: Each panel presents the slice of the estimated quadratic coefficient function $\hat{\Gamma}(r, s)$ at each child's age r for $r = 10, \dots, 29$. Each slice is a univariate function (a line) at each child's age r , $\hat{\Gamma}(r, \cdot)$. This figure displays 20 slices from $\hat{\Gamma}(10, \cdot)$ (upper-left) to $\hat{\Gamma}(29, \cdot)$ (lower-right). The vertical dashed line in each panel indicates the child's age r . The shaded area indicates 68% bootstrap confidence interval.

E.3 Interaction Effect between Parental Income and Family Structure Trajectories

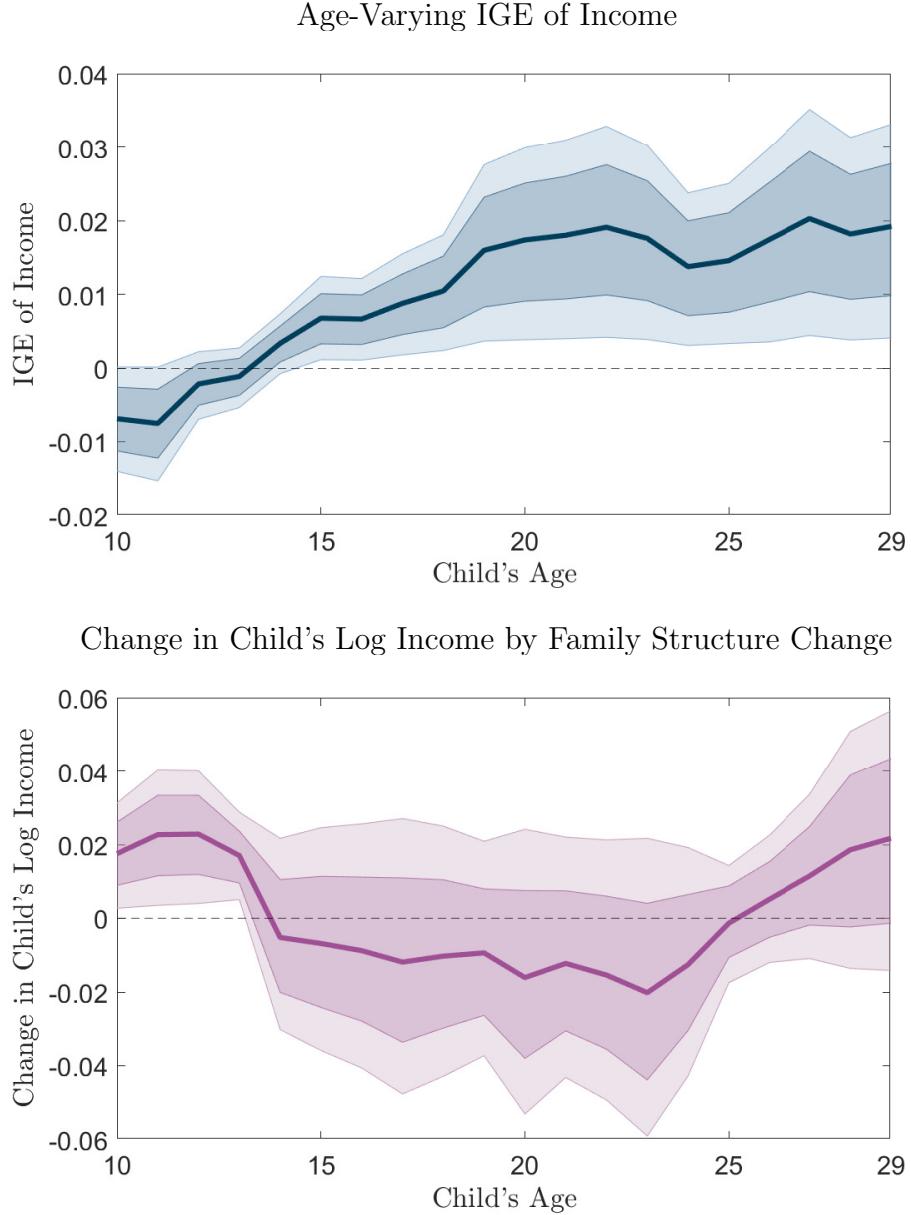
We also consider the following regression model to reflect the interaction effect ([Usset et al. \(2016\)](#)) between parental income and family structure trajectories.

$$y_i = \alpha + \int_{10}^{29} \beta(r) f_i(r) dr + \int_{10}^{29} \gamma(r) g_i(r) dr + \int_{10}^{29} \int_{10}^{29} f_i(r) g_i(s) \Delta(r, s) ds dr + \varepsilon_i \quad (33)$$

where (y_i) is the 3-year average log income of child i , (f_i) is a trajectory of i 's log parental income for 20 years, and (g_i) is a trajectory of family structure for 20 years. A bivariate function in the regression (33), $\Delta(r, s)$ is the coefficient of interaction between (f_i) and (g_i) . Note that, in contrast to $\Gamma(r, s)$ in model (32), $\Delta(r, s)$ is asymmetric bivariate function. If we take the view that family structure is a proxy of parents' time investment, $\Delta(r, s) < 0$ implies the evidence of dynamic complementarities between two different types of parental investment at child's age r and s .

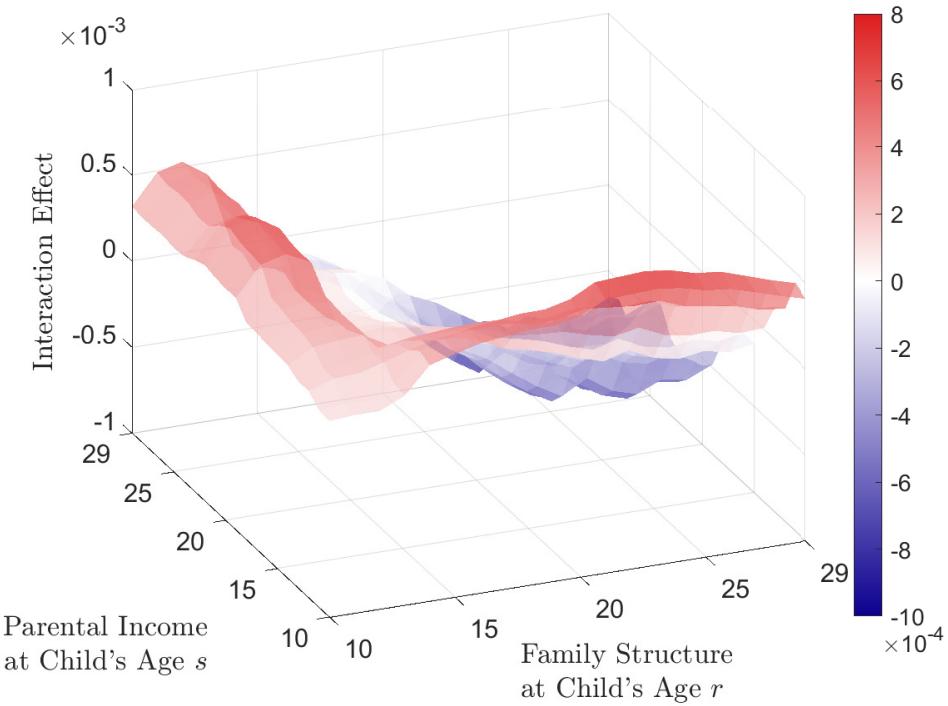
Figure 20 presents the estimated $\hat{\beta}$ and $\hat{\gamma}$ in regression (33). Even though interaction term is added to the regression model, these estimates are very similar to those without interaction term as in Figure 6. Figure 21 shows the estimated bivariate function $\Delta(r, s)$. This estimated function is not statistically significant at all child's ages, so in our analysis, we cannot find the evidence of dynamic complementarities between parental investments at different times. Each slice in Figure ?? and 23 presents the estimated bivariate function and 68% bootstrap confidence interval.

Figure 20: Parental Income and Family Structure Trajectories and Offspring's Income



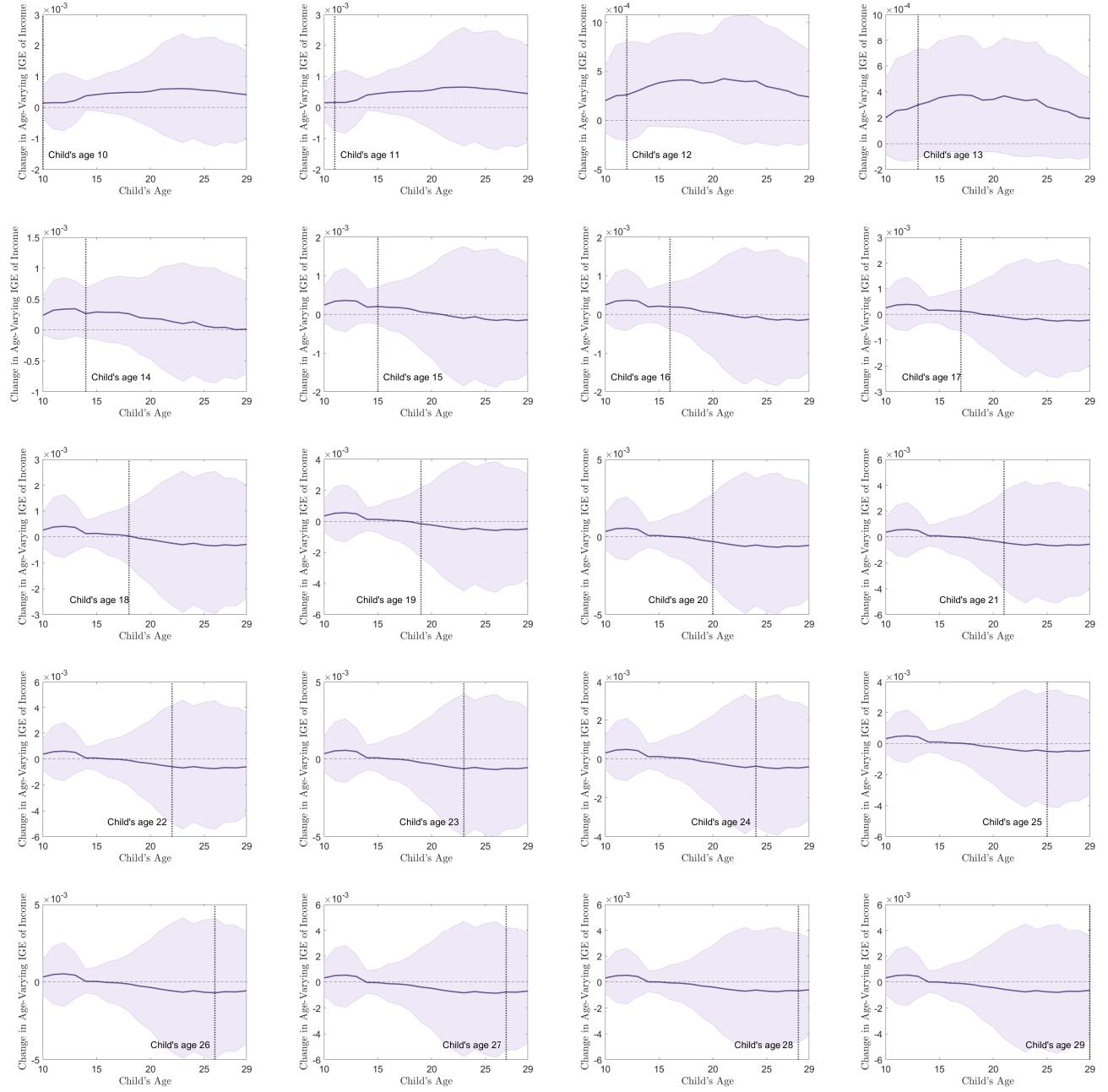
Notes: The panels in this figure present $\hat{\beta}(\cdot)$ and $\hat{\gamma}(\cdot)$ in (33), respectively. The upper panel presents the estimated $\hat{\beta}(\cdot)$, which measures age-varying IGE of income from a child's age 10 to 29. The lower panel plots $\hat{\gamma}(\cdot)$. This estimate measures the association between a trajectory of family structure (from a child's age 10 to 29) and a child's log income. The darker (lighter) shaded areas indicate 68% (90%) bootstrap confidence intervals.

Figure 21: Interaction Effect on Offspring's Income



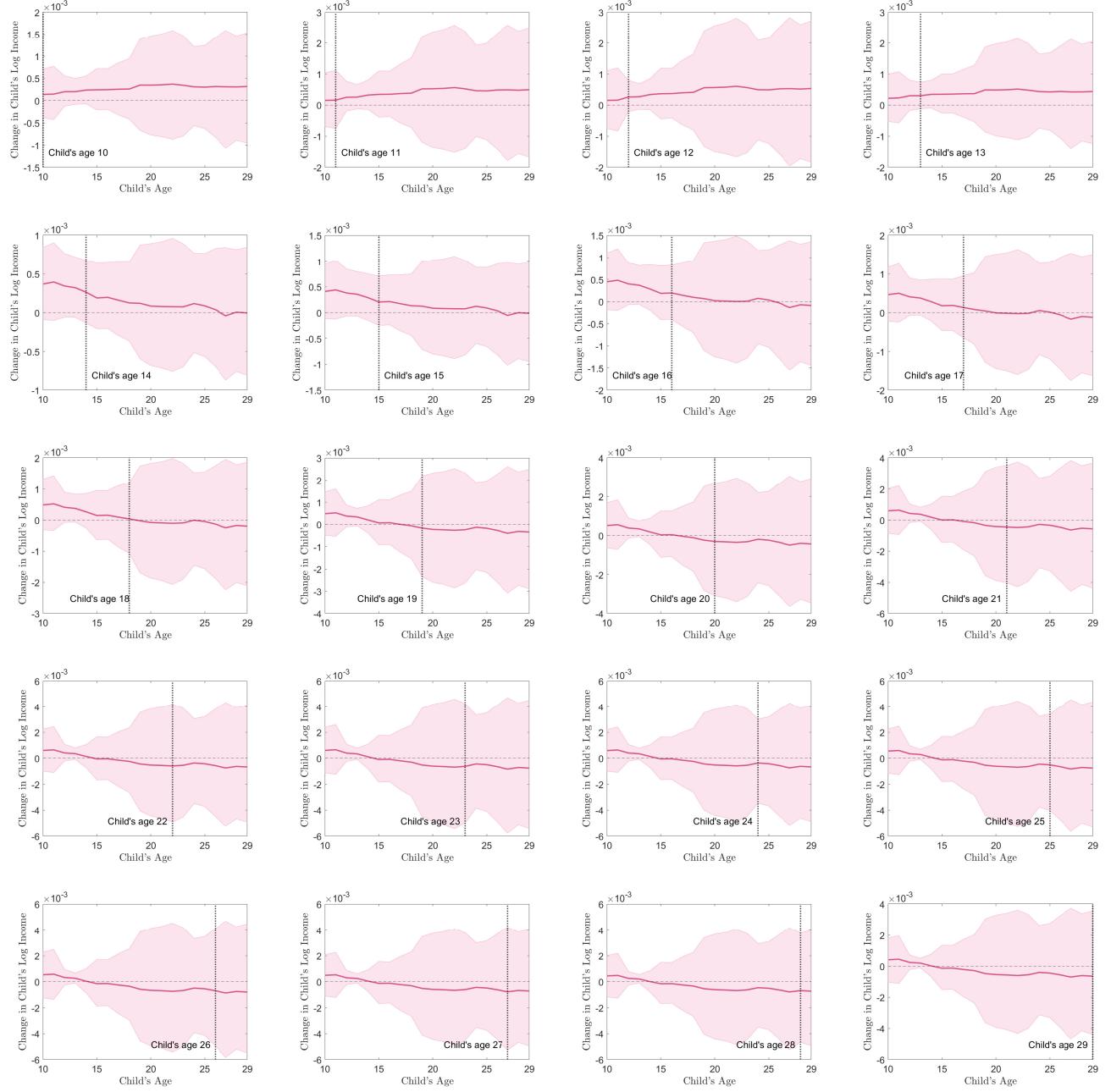
Notes: This figure presents the estimated interaction effect on a child's income, which is equivalent to the effect of interaction between parental income and family structure trajectories on child's permanent income.

Figure 22: Slices of Interaction Coefficient Function, $\hat{\Delta}(r, \cdot)$ for $r = 10, \dots, 29$



Notes: Each panel presents the slice of the estimated quadratic coefficient function $\hat{\Delta}(r, s)$ at each child's age r for $r = 10, \dots, 29$. Each slice is a univariate function (a line) at each child's age r , $\hat{\Delta}(r, \cdot)$. This figure displays 20 slices from $\hat{\Delta}(10, \cdot)$ (upper-left) to $\hat{\Delta}(29, \cdot)$ (lower-right). The vertical dashed line in each panel indicates the child's age r . The shaded area indicates 68% bootstrap confidence interval.

Figure 23: Slices of Interaction Coefficient Function, $\hat{\Delta}(\cdot, s)$ for $s = 10, \dots, 29$

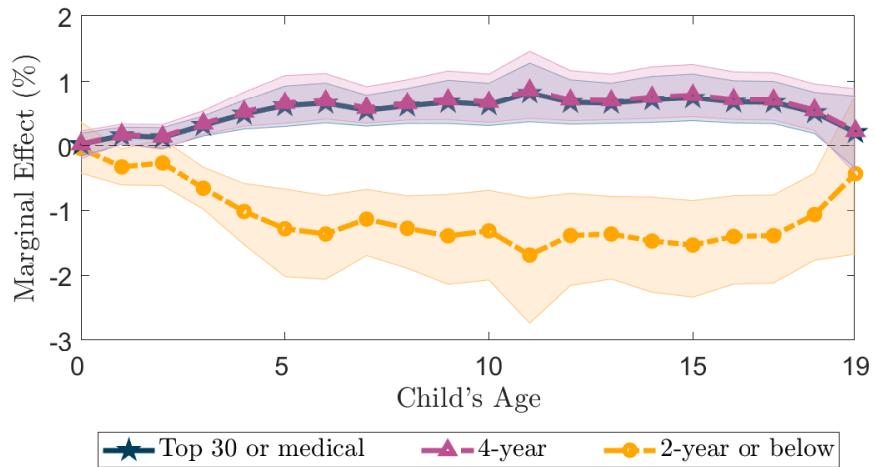


Notes: Each panel presents the slice of the estimated quadratic coefficient function $\hat{\Delta}(r, s)$ at each child's age s for $s = 10, \dots, 29$. Each slice is a univariate function (a line) at each child's age r , $\hat{\Delta}(\cdot, s)$. This figure displays 20 slices from $\hat{\Delta}(\cdot, 10)$ (upper-left) to $\hat{\Delta}(\cdot, 29)$ (lower-right). The vertical dashed line in each panel indicates the child's age s . The shaded area indicates 68% bootstrap confidence interval.

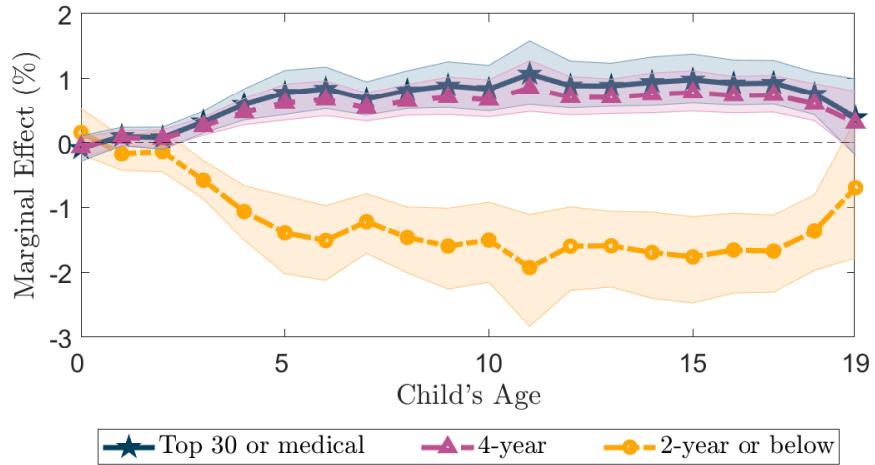
E.4 Educational Attainment

As a robustness check of our result, we compare the marginal effect of parental income trajectory on their children's educational attainment under various models, ordered logit, ordered probit, and multinomial logit. However, we do not consider the multinomial probit model since there is no closed-form solution. For this reason, we compare the results from the three models.

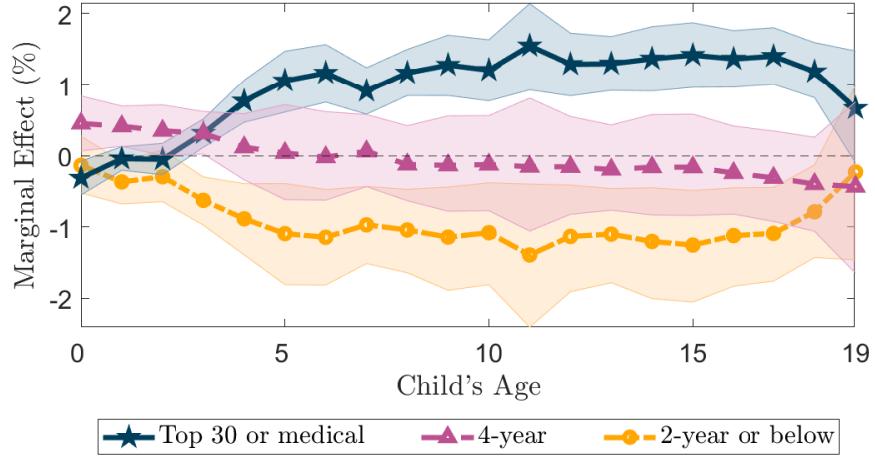
1. Ordered Logit (Figure 9)



2. Ordered Probit



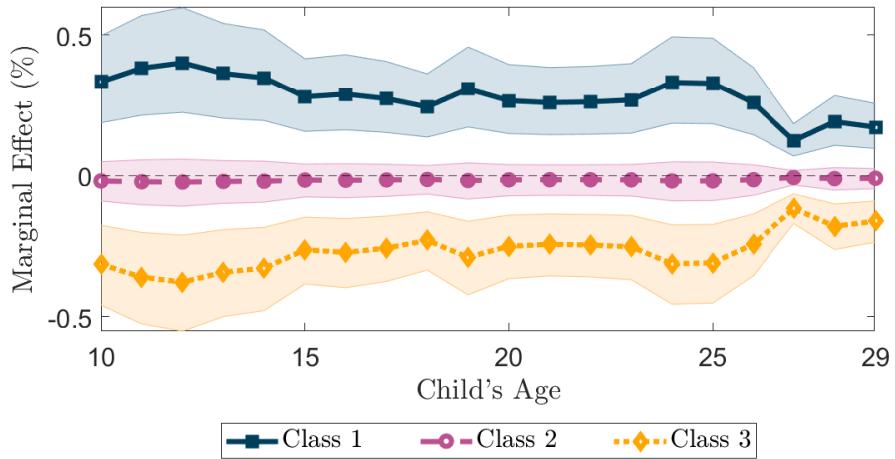
3. Multinomial Logit



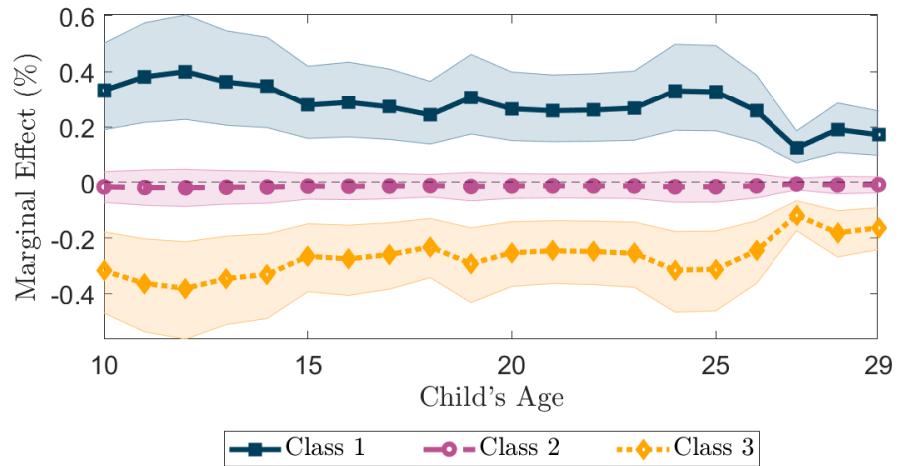
E.5 Occupational Choice

As a robustness check of our result, we compare the marginal effect of parental income trajectory on their children's occupational choice under various models, ordered logit, ordered probit, and multinomial logit. However, we do not consider the multinomial probit model since there is no closed-form solution. For this reason, we compare the results from the three models.

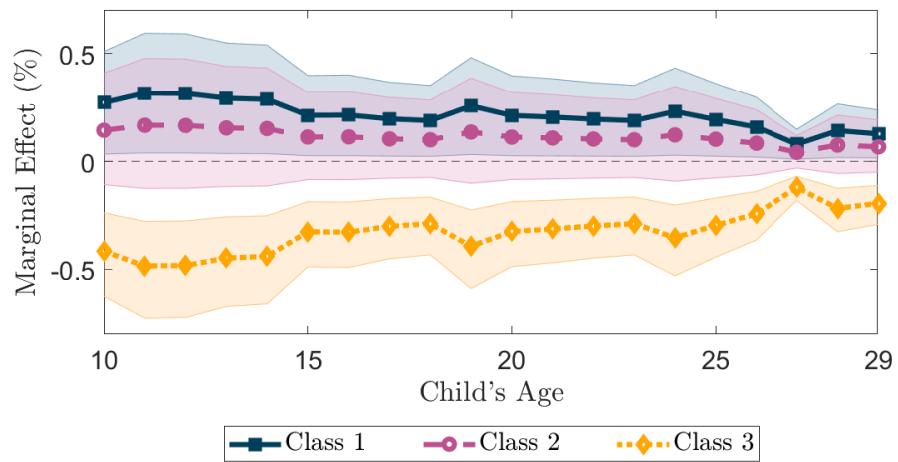
1. Ordered Logit



2. Ordered Probit



3. Multinomial Logit (Figure 24)



References

- [1] Becker, G. S. and N. Tomes (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy* 87(6), 1153–1189.
- [2] Biblarz, T. J., A. E. Raftery, and A. Bucur (1997). Family structure and social mobility. *Social Forces* 75(4), 1319–1341.
- [3] Bloome, D. (2017). Childhood family structure and intergenerational income mobility in the united states. *Demography* 54(2), 541–569.
- [4] Bosq, D. (2000). *Linear Processes in Function Spaces*. Springer-Verlag.
- [5] Carneiro, P., I. L. García, K. G. Salvanes, and E. Tominey (2021). Intergenerational mobility and the timing of parental income. *Journal of Political Economy* 129(3), 757–788.
- [6] Chang, Y., M. Choi, and J. Y. Park (2021). A factor model for functional time series. Working Paper, Indiana University.
- [7] Chang, Y., J. Y. Park, and D. Pyun (2021). From functional autoregressions to vector autoregressions. Working Paper, Indiana University.
- [8] Chaussé, P., T. Chen, and K. A. Couch (2021). Assessing inter-generational mobility: A functional data regression approach. Working Paper.
- [9] Chetty, R., N. Hendren, P. Kline, and E. Saez (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *Quarterly Journal of Economics* 129(4), 1553–1623.
- [10] Cholli, N. and S. N. Durlauf (2021). Understanding the heterogeneity of intergenerational mobility across neighborhoods. Working Paper, University of Chicago.
- [11] Chun, S.-H. (2010). Strategy for industrial development and growth of major industries in korea. *Energy & Mineral Resources Development Association of Korea*.
- [12] Cunha, F. and J. Heckman (2007). The technology of skill formation. *American Economic Review* 97(2), 31–47.
- [13] Cunha, F., J. J. Heckman, and S. M. Schennach (2010). Estimating the technology of cognitive and noncognitive skill formation. *Econometrica* 78(3), 883–931.

- [14] Duncan, G. J. and W. Rodgers (1991). Has children's poverty become more persistent? *American Sociological Review*, 538–550.
- [15] Durlauf, S. N. (1996). A theory of persistent income inequality. *Journal of Economic growth* 1(1), 75–93.
- [16] Durlauf, S. N., A. Konstantinidi, and A. Kourtellos (2021). Intergenerational trajectories. Working Paper, University of Chicago.
- [17] Emran, M. S. and F. Shilpi (2011). Intergenerational occupational mobility in rural economy evidence from Nepal and Vietnam. *Journal of Human Resources* 46(2), 427–458.
- [18] Heckman, J. J. and S. Mosso (2014). The economics of human development and social mobility. *Annual Review of Economics* 6(1), 689–733.
- [19] Jäntti, M. and S. P. Jenkins (2015). Income mobility. In *Handbook of income distribution*, Volume 2, pp. 807–935. Elsevier.
- [20] Kim, H. (2009). Analysis on intergenerational economic mobility in korea. *KDI Policy Study*.
- [21] Lee, K. and I. Min (2016). Study on intergenerational income and occupational mobility. *Korea Labor Institute*.
- [22] Long, J. and J. Ferrie (2013). Intergenerational occupational mobility in great britain and the united states since 1850. *American Economic Review* 103(4), 1109–37.
- [23] Loury, G. C. (1981). Intergenerational transfers and the distribution of earnings. *Econometrica*, 843–867.
- [24] Musick, K. and R. D. Mare (2004). Family structure, intergenerational mobility, and the reproduction of poverty: Evidence for increasing polarization? *Demography* 41(4), 629–648.
- [25] Park, J. Y. and J. Qian (2012). Functional regression of continuous state distributions. *Journal of Econometrics* 167, 397–412.
- [26] Phang, H. and K.-H. Kim (2003). Trends and differentials in educational stratification in korea. *Korean Journal of Sociology* 37(4), 31–65.
- [27] Ramsay, J. O. and B. W. Silverman (2005). *Functional Data Analysis*. Springer.

- [28] Ramsay, J. O. and B. W. Silverman (2007). Applied Functional Data Analysis: Methods and Case Studies. Springer.
- [29] Shin, Y.-J., M.-J. Lee, and S. Park (2019). Analysis and outlook of korea's long-term trends in lowest-low fertility. KIHASA Research Monographs.
- [30] Solon, G. (1999). Intergenerational mobility in the labor market. In Handbook of Labor Economics, Volume 3, pp. 1761–1800. Elsevier.
- [31] Solon, G. (2002). Cross-country differences in intergenerational earnings mobility. Journal of Economic Perspectives 16(3), 59–66.
- [32] Solon, G. (2004). A model of intergenerational mobility variation over time and place. Generational Income Mobility in North America and Europe 2, 38–47.
- [33] Usset, J., A.-M. Staicu, and A. Maity (2016). Interaction models for functional regression. Computational Statistics & Data Analysis 94, 317–329.
- [34] Yao, F. and H.-G. Müller (2010). Functional quadratic regression. Biometrika 97(1), 49–64.