The Effects of Parental Income and Family Structure on Intergenerational Mobility: A Trajectories-Based Approach*

Yoosoon Chang† Steven N. Durlauf‡ Bo Hu§ Joon Y. Park¶

July 2025

Abstract

We examine how parental income and family structure during childhood and adolescence affect adult income, emphasizing the timing of these effects. Using an ordered multinomial probability model with functional covariates, we find that these familial influences are strongest in middle childhood and adolescence. We also uncover a complementary relationship in the effects of income and family structure trajectories during key developmental periods. By flexibly controlling for personal and family characteristics using nonparametric methods, our approach effectively handles high-dimensional covariates. The results advance the understanding of intergenerational income mobility and highlight the long-term importance of early-life familial conditions for adult economic success.

JEL Classification: C10, C14, C25, C50, D10, J12, J62

Keywords and phrases: intergenerational mobility, income trajectory, family structure trajectory, complementarity between family income and family structure, ordered multinomial probability model, reproducing kernel Hilbert space

*We thank the James M. and Cathleen D. Stone Foundation for support, and Renato De Angelis and Weiqi Wang for helpful comments. We are also grateful to Simon Sokbae Lee and Guo Yan for helpful discussions.

†Indiana University, yoosoon@iu.edu

‡University of Chicago, sdurlauf@uchicago.edu

§Indiana University, hu21@iu.edu

¶Indiana University, joon@iu.edu

1 Introduction

One of the most active areas of current inequality research concerns intergenerational mobility. This paper contributes to this literature by integrating the study of the effects of family income and the effects of family structure on income. Specifically, this paper is designed to extend two distinct sets of analyses.

First, we build on recent studies that demonstrate how aggregate measures of intergenerational elasticity (IGE) may obscure a critical role for the timing of parental income. See Muller (2010), Duncan et al. (2010), Cheng and Song (2019), and Chang et al. (2024) for a range of findings. The sensitivity of child outcomes to the timing of parental income can be expected given factors ranging from credit constraints (Grawe, 2004) to periods of sensitive dependence of childhood skill formation (Heckman and Mosso, 2014). Chang et al. (2024) provide the most general evidence that income timing matters by tracing out how distinct family income effects occur at each age of childhood and adolescence. We follow this approach by exploring how different family income trajectories predict adult outcomes, doing so in a way that considers the evolution of income classes rather than income levels, which are the focus of most of the literature.

Second, we build on analyses that focus on the distinct role of family structure on child development and adult outcomes. Family effects have long been documented: important examples include Björklund and Sundström (2006), Francesconi et al. (2010), Moschion and van Ours (2019), and Frimmel et al. (2024). Joint consideration of family income and family structure is less common, but some noteworthy advances have been made. Bloome (2017) finds that children in single parent households are more likely to exhibit downward mobility. Ermisch et al. (2004) find that growing up in a single-parent family arguably has a stronger negative impact on young adult outcomes than experiencing parental joblessness, with early childhood family disruption (ages 0–5) being particularly influential. These types of findings have evident explanations. Parental time augments income in defining family resources. Moreover, family structure impacts a child's stability, emotional well-being, and mental health. In parallel to our analysis of income, we consider how trajectories of family structure predict adult outcomes.

Our analysis builds on this prior literature by developing a statistical framework to simultaneously analyze the effects of family income and family structure on children. We do this by generalizing a framework originating in Chang et al. (2025) that models the evolution of low, medium, and high income classes across generations. This formulation allows us to interpret intergenerational mobility via a multinomial probability model that generalizes classical Markov chain formulations of intergenerational mobility dynamics. The Markov chain structure allows us to analyze very general specifications of the effects of family income and family structure trajectories on the probabilities of adult income classes. Methodologically, this paper develops an ordered multinomial probability model that incorporates functional covariates. This model means that for each family there is a unique Markov transition matrix that maps income classes of parents to children. The probabilities of children attaining different adult income classes are linked to a range of individual, familial, and environmental factors, with particular attention to income and family structure trajectories which are modeled as functional covariates. The model also accommodates additional covariates in a flexible, nonparametric manner. Departing from conventional probit or logit approaches, we relax parametric assumptions on the distribution of unobserved heterogeneity by adopting a fully nonparametric specification. Drawing on tools from functional data analysis, nonparametric estimation, and machine learning, we propose a maximum likelihood estimation procedure for the model parameters and discuss its theoretical properties.

One key feature of our analysis is the exploration of interactions between family economic conditions and family structure in determining children's outcomes. Family income and family structure may function as either substitutes or complements; there are no strong a priori theoretical grounds for one or the other to hold. Further, the nature of this interaction may vary across different stages of child development; there is no reason why uniform complementarity or substitutability needs to hold. A major methodological innovation of this paper is the development of a way to model trajectory interactions that allows them to vary across the life course.

We apply these methods to the Panel Study of Income Dynamics (PSID). We show that parental income timing plays a critical role in future adult outcomes. While parental income and two-parent family status positively influence a child's adult income at all stages of childhood, the effects are strongest during the middle childhood and adolescent years. These adolescence findings align with existing literature such as Carneiro et al. (2021), Duncan et al. (2010), and Chang et al. (2024). With reference to early childhood, our results confirm the findings in Chang et al. (2024) that early childhood is relatively less important than later years. This finding is different from other studies; we discuss the reasons why our results are credible below.

Our model also allows us to empirically investigate whether family economic conditions and family structure influence future offspring outcomes as complements or substitutes. Our findings suggest that, in general, these factors always act as complements. Notably, the complementary effect is most pronounced between parental income during middle to late childhood and family structure during late childhood. Further, complementarity holds between family structure and family income measured at different ages. These results are novel in the literature, offering new insights into the interplay between financial resources and family structure dynamics.[1] These results underscore the importance of coordinated policies and interventions that address both economic and familial stability, particularly

[1]Of course, there are findings in many papers that can speak to the interactions between family structure and family income. For example, Carneiro et al. (2021) find that income effects are smaller when family structure is added as a control. Our approach evaluates interactions systematically and with minimal structure on the nature of the interactions.

during critical developmental periods, to maximize the likelihood of positive outcomes for children.

2 Income Classes vs Income Levels

Our analysis of family income and family structure trajectories focuses on the way they affect the likelihood of a child being a member of one of three income classes—low, middle, or high—in adulthood. We consider income categories instead of the income levels of individuals as our dependent variable for the following reasons.

First, particular income categories such as middle class are of substantive interest to the general public and policymakers. While our definition may not correspond to the definition used by the general public, it mirrors the popular concern about relative status.

Second, our generalized Markov chain approach allows for very flexible modeling of the effects of trajectories on adult outcomes.

Third, publicly available income data usually are left and/or right censored.[2] While bottom and/or top coding could significantly distort the absolute level of income of certain groups of people, it does not change the income class classifications of individuals, if income categories are reasonably designed.

Fourth, this strategy helps deal with zeros in income series. It is conventional in mobility analyses to take logarithms. Standard approaches such as "adding one before taking logs" can lead to significant bias (Chen and Roth, 2024). The use of categories eliminates this issue.

Fifth, in the study of intergenerational mobility, it is customary to link the permanent income of parents with that of their children. However, since permanent income is not

[2]The inconsistent censoring threshold in PSID further complicates the issue. For example, the household income in PSID is top-coded at 99,999 from 1968 to 1979, at 999,999 in 1980, increased up to 9,999,999 from 1981 to 1983, back to 999,999 in 1984 and 1985, increased to 9,999,999 again from 1986 to 1996, and back to 999,999 from 1997 on.

directly observable, it has to be proxied. One common proxy for permanent income is the average income over an extended period, ideally a lifetime. Although the PSID panel that we use in this paper spans multiple decades and includes families tracked across multiple generations, it does not cover a large enough number of parent–child pairs with complete income histories to reliably estimate lifetime income for both generations. Consequently, for some of the younger generation, the available income data is limited to the early years, making a lifetime average infeasible and requiring the use of early adulthood income as a potentially imprecise proxy.

To evaluate the suitability of early adulthood income as a proxy for permanent income levels or income categories, we restrict the sample to individuals with income data spanning ages 31 to 60, and assess whether average income in their thirties reliably represents income in later life stages and over the entire observed period. Specifically, we compare early and later life income levels by calculating the percentage difference between average income in a later stage and that in the early stage. In addition, to examine categorical stability, we compute the proportion of individuals who remain in the same income category across stages. We define three income categories—low, middle, and high—based on relative position within the distribution: low income is below two-thirds of the median, high income exceeds twice the median, and middle income lies in between. These categories are constructed relative to either the full sample or specific birth cohorts, as discussed below.

Table 1 presents the proportions of individuals whose late life stage income levels differ from their 30s income by less than 20% or 40%, as well as the proportion who remain in the same income category (low, middle, or high) when classification is based on income in their 30s versus later life stages. Across all cohorts, only slightly more than one-third of individuals have an average income in their forties that differs from their thirties average by less than 20%, and less than two-thirds differ by less than 40%—indicating substantial variation across life stages. When comparing individuals' incomes in their thirties to those

Table 1: Income Level vs Income Class

	40s vs 30s income			50s vs 30s income			life vs 30s income		
	<20%	<40%	same class	<20%	<40%	same class	<20%	<40%	same class
all cohorts	0.359	0.628	0.782	0.284	0.506	0.676	0.520	0.796	0.833
1936-1945	0.392	0.669	0.788	0.332	0.584	0.723	0.562	0.863	0.852
1946-1955	0.352	0.622	0.786	0.284	0.508	0.685	0.546	0.804	0.852
1956-1965	0.336	0.583	0.770	0.255	0.456	0.645	0.501	0.773	0.828
1966-1975	0.398	0.680	0.753	0.272	0.505	0.702	0.593	0.850	0.842

Notes: This table compares individuals' average income in their 40s, 50s, and over their lifetime to their average income in their 30s. It reports the proportions (ranging from 0 to 1) of individuals whose income levels differ from their 30s income by less than 20% or 40%, as well as the proportion who remain in the same income category (low, middle, or high) when classification is based on income in their 30s versus in their 40s, 50s, and their lifetime average.

incomes in their fifties, the proportion of individuals whose average income differs by less than 20% and by less than 40% drops further to 28.4% and 50.6%, respectively. Additionally, for half of the sample, average lifetime income differs from income in their twenties by more than 20%, and for 20% of individuals, lifetime income differs from income in their thirties by more than 40%.

In contrast, 83.3% of individuals are assigned to the same income category when comparing classifications based on average income in their thirties versus their lifetime average. This suggests that using income categories as a proxy for permanent income is nearly as robust as tolerating a 40% error margin in income level. The patterns are broadly consistent across different cohorts. These findings highlight that modeling income in categorical terms provides a more stable and reliable approach when dealing with incomplete income histories. Motivated by this insight, we proceed to develop a nonparametric multinomial model that incorporates income trajectories along with other covariates that may influence an individual's long-term income status.

3 Model Specification and Estimation

3.1 Basic Setup

Denote by $i = 1, \ldots, n$ and $j = 1, 2, \ldots, m$ individuals and their income categories, respectively. The conditional probability of individual i being in category j is determined by a latent variable y_i^* representing permanent income of a child through

$$\pi_j(f, g, x) = \mathbb{P}\left\{\mu_{j-1} < y_i^* \leqslant \mu_j \,\middle|\, f_i = f, g_i = g, x_i = x\right\}, \tag{1}$$

where $(\mu_j)_{j=0}^m$ is a set of thresholds defined as $-\infty = \mu_0 < \mu_1 < \cdots < \mu_{m-1} < \mu_m = \infty$ with an $(m-1)$-dimensional parameter $\mu = (\mu_1, \ldots, \mu_{m-1})'$, $(f_i)_{i=1}^n$ and $(g_i)_{i=1}^n$ are functional covariates measuring the trajectories of parental income and family structure, respectively, and $(x_i)_{i=1}^n$ is an ℓ-dimensional vector of covariates including gender, race, parental education, health conditions and environments.

We let the latent variable y_i^* be generated from these covariates as

$$y_i^* = \int_a^b \alpha(t) f_i(t) dt + \int_a^b \beta(t) g_i(t) dt + \int_a^b \int_a^b \Gamma(t,s) f_i(t) g_i(s) dt ds + \tau(x_i) + \varepsilon_i,$$

where $\alpha(\cdot)$ and $\beta(\cdot)$ are functional coefficients on f_i and g_i respectively, $\Gamma(\cdot,\cdot)$ is a function from $\mathbb{R} \times \mathbb{R}$ to $\mathbb{R}$, a and b are two real numbers indicating age range, $\tau(\cdot)$ is a function from $\mathbb{R}^\ell$ to $\mathbb{R}$, and ε_i is a random error with unknown density function ρ. Subsequently, we view (f_i) and (g_i) as elements in the Hilbert space $\mathcal{H} = L^2([a,b))$ of square-integrable real functions on $[a,b)$ with inner product $\langle u, v \rangle = \int_a^b u(t)v(t)\,dt$ for $u, v \in \mathcal{H}$ and squared norm $\|v\|^2 = \int_a^b v(t)^2 dt$ for $v \in \mathcal{H}$. Furthermore, we interpret Γ as a linear operator on $\mathcal{H}$, which maps v to Γv with Γv given by $(\Gamma v)(t) = \int_a^b \Gamma(t,s) v(s)\,ds$ for $v \in \mathcal{H}$. Throughout the paper, we will assume that Γ is a compact operator and simply write $\Gamma \in L(\mathcal{H})$.[3] Under this setting, we may rewrite our model more compactly as

$$y_i^* = \langle \alpha, f_i \rangle + \langle \beta, g_i \rangle + \langle f_i, \Gamma g_i \rangle + \tau(x_i) + \varepsilon_i. \tag{2}$$

The model introduced here is quite flexible in representing the effects of our functional covariates f_i and g_i on the latent variable y_i^*, while allowing for arbitrary nonlinear effects of other covariates (x_i) included as controls. In particular, we allow for complementarity or substitutability of f_i and g_i, as well as their individual linear effects, on y_i^*.[4] Throughout the paper, we assume that $(f_i, g_i, x_i, \varepsilon_i)$ are independent and identically distributed for $i = 1, \ldots, n$.

Note that the functional coefficients $\alpha(t), \beta(t)$, and the bivariate function $\Gamma(t,s)$, which we call the interaction function, can be approximated by

$$\alpha(t) \approx \frac{1}{\delta} \int_{t-\frac{\delta}{2}}^{t+\frac{\delta}{2}} \alpha(z) dz, \qquad \beta(t) \approx \frac{1}{\delta} \int_{t-\frac{\delta}{2}}^{t+\frac{\delta}{2}} \beta(z) dz$$

and

$$\Gamma(t,s) \approx \frac{1}{\delta^2} \int_{t-\frac{\delta}{2}}^{t+\frac{\delta}{2}} \int_{s-\frac{\delta}{2}}^{s+\frac{\delta}{2}} \Gamma(z,w)\, dz dw$$

when δ is small. Using standard terminology in the regression with interaction terms, we may call $\alpha(t)$ and $\beta(t)$ the main effect of parental income and family structure at age t, and $\Gamma(t,s)$ the interaction effect of parental income at age t and family structure at age s. In this

[3]For notational brevity, we use the same symbol, Γ, to denote both the binary function and the operator; respective meanings will be clear from the context. Intuitively, a compact operator may be thought of as an infinite dimensional matrix and shares many common features with finite dimensional matrices. In functional data analysis, any parameter given as an operator is routinely assumed to be compact. It is well known that the operator Γ induced by a kernel function $\Gamma(\cdot,\cdot)$ becomes compact if for instance $\int_a^b \int_a^b \Gamma(t,s)^2 dtds < \infty$ (Conway, 2000, p. 267).

[4]The Riesz representation theorem states that any linear mapping λ from $\mathcal{H}$ to $\mathbb{R}$ can be represented uniquely as an inner product with a function u in $\mathcal{H}$, i.e., $\lambda(v) \equiv \langle u, v \rangle$ for all $v \in \mathcal{H}$ with some $u \in \mathcal{H}$ fixed.

study, we shall focus on the child's parental income and family structure from birth to age 20. We therefore consider (f_i) and (g_i) as functions on $[0, 20)$ by setting $a = 0$ and $b = 20$.

We let the function $\tau(\cdot)$ and the density $\rho(\cdot)$, which are introduced to control the effects of other covariates x_i on y_i^* and represent the distribution of an error term, respectively, be given nonparametrically without imposing any restrictive parametric form. In what follows, we assume that $\tau \in \mathcal{H}_K$, where $\mathcal{H}_K$ is defined as the reproducing kernel Hilbert space (RKHS) generated by a kernel function $K : \mathbb{R}^\ell \times \mathbb{R}^\ell \to \mathbb{R}$. This is to estimate τ effectively while accommodating the presence of functional terms and the full set of our controls in the model.[5] Moreover, we assume that ρ belongs to a class $\mathcal{P}$ of density functions that can be expressed using Hermite polynomials. This assumption is made so that we may use the approach developed by Gallant and Nychka (1987). Although this class is not fully nonparametric, it is large enough to encompass many density functions with a variety of different characteristics. The details of our specifications for τ and ρ will be provided later in Sections 3.4 and 3.5, respectively.[6]

3.2 Identification

In order to understad the conditions for identifiability of our model, let $\theta = (\mu, \alpha, \beta, \Gamma, \tau, \rho)$ be given by the set of unknown parameters and note that our specification of the latent variable y_i^* representing permanent income of a child in (2) yields the conditional income class probabilities $\pi_j(f, g, x; \theta)$ in (1) as a function of θ for each $j = 1, \ldots, m$. By identification of $\theta \in \Theta$ for some parameter space Θ, we mean that if θ_0 is the true value of θ and there exists $\theta \in \Theta$ such that $\pi_j(f_i, g_i, x_i; \theta) = \pi_j(f_i, g_i, x_i; \theta_0)$ a.s. for all $j = 1, \ldots, m$, then $\theta = \theta_0$. We

[5] To estimate our model, we use a machine learning technique called the *kernel trick*, which is built up on a reproducing kernel Hilbert space (RKHS). The kernel trick allows us to accommodate the presence of general additional terms as well as a large-dimensional set of covariates.

[6] Our specifications of τ and ρ are based on Yan (2024), who considers the binary choice model with an additional simple linear term for the covariate of interest, in place of our linear and quadratic terms involving functional covariates.

define $\Theta = \mathbb{R}^{m-1} \times \mathcal{H} \times \mathcal{H} \times L(\mathcal{H}) \times \mathcal{T} \times \mathcal{P}$ and let $\alpha \in \mathcal{H}, \beta \in \mathcal{H}, \Gamma \in L(\mathcal{H})$, $\tau \in \mathcal{T}$ and $\rho \in \mathcal{P}$.

For identification of the parameter θ in our model, it is sufficient to assume that (a) (ε_i) is independent of (f_i, g_i, x_i), (b) (f_i, g_i, x_i) has support $\mathcal{H} \times \mathcal{H} \times \mathcal{X}$, (c) $\mathcal{T}$ is a set of continuous functions with zero infimum on $\mathcal{X}$, and (d) $\mathcal{P}$ is a set of strictly positive probability density functions on $\mathbb{R}$ representing distributions with zero mean and unit variance.

Our assumptions for identification are directly comparable to those employed in two earlier papers by Matzkin (1992) and Yan (2024), which consider identification of the binary choice model given by $y_i^* = \lambda w_i + \tau(x_i) + \varepsilon_i$, where w_i is a covariate of interest added separably from x_i as a linear term. The former allows τ to be any continuous function, while the latter assumes that τ belongs to a RKHS. Of the two, the latter is more directly related to our paper. The proof of identification of the parameter θ in our model is rather straightforward given their results. To save the space, it is therefore provided in an online appendix to our paper.

The condition in (a) is standard, although it is not necessary. The conditions in (b), (c) and (d) ensure that the parameters (α, β, Γ), τ and ρ are identified, respectively. In particular, it follows from (b) that if the functional component $\langle \alpha, f_i \rangle + \langle \beta, g_i \rangle + \langle f_i, \Gamma g_i \rangle$ is identified in our model, the parameters α, β and Γ are individually identified. Moreover, $\inf_{x \in \mathcal{X}} \tau(x) = 0$ sets the level of τ, identifying all $(m-1)$ threshold parameters $(\mu_j)_{j=1}^{m-1}$. Without this condition, we need to fix one of these parameters and may identify only $(m-2)$ threshold parameters. The restriction we impose on the mean and variance of the distribution given by ρ identify the location and scale of the systematic component $\langle \alpha, f_i \rangle + \langle \beta, g_i \rangle + \langle f_i, \Gamma g_i \rangle + \tau(x_i)$ of our model.

3.3 Main Part

The main part of our model represents the effects of functional covariates (f_i) and (g_i) measuring the trajectories of parental income and family structure, respectively, on the permanent income of a child (y_i^*). This part includes two unknown functions α and β and one linear operator Γ to be estimated. To estimate these parameters, we approximate our functional covariates by finite dimensional functions and subsequently represent them as finite dimensional vectors. The required procedure involves two steps. In the first step, we consider our functional observations as elements in a Hilbert space $\mathcal{H}$ of square-integrable functions and project them on a finite dimensional subspace $\mathcal{V}_p$ of dimension p, say. In the second step, we represent the functional observations projected on an p-dimensional subspace $\mathcal{V}_p$ of $\mathcal{H}$ as p-dimensional vectors in $\mathbb{R}^p$ using an isometry between $\mathcal{V}_p$ and $\mathbb{R}^p$.[7] Once these two steps are done, we may approximate, represent and estimate our functional parameters α and β as finite dimensional vectors and Γ as a matrix accordingly. This approach has been used widely in functional data analysis, and its basic theory has been developed by Chang et al. (2023).

For a more precise description of our procedure, we let $h = f$ or g and define

$$\Sigma = \frac{1}{n}\sum_{i=1}^{n}\left((h_i - \overline{h}) \otimes (h_i - \overline{h})\right),$$

with $\overline{h} = (1/n)\sum_{i=1}^{n} h_i$, as the sample variance operator of $(h_i)_{i=1}^{n}$.[8] Subsequently, we denote by $(v_k)_{k=1}^{p}$ the set of p-leading functional principal components of Σ and by Π the orthogonal projection on the p-dimensional subspace $\mathcal{V}_p$ of $\mathcal{H}$, and define a map $\varphi : h \mapsto [h]$, where $[h]$

[7] An isometry is a map, which is bijective and preserves norm.

[8] Here $\otimes$ denotes the tensor product in $\mathcal{H}$, which reduces to the outer product if (h_i) are vectors instead of functions.

defined as

$$[h] = \begin{bmatrix} \langle v_1, h \rangle \\ \vdots \\ \langle v_p, h \rangle \end{bmatrix}$$

is a p-dimensional vector in $\mathbb{R}^p$. The map φ is indeed an isometry between $\mathcal{V}_p$ and $\mathbb{R}^p$.[9] We define the sample variance operator Σ, the leading functional principal components (v_k) of dimension p, the orthogonal projection Π on the linear span $\mathcal{V}_p$ of (v_k) and the isometry φ separately for (f_i) and (g_i), and we use the subscript or superscript f and g to denote them specifically for (f_i) and (g_i), respectively. Our finite-dimensional approximation and representation can be further extended to the operator Γ by introducing the corresponding map $\Phi : \Gamma \mapsto [\Gamma]$, where $[\Gamma]$ defined as

$$[\Gamma] = \begin{bmatrix} \langle v_1^f, \Gamma v_1^g \rangle & \dots & \langle v_1^f, \Gamma v_{p_g}^g \rangle \\ \vdots & & \vdots \\ \langle v_{p_f}^f, \Gamma v_1^g \rangle & \dots & \langle v_{p_f}^f, \Gamma u_{p_g}^g \rangle \end{bmatrix},$$

is a $p_f \times p_g$ matrix, i.e., an element in $\mathbb{R}^{p_f \times p_g}$. We choose p_f and p_g jointly by a cross-validation for our empirical analysis in the paper.

Now we may deduce that

$$\langle \alpha, f_i \rangle \approx \langle \alpha, \Pi_f f_i \rangle = [\alpha]'[f_i], \qquad \langle \beta, g_i \rangle \approx \langle \beta, \Pi_g g_i \rangle = [\beta]'[g_i]$$

and

$$\langle f_i, \Gamma g_i \rangle \approx \langle \Pi_f f_i, \Gamma \Pi_g g_i \rangle = [f_i]'[\Gamma][g_i],$$

[9]For $h \in \mathcal{V}_p$, we have $h = \sum_{k=1}^p \langle v_k, h \rangle v_k$, from which it follows that $\|h\|^2 = \sum_{k=1}^p \langle v_k, h \rangle^2 = \|[h]\|^2$, where we use the same generic notation for the norms in $\mathcal{V}_p$ and $\mathbb{R}^p$. This implies, in particular, that any distance minimization problem in a function space $\mathcal{V}_p$ can be formulated as an equivalent problem in an Euclidean space $\mathbb{R}^p$.

so that the main part of our model can be approximated as

$$\langle\alpha, f_i\rangle + \langle\beta, g_i\rangle + \langle f_i, \Gamma g_i\rangle \approx [\alpha]'[f_i] + [\beta]'[g_i] + [f_i]'[\Gamma][g_i]$$

for $i = 1, \ldots, n$. This is because

$$\Pi_f f_i = \sum_{k=1}^{p_f} \langle v_k^f, f_i\rangle v_k^f \quad \text{and} \quad \Pi_g g_i = \sum_{k=1}^{p_g} \langle v_k^g, g_i\rangle v_k^g,$$

from which it follows that

$$\langle\alpha, \Pi_f f\rangle = \sum_{k=1}^{p_f} \langle v_k^f, \alpha\rangle\langle v_k^f, f_i\rangle, \qquad \langle\beta, \Pi_g g_i\rangle = \sum_{k=1}^{p_g} \langle v_k^g, \beta\rangle\langle v_k^g, g_i\rangle$$

and

$$\langle\Pi_f f_i, \Gamma\Pi_g g_i\rangle = \sum_{a=1}^{p_f}\sum_{b=1}^{p_g} \langle v_a^f, f_i\rangle\langle v_b^g, g_i\rangle\langle v_a^f, \Gamma v_b^g\rangle.$$

for $i = 1, \ldots, n$. Our finite-dimensional approximations and representations ensure that the estimation procedure remains computationally tractable while preserving and reflecting major variations in functional observations. Chang et al. (2023) provide the background theory of our approach here, and Chang et al. (2024) show that it has some well defined optimality properties.

3.4 Control Part

To estimate the function τ of control variable x_i nonparametrically, we employ a machine learning technique called the kernel trick with a suitable regularization procedure. The kernel trick relies heavily on the theory of reproducing kernel Hilbert space (RKHS), providing a structured and computationally efficient approach to modeling and estimating a general nonlinear relationship. For our analysis, we assume that τ is a continuous function belonging to the RKHS $\mathcal{H}_K$ associated with a kernel function $K : \mathbb{R}^\ell \times \mathbb{R}^\ell \to \mathbb{R}$ given by $K(x, y) = \exp\left(-\kappa \|x - y\|^2\right)$ for $x, y \in \mathbb{R}^\ell$, called a *radial kernel*, where κ is a scaling parameter. Intuitively, the RKHS $\mathcal{H}_K$ may be thought of as a Hilbert space of functions spanned by a set of basis functions $\big(K(\cdot, x_i)\big)_{i=1}^{\infty}$ with the scaling parameter κ controlling the resolution of these basis functions. The RKHS $\mathcal{H}_K$ is endowed with the inner product $\langle \cdot, \cdot \rangle_K$ defined as $\big\langle K(\cdot, x), K(\cdot, y) \big\rangle_K = K(x, y)$ for all $x, y \in \mathbb{R}^\ell$.[10]

Specifically, we let τ be given by

$$\tau(\cdot) = \sum_{i=1}^{n} c_i K(\cdot, x_i), \tag{3}$$

where $c = (c_i)_{i=1}^n$ is a vector of coefficients for the basis functions $\big(K(\cdot, x_i)\big)_{i=1}^n$. It follows immediately from (3) that

$$\begin{bmatrix} \tau(x_1) \\ \tau(x_2) \\ \vdots \\ \tau(x_n) \end{bmatrix} = \begin{bmatrix} K(x_1, x_1) & K(x_1, x_2) & \cdots & K(x_1, x_n) \\ K(x_2, x_1) & K(x_2, x_2) & \cdots & K(x_2, x_n) \\ \vdots & \vdots & & \vdots \\ K(x_n, x_1) & K(x_n, x_2) & \cdots & K(x_n, x_n) \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix},$$

or more succinctly, $\tau_v = K_m c$ with τ_v defined as an n-dimensional vector with $\tau(x_i)$ in its i-th entry for $i = 1, \ldots, n$, and K_m as an $n \times n$ matrix with $K(x_i, x_j)$ in its (i, j)-th entry for $i, j = 1, \ldots, n$.[11] Note that, as long as K_m is of full rank, we may perfectly fit τ_v by choosing c appropriately. This means that we need a regularization procedure to avoid overfitting and estimate τ reliably.

In line with our approach to deal with infinite dimensionality of the functional terms, we use a method of effectively reducing the dimension of an n-dimensional unknown parameter

[10]The kernel function $K(x, y)$ may thus be *reproduced* in a RKHS by the inner product of basis functions $K(\cdot, x)$ and $K(\cdot, y)$ for all $x, y \in \mathbb{R}^\ell$.

[11]The $n \times n$ matrix K_m is called the Gram matrix of a kernel K.

$[c]$ to regularize our estimator for τ. More explicitly, we set $c = H_q\gamma$ for $\gamma \in \mathbb{R}^q$, where H_q is an $n \times q$ semi-orthogonal matrix of leading eigenvectors of K_m with $q \ll n$. This will reduce the number of parameters involved in the estimation of τ from n to $q \ll n$. To see that our dimension reduction here is effective, note that the spectral representation of K_m yields $K_m = H\Lambda H'$, where Λ is an n-dimensional diagonal matrix with the eigenvalues of K_m in its diagonal entries in descending order and H is an orthogonal matrix of the corresponding eigenvectors in its columns. Note that $K_m c = K_m H_q \gamma = H_q \Lambda_q \gamma$, where H_q is as defined above and Λ_q is a diagonal matrix with the q-leading eigenvalues of K_m in its diagonal entries. Our reparametrization from c to γ thus allows us to focus on the q-leading principal components of K_m – ignoring all other relatively minor components – which can be estimated with highest precision for any given q.[12] Under our reparametrization, τ is estimated simply by choosing a q-dimensional parameter γ, and for this reason we denote it as $[\tau]$, i.e., $\gamma \equiv [\tau]$ in our subsequent discussions. We choose the dimension q of the parameter $[\tau]$ by a cross-validation, jointly with the dimension p of the finite dimensional vectors we need to select to approximate our functional observations as discussed in the previous section.

3.5 Error Term

To model the error distribution nonparametrically, we follow the approach of Gallant and Nychka (1987) and consider a class of density functions $\mathcal{P}$ specified by

$$\rho(\varepsilon) = \frac{1}{c(\delta, \epsilon)} \left(1 + \sum_{k=1}^{r} \delta_k \varepsilon^k \right)^2 \phi(\varepsilon) + \epsilon\psi(\varepsilon), \tag{4}$$

[12] Alternatively, we may set $\gamma = c$ and include a penalty term $\lambda\|\tau\|_K^2$ to the likelihood function, where $\|\tau\|_K^2 = \langle \tau, \tau \rangle_K = \gamma'[K]\gamma$ is the squared norm introduced in the RKHS $\mathcal{H}_K$. However, we do not take this approach in the paper, since all other components are regularized by reducing their ranks instead of introducing penalty terms.

where ϕ is the standard normal density function, $\delta = (\delta_k)_{k=1}^r$ is a vector of polynomial coefficients to be estimated, $\epsilon > 0$ is some number and ψ is a strictly positive density function. The constant $c(\delta, \epsilon)$ serves as a normalization factor, and the density ρ is standardized to satisfy $\int \varepsilon \, \rho(\varepsilon) \, d\varepsilon = 0$ and $\int \varepsilon^2 \rho(\varepsilon) \, d\varepsilon = 1$ for identification.[13] The term $\epsilon\psi$ serves as a lower bound to ensure that the log density is uniformly dominated by an integrable function, which is required to establish a uniform law of large numbers for the log-likelihood function. In practice, ϵ can be chosen arbitrarily small, allowing this term to be effectively disregarded during implementation.

The flexibility of this density specification lies in its ability to approximate a wide variety of densities arbitrarily well by adjusting the order of the polynomial r and the corresponding coefficients δ. This adaptability makes it particularly suitable for capturing non-standard error distributions in many empirical applications. Following our convention throughout the paper, we subsequently denote the r-dimensional parameter δ by $[\rho]$, i.e., $\delta \equiv [\rho]$. We choose the dimension r of $[\rho]$ by a cross-validation jointly with all other tuning parameters in our empirical analysis.

3.6 Estimation

The estimation of θ is based on maximizing the log-likelihood function given by

$$\log \mathcal{L} = \sum_{i=1}^{n} \sum_{j=1}^{m} y_{ij} \log \pi_j(f_i, g_i, x_i; \theta),$$

where y_{ij} is an indicator function that equals 1 if the i-th observation belongs to the j-th class, and $\pi_j(f_i, g_i, x_i; \theta)$ is the corresponding class probabilities, as a function of the parameter $\theta \in \Theta$. The parameter θ includes functional and operative parameters α, β, Γ, τ and ρ, as

[13]The interested reader is referred to, e.g., Yan (2024) and Chang et al. (2025) for more details on how to normalize ρ by setting the normalizing constant appropriately and standardize ρ by imposing the required moment restrictions.

well as a vector parameter μ, which appear due to the presence of functional covariates and the nonparametric nature of our model. Specifically, we have

$$\theta = (\mu, \alpha, \beta, \Gamma, \tau, \rho) \in \mathcal{H} \times \mathcal{H} \times L(\mathcal{H}) \times \mathcal{T} \times \mathcal{P} = \Theta, \tag{5}$$

as defined earlier. A direct maximization of the log-likelihood function $\log \mathcal{L}$ with respect to the parameter θ consisting of such a diverse set of parameters over a multiple products of infinite dimensional parameter space Θ seems impossible.

We approximate and represent the functional and operative parameters $\alpha, \beta, \Gamma, \tau$ and ρ by finite dimensional vectors and matrices and denote them as $[\alpha], [\beta], [\Gamma], [\tau]$ and $[\rho]$, respectively, and define $[\theta] \in [\Theta]$ by

$$[\theta] = \big(\mu, [\alpha], [\beta], [\Gamma], [\tau], [\rho]\big) \in \mathbb{R}^{m-1} \times \mathbb{R}^{p_f} \times \mathbb{R}^{p_g} \times \mathbb{R}^{p_f \times p_g} \times \mathbb{R}^{q} \times \mathbb{R}^{r} = [\Theta], \tag{6}$$

correspondingly as $\theta \in \Theta$ in (5). Note that all functional and operative parameters are now given by the corresponding finite dimensional parameters denoted with brackets. Then we maximize the approximate log-likelihood function

$$\log[\mathcal{L}] = \sum_{i=1}^{n} \sum_{j=1}^{m} y_{ij} \log \pi_j([f_i], [g_i], x_i; [\theta]),$$

with respect to a finite-dimensional parameter $[\theta]$ over the finite-dimensional parameter space $[\Theta]$ introduced in (6), in place of the original log-likelihood function $\log \mathcal{L}$, to obtain an approximate maximum likelihood estimator $\widehat{[\theta]}$ of $[\theta]$. Given an estimator $\widehat{[\theta]}$ of $[\theta]$, it is rather straightforward to obtain the corresponding estimator $\hat{\theta}$ of θ from the one-to-one relationship between $[\theta]$ and θ over $[\Theta]$ and a subset of Θ defined by our approximation of functional and operative parameters as finite dimensional vectors and matrices.

The confidence bands of our estimators are obtained by the bootstrap. For our bootstrap,

(i) we resample (y_i^*) using a regression bootstrap with fixed design based on our model (2) given by the estimated functional and operative parameters $\widehat{\alpha}$, $\widehat{\beta}$, $\widehat{\Gamma}$, $\widehat{\tau}$, and the errors generated as i.i.d. draws from the probability density $\widehat{\rho}$ assuming (f_i, g_i, x_i) are fixed for $i = 1, \ldots, n$, and (ii) (y_i) are resampled subsequently by defining y_i to be in category j if and only if $\widehat{\mu}_{j-1} < y_i^* \leqslant \widehat{\mu}_j$ for $i = 1, \ldots, n$ and $j = 1, \ldots, m$.

We may establish consistency of our maximum likelihood estimator $\widehat{\theta}$ and their bootstrapped confidence bands under suitable regularity conditions if we let $p_f, p_g \to \infty$, $q \to \infty$ and $r \to \infty$ as $n \to \infty$ at appropriate rates. For a binary choice model given by $y_i^* = \lambda w_i + \tau(x_i) + \varepsilon_i$, consistency of $\left(\widehat{\tau}, \widehat{\rho}\right)$ has been established by Yan (2024). Moreover, Chang et al. (2023), which has obtained consistency of our functional and operative parameter estimators and their bootstrap confidence bands for general linear functional models, suggest that the functional and operative estimators $\left(\widehat{\alpha}, \widehat{\beta}, \widehat{\Gamma}\right)$ and their confidence bands are consistent under mild regularity conditions.[14]

4 Data

Our analysis is based on data from the Panel Study of Income Dynamics (PSID), a comprehensive longitudinal survey conducted in the United States. Initiated in 1968, the PSID tracks individuals and their descendants over several decades, providing annual updates through 1997, with biannual updates thereafter until 2021.[15] Our sample consists of individuals who were born no earlier than 1967, and reach at least age 31 by 2021. We also track their parents, ultimately analyzing child-parent pairs.

The dependent variable, the child's adult income class, is constructed from the child's average income during ages 31 to 35 (inclusive). We only include children who have at least

[14] A formal proof for consistency of our procedure is possible, it will be quite lengthy and so is omitted here.

[15] At the time of our analysis, the 2021 wave was the most recent release of PSID data. The subsequent 2023 wave, which became available in June 2025, was not incorporated into the construction of our sample.

three incomes reported from age 31 to 35. We use the Pew Research Center's methodology to classify children into low-, middle-, and high-income categories (Pew Research Center, 2020). To do this, we calculate the median income of the children and define the low-income threshold as two-thirds of this median, while the high-income threshold is set at twice the median income.

The independent variables we use are parental income trajectory during the child's age from 0 to 19 (both included), family structure trajectory during the child's age from 0 to 19 (inclusive), child gender, child race (Black or non-Black), whether the child is underweight at birth, parental college education, parental average age at childbirth, and region (Northeast, North, South, West, and Alaska and Hawaii).

Parental income trajectories are constructed from parental income at the child's ages from 0 to 19. We require that parental incomes for each individual have at most one missing during the 20 years, and income must be available at age 0 and at age 19. If income data are missing, we use linear interpolation to impute the missing values. For each individual, to construct a continuous income trajectory f_i from these incomes reported at discrete time points, we take the domain of the trajectory function to be $[0, 20)$, and construct a step function $f_i^s(t)$ so that $f_i^s(t)$ takes the value of the income reported at age $[t]$, where $[t]$ denotes the integer part of t. We then smooth this step function by projecting it onto a space spanned by 14 B-spline basis functions. The projected function is used as the income trajectory f_i in the estimation. Similarly, we use the family structure indicator (1 for two-parent families, 0 otherwise) from the child's ages 0 to 19, together with B-spline basis projections, to construct the family structure trajectory, allowing for at most one missing observation over the 20-year period.

Our sample consists of a total of 1058 child-parent pairs. We deflate income by the Consumer Price Index for All Urban Consumers (CPI-U-RS, 1977 = 100), following standard practice. Table 2 provides the summary statistics of the variables we use in our study.

Table 2: Summary Statistics

Variable	Type	Mean	Std	Min	Max
log child income	continuous	10.351	0.743	5.556	13.104
low income class	dummy	0.284	0.451	0.000	1.000
middle income class	dummy	0.611	0.488	0.000	1.000
high income class	dummy	0.106	0.308	0.000	1.000
male	dummy	0.510	0.500	0.000	1.000
Black	dummy	0.193	0.395	0.000	1.000
parent college degree	dummy	0.373	0.484	0.000	1.000
underweight at birth	dummy	0.060	0.238	0.000	1.000
parental age at birth	continuous	26.974	5.114	17.000	44.500
region	categorical				
northeast	dummy	0.147	0.354	0.000	1.000
north central	dummy	0.289	0.453	0.000	1.000
south	dummy	0.397	0.489	0.000	1.000
west	dummy	0.164	0.370	0.000	1.000
Alaska & Hawaii	dummy	0.003	0.053	0.000	1.000
N = 1058					

The upper left panel of Figure 1 illustrates the age-specific distribution of family income trajectories for our sample, presenting the 10th, 30th, 50th, 70th, and 90th percentiles as a function of child age from birth to age 20. Family income increases steadily across all percentiles as children grow older; however, the rate of increase is slower at the 10th percentile compared to the 90th percentile. As a result, the income distribution becomes more dispersed with age, reflecting increasing income inequality over the child's developmental period. When children are young (or when parents are early in their life cycle), observed incomes are more concentrated so that high permanent incomes tend to be underrepresented, while low permanent incomes are overrepresented. As children age, observed incomes provide a more accurate reflection of permanent income levels, contributing to the widening distribution observed over time.

The lower left panel of Figure 1 illustrates the age-specific mean of the family structure trajectory, capturing the proportion of children living in two-parent households at each age. The plot reveals an initial increase in the proportion of two-parent families during the first two years, followed by a steady decline throughout childhood and adolescence, continuing (at least) until the child reaches age 20. This pattern reflects substantial changes in family structure over the course of a child's development. In particular, after early childhood, family transitions such as parental separation, divorce, or changes in cohabitation status become more prevalent. As a result, the proportion of children living in two-parent households steadily declines. These dynamics highlight the importance of examining family structure as a trajectory over time rather than relying solely on measures that aggregate over different years.

5 Empirical Results

In order to implement our multinomial probability framework we need to select the tuning parameters outlined in Sections 3.3–3.6. We do this by 10-fold cross-validation. This procedure selects one functional principal component (FPC) for the parental income trajectory ($p_f = 1$), one FPC for the family structure trajectory ($p_g = 1$), five principal components for the Gram matrix ($q = 5$), and a second-order polynomial ($r = 2$) in the representation of the density of the random term.

The upper and lower right panels of Figure 1 display respective scree plots for the parental income and family structure trajectories. These plots illustrate the cumulative proportion of variance explained by each FPC. For both trajectories, the first FPC accounts for slightly more than half of the total variance, while the first two FPCs together explain just under 70%. Our benchmark results are based on the selection of $p_f = p_g = 1$ through cross-validation; however, we also report results for alternative choices of the number of FPCs to assess robustness.

5.1 Main Effects of Parental Income and Family Structure

The functional parameters $\alpha(t)$ and $\beta(t)$, defined as functions of age t, capture the age-specific effects of parental income and family structure, respectively, on a child's latent permanent income, which serves as the underlying determinant of the child's income class in adulthood. The estimates of these functional coefficients, $\alpha(t)$ and $\beta(t)$, along with 90% and 95% pointwise confidence bands, are presented in the top-left and bottom-left panels of Figure 2.

Our results indicate that parental income has a consistently positive effect on a child's adult income across all stages of childhood, as would naturally be expected and which is an implication of models that map permanent income (measured as average income) of parents to children. However, the magnitude of this effect varies by developmental period, a relationship which is obscured by conditioning on a single income statistic as opposed to the parental income trajectory. Notably, parental income during middle childhood and adolescence exerts a significantly stronger influence on a child's future income than parental income during early childhood. This finding contrasts with the tenor of the early childhood development literature (e.g., Heckman and Mosso, 2014), which emphasizes early childhood as a critical period for skill formation and long-term outcomes. Our evidence does not logically contradict the early childhood literature, as claims about early childhood involve the skills production function and not parental income effects per se. However, the sensitivity of adult outcomes to adolescent incomes suggests that the uses of this income have distinct effects from early childhood investment. As discussed in Chang et al. (2024), one reason why this finding is credible is that incomes in later childhood and adolescence influence the schools and neighborhoods, which are social determinants of development that are not operational for younger years; see Wodtke et al. (2016) for evidence to this effect. Our essentially monotonically increasing functional coefficient curve contrasts with some other studies, notably Duncan et al. (2010) and Carneiro et al. (2021) that argue that middle childhood has weaker effects than early childhood and adolescence. However, our finding is consistent with Chang et al. (2024). This discrepancy illustrates, in our judgment, the benefit of the trajectory approach as other methods place a priori restrictions on the ways that parental income trajectories affect children that do not have strong theoretical foundations.

Family structure plays a distinct role in shaping long-term economic outcomes. Children raised in two-parent households at any stage of childhood tend to achieve higher adult income status than those raised in single-parent households. While this advantage is relatively modest during early childhood, it becomes pronounced in middle childhood and adolescence, peaking during the late high school years at aournd age 17.

5.2 Interactions

The upper-right panel of Figure 2 displays the estimated interaction function $\widehat{\Gamma}(t, s)$ as a bivariate function of ages t and s using a heatmap. The lower-right panel displays heatmap colors for the grid cells that are statistically significantly different from zero at the 5% level, leaving others blank. The results reveal that parental income and family structure generally exhibit a complementary relationship in influencing a child's long-term economic outcomes. Rather than acting as independent determinants, these factors interact synergistically, amplifying their combined impact on a child's adult income. This complementarity underscores the importance of considering the joint influence of family financial resources and household structure when evaluating the determinants of economic mobility.

The observed synergy between parental income and family structure is evident across all stages of childhood. At nearly every developmental stage, higher parental income is more beneficial when paired with a stable two-parent family structure, and likewise, the benefits of a stable family structure are amplified in the presence of greater financial resources.

This pattern suggests that these two factors operate not in isolation but in a mutually reinforcing manner, jointly shaping children's development and future outcomes. Rather than functioning as substitutes, financial resources and family stability appear to complement each other in promoting long-term success.

This complementary effect grows stronger as children age and is particularly pronounced between parental income during middle to late childhood and family structure during late childhood, identifying a key developmental window in which their joint influence on adult income is greatest. This stage coincides with increasing academic demands, the formation of long-term aspirations, and heightened exposure to peer, social, and institutional influences. It is also a period of growing autonomy, when children begin to make consequential decisions about their education, relationships, and identity. In this context, financial resources can facilitate access to enriching experiences, high-quality education, and reduced stress, while family stability can provide emotional security, supervision, and guidance.

These findings underscore the importance of modeling child development as a dynamic, stage-sensitive process shaped by interacting inputs, rather than static effects of isolated factors.

5.3 Robustness Checks

We report in this section estimation results obtained under alternative specifications to our benchmark setting. To preserve space, the corresponding figures are provided in the online appendix. Specifically, instead of estimating the density ρ of the error term ε flexibly as in (4), we impose parametric assumptions by specifying ε to follow either a standard normal distribution or a standard logistic distribution. We then estimate the parameters $[\theta]$ as in (6), excluding $[\rho]$, and present the estimated functional coefficients $\widehat{\alpha}(t), \widehat{\beta}(t)$ and the interaction function $\widehat{\Gamma}(t, s)$ in Figures A1 and A2, respectively, in the online appendix.

These two specifications correspond to the widely used probit and logit models in economics. It is important to note that, although the distribution of the error term is fixed in these alternative settings, the control function $\tau(\cdot)$ continues to be estimated nonparametrically, and the model retains its functional components.

The results indicate that replacing the nonparametric specification of the error distribution with either of these parametric alternatives has little impact on the overall findings. This robustness suggests that our main results are not driven by the flexible specification of the unobserved heterogeneity but rather by the functional structure of the model and the information contained in the observed trajectories.

We also examine the sensitivity of our results to the number of FPCs used to represent the functional data, which in our benchmark analysis is determined via cross-validation. Specifically, we set (p_f, p_g)—the numbers of FPCs for family income trajectories and family structure trajectories—to (2, 1), (1, 2), and (2, 2), and report the corresponding estimation results in Figures A3, A4, and A5, respectively, in the online appendix. Our main qualitative conclusions remain robust across these alternative specifications: timing plays a critical role, with parental income and two-parent status exerting the strongest positive influence on a child's adult income during middle childhood and adolescence. Further, family income and family structure complement each other, with the complementarity effect being most pronounced between parental income during middle to late childhood and family structure during late childhood.

Quantitatively, the estimates exhibit some variation with respect to the choice of tuning parameters. Increasing the number of FPCs for family income trajectories from 1 to 2 flattens the estimated functional coefficient $\widehat{\alpha}(t)$ during early to middle childhood and widens its confidence band in this period, rendering the effect of parental income statistically insignificant at the 5% level. The coefficient then rises more steeply from middle childhood onward and reaches a higher peak in later childhood compared to the benchmark estimate. A similar pattern arises for the estimated functional coefficient $\widehat{\beta}(t)$ for family structure trajectories when increasing the number of FPCs from 1 to 2: the coefficient function becomes flatter, and in some cases even downward-sloping, with reduced statistical significance during early to middle childhood, but picks up more sharply and peaks higher in late childhood relative to the benchmark results.

The estimated interaction effects remain broadly similar to those in the benchmark setting during late childhood. However, the complementarity effect between family income and family structure is weaker and less statistically significant during early childhood under these alternative parameter choices.

This pattern reflects the classic bias-variance trade-off in nonparametric estimation. Including more FPCs reduces approximation bias but increases estimator variance, leading to wider confidence bands and reduced statistical significance for certain effects. Nevertheless, the primary qualitative findings are unchanged. In fact, the increased distinction between early childhood and later developmental stages in the estimates under these alternative specifications reinforces our conclusion that middle to late childhood is a particularly critical period for shaping adult economic outcomes.

Further, we estimate the model using the survey weights provided by the PSID, weighting the log-likelihood contribution of each observation in the estimation accordingly. The results are presented in Figure A6 of the online appendix. All qualitative features are preserved, while the estimates are quantitatively slightly smaller. This attenuation likely reflects that, within our sample and for the specific variables examined, observations with lower response values tend to receive higher weights under the PSID sample design, thereby reducing the weighted estimates relative to the unweighted ones. Nonetheless, the overall patterns and significance levels remain largely unchanged, indicating that the main conclusions are robust to the use of survey weights.

6 Conclusion

This study contributes to the literature on intergenerational mobility by highlighting the importance of both family economic conditions and family structure across various stages of childhood and adolescence. Our findings emphasize the timing and dynamic nature of these influences, demonstrating that their impact on a child's adult income is not uniform throughout childhood. We find that middle childhood and adolescence are particularly critical periods, during which parental resources and family stability have the most pronounced effects on long-term outcomes. By focusing on developmental trajectories rather than aggregate measures, this study provides a more nuanced understanding of how family dynamics influence economic mobility across generations.

Our analysis reveals that family economic conditions and family structure generally function as complements, with their interaction playing a crucial role in shaping adult outcomes. This complementary relationship is especially evident between parental income during middle to late childhood and family structure during late childhood, a novel finding that expands current knowledge on the interplay between financial resources and familial stability. These results have important policy implications, suggesting that interventions aimed at improving both economic resources and family environments should be coordinated and targeted toward critical developmental stages. Such policies could maximize their effectiveness, fostering greater intergenerational mobility and improving the long-term well-being of future generations.

While the current analysis provides evidence of the interactive effects of parental income and family structure trajectories on children's long-term economic outcomes, an important direction for future research is to explore heterogeneity in these patterns across subpopulations. The strength and timing of these effects may differ by gender, race and ethnicity, parental education, or other sociodemographic characteristics. For instance, boys and girls may respond differently to family structure transitions or financial stress during adolescence, and minority children may face compounding disadvantages or varying institutional contexts that shape how family and economic environments translate into adult outcomes. Moreover, the interaction between income and family structure may operate differently depending on cultural norms, access to extended kin networks, or community-level supports. Incorporating such heterogeneity would deepen our understanding of the mechanisms underlying intergenerational mobility and help identify targeted intervention strategies that are responsive to the needs of diverse groups.

References

Björklund, A. and Sundström, M. (2006). Parental separation and children's educational attainment: A siblings analysis on swedish register data. *Economica*, 73(292):605–624.

Bloome, D. (2017). Childhood family structure and intergenerational income mobility in the United States. *Demography*, 54(2):541–569.

Carneiro, P., Lopez-Garcia, I., Salvanes, K., and Tominey, E. (2021). Intergenerational mobility and the timing of parental income. *Journal of Political Economy*, 129(3):757–788.

Chang, Y., Durlauf, S., Hu, B., and Park, J. (2025). Accounting for individual-specific heterogeneity in intergenerational income mobility. *Sociological Methods & Research*.

Chang, Y., Durlauf, S., Lee, S., and Park, J. (2024). A trajectories-based approach to measuring intergenerational mobility. Manuscript.

Chang, Y., Park, J., and Pyun, D. (2023). From functional autoregressions to vector autoregressions. Manuscript.

Chen, J. and Roth, J. (2024). Logs with zeros? Some problems and solutions. *Quarterly Journal of Economics*, 139(2):891–936.

Cheng, S. and Song, X. (2019). Linked lives, linked trajectories: Intergenerational association of intragenerational income mobility. *American Sociological Review*, 84(6):1037–1068.

Conway, J. (2000). *A Course in Functional Analysis*. Springer, 2nd edition.

Duncan, G., Ziol-Guest, K., and Kalil, A. (2010). Early-childhood poverty and adult attainment, behavior, and health. *Child Development*, 81(1):306–325.

Ermisch, J., Francesconi, M., and Pevalin, D. (2004). Parental partnership and joblessness in childhood and their influence on young people's outcomes. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 167(1):69–101.

Francesconi, M., Jenkins, S., and Siedler, T. (2010). Childhood family structure and schooling outcomes: Evidence for germany. *Journal of Population Economics*, 23:1073–1103.

Frimmel, W., Halla, M., and Winter-Ebner, R. (2024). How does parental divorce affect children's long-term outcomes? *Journal of Labor Economics*, 239.

Gallant, A. R. and Nychka, D. (1987). Semi-nonparametric maximum likelihood estimation. *Econometrica*, 55(2):363–390.

Grawe, N. D. (2004). Reconsidering the use of nonlinearities in intergenerational earnings mobility as a test for credit constraints. *Journal of Human Resources*, 39:813–827.

Heckman, J. and Mosso, S. (2014). The economics of human development and social mobility. *Annual Review of Economics*, 6:689–733.

Matzkin, R. (1992). Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models. *Econometrica*, 60(2):239–270.

Moschion, J. and van Ours, J. C. (2019). Do childhood experiences of parental separation lead to homelessness? *European Economic Review*, 111:211–236.

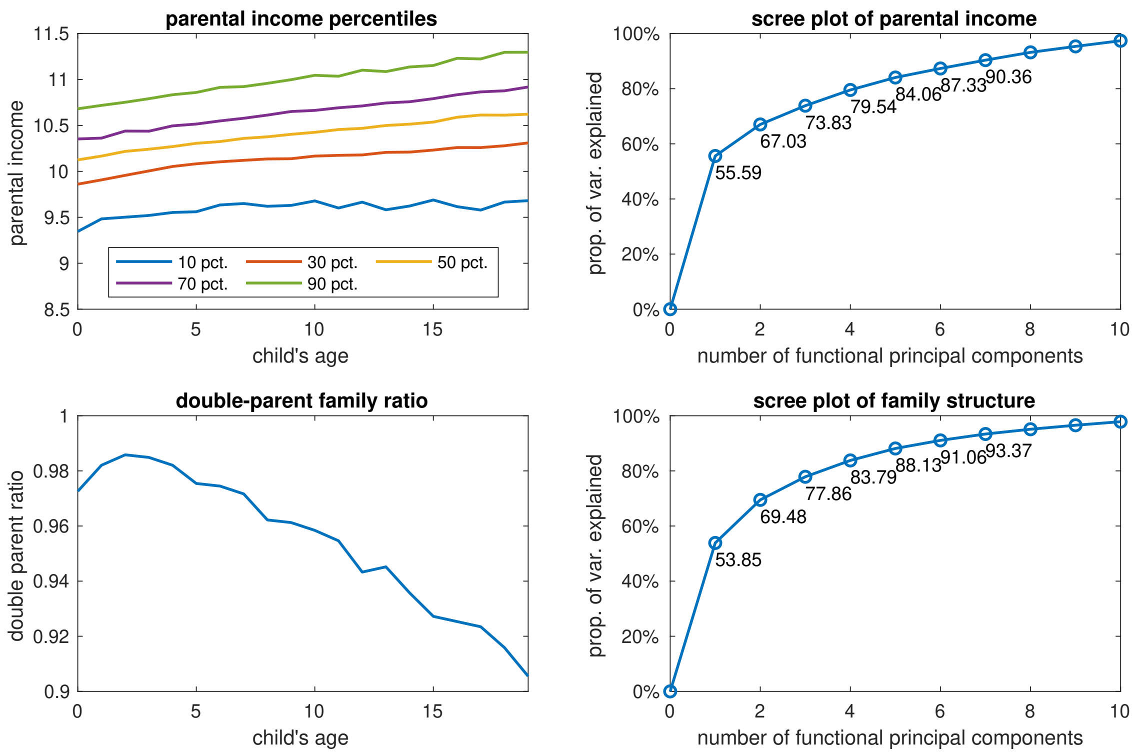
Muller, S. (2010). Another problem in the estimation of intergenerational income mobility. *Economic Letters*, 108(1):291–295.

Pew Research Center (2020). Most americans say there is too much economic inequality in the U.S., but fewer than half call it a top priority. Research Report.

Wodtke, G., Harding, D., and Elwert, F. (2016). Neighborhood effect heterogeneity by family income and developmental period. *American Journal of Sociology*, 121:1168–1222.

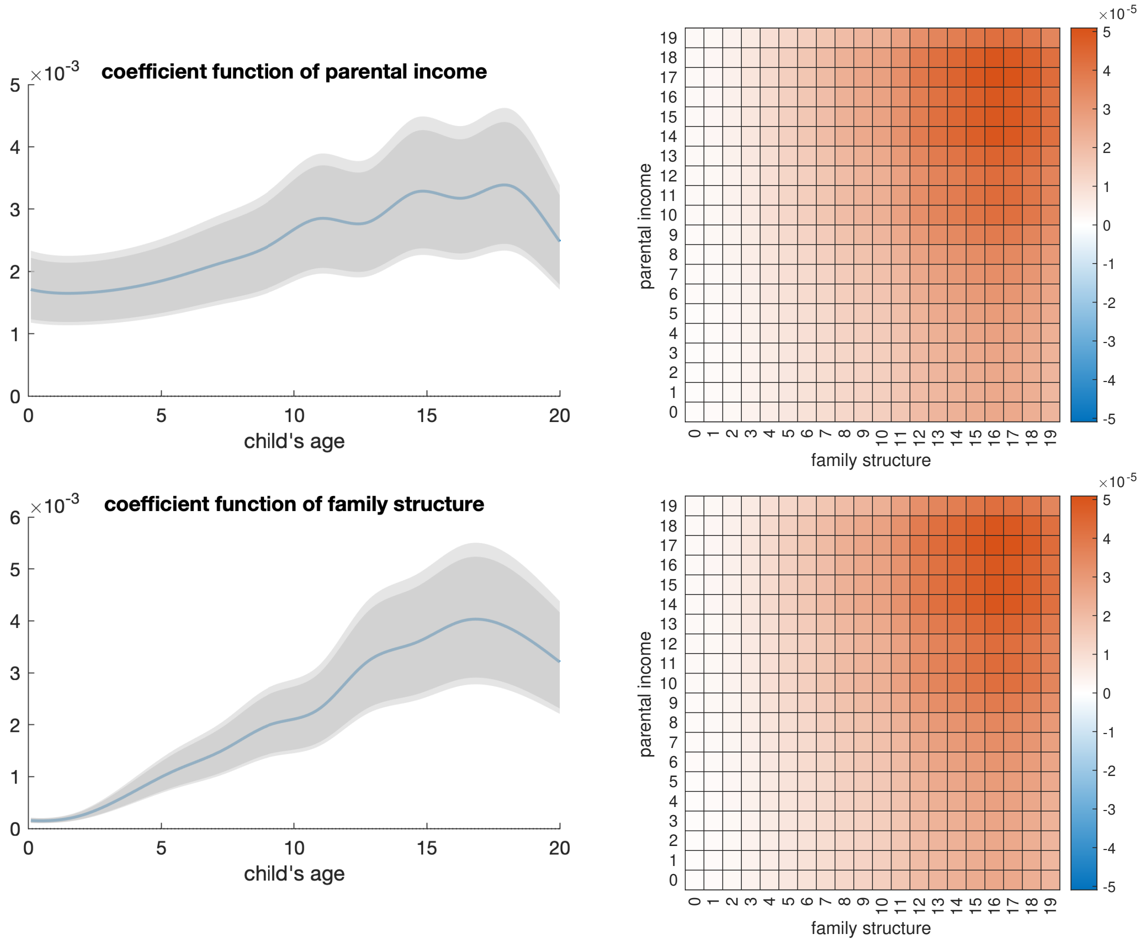
Yan, G. (2024). A kernelization-based approach to nonparametric binary choice models. Manuscript.

Figure 1: Parental Income and Family Structure Trajectories



Notes: The upper-left panel displays the age-specific 10th, 30th, 50th, 70th, and 90th percentiles of family income trajectories as a function of child age, ranging from birth to age 20. The upper-right panel presents a scree plot of parental income trajectories, illustrating the proportion of variance explained by each principal component (cumulatively). The lower-left panel depicts the age-specific mean of the family structure trajectory, which represents the proportion of two-parent families in the sample at each child age. The lower-right panel presents a scree plot of family structure trajectories.

Figure 2: Main and Interaction Effects of Parental Income and Family Structure



Notes: The upper-left panel shows the estimated functional coefficient $\widehat{\alpha}(t)$ for family income trajectories, while the lower-left panel shows the estimated functional coefficient $\widehat{\beta}(t)$ for family structure trajectories. The dark and light shaded areas represent the 90% and 95% confidence bands, respectively. The upper-right panel displays the heat map of the estimated interaction function $\widehat{\Gamma}(t,s)$ between family income and family structure trajectories. The lower-right panel displays the corresponding heat map, retaining only the grid cells with values significantly different from zero at the 5% significance level.

Appendix: Data Details

Since interviews are conducted at different times in different years, we use the individual's year of birth to calculate the age of the individual instead of using the individual's reported age directly. The year of birth is reported annually or biannually by each individual and could be inconsistent across years. If this reported year of birth remains consistent across years, we use that report. In cases where the reported years of birth vary but only by one year, the mode year is selected. If multiple mode years exist, we use the earliest reported year. If the mode year is not unique and does not have a clear majority, 70% frequency and a minimum of three occurrences are used as a threshold for a report to be used.

In scenarios where the year of birth cannot be determined from reported data, we also use information from individuals' reported ages. By subtracting the reported age from the interview year, we get an implied year of birth corresponding to the individual. Similar to the reported years of birth, consistency checks are applied. If the implied year of birth is unique across age reports in different years, it is adopted. When the inferred year of birth varies by one year across age reports, the mode year is selected, prioritizing the first reported instance in cases of ties. In other cases, the same criteria of frequency and minimum occurrences are applied. We construct income profiles based on the year of birth constructed above. For example, an individual's income in the year 1990 will be recorded as his or her age-30 income if his or her year of birth is 1960.

Online Appendix to
The Effects of Parental Income and Family Structure on Intergenerational Mobility: A Trajectories-Based Approach

Yoosoon Chang, Steven N. Durlauf, Bo Hu and Joon Y. Park

July 2025

1 Identification

In this section, we give a formal exposition of the identification of our model. Let (f_i, g_i, x_i) be i.i.d. observations as assumed in the text.

Assumption 1. We assume that
(a) (ε_i) is independent of (f_i, g_i, x_i).
(b) (f_i, g_i, x_i) has support $\mathcal{H} \times \mathcal{H} \times \mathcal{X}$.
(c) $\mathcal{T}$ is a set of continuous functions with zero infimum on $\mathcal{X}$.
(d) $\mathcal{P}$ is a set of strictly positive probability density functions on $\mathbb{R}$ with zero mean and unit variance.

Many of the conditions in Assumption 1 are directly comparable to those used in Matzkin (1992) and Yan (2024) for identification of binary choice models with no functional and operative arguments.

Theorem 1.1. *Under Assumption 1, θ is identified in Θ.*

Proof of Theorem 1.1. To establish identifiability of our model, we need to show that $\theta = \theta_0$ whenever $\pi_j(f, g, x; \theta) = \pi_j(f, g, x; \theta_0)$ for $j = 1, \ldots, m-1$ for all $(f, g, x) \in \mathcal{H} \times \mathcal{H} \times \mathcal{X}$, the support given by the condition (b) in Assumption 1. We will show that our model is identified for $m = 2$, from which it follows that our model is generally over-identified for $m > 2$. Subsequently, we let F and F_0 be the distribution functions of ρ and ρ_0, respectively. We break our proof into two parts.

Identification of α, β and Γ. Let $\pi_j(f, g, x; \theta) = \pi_j(f, g, x_i; \theta_0)$ for all $(f, g, x) \in \mathcal{H} \times \mathcal{H} \times \mathcal{X}$ and for any $j = 1, \ldots, m-1$ given. Assume $m = 2$, in which case μ becomes a scalar. It follows that

$$F\big(\mu-\langle\alpha,f\rangle-\langle\beta,g\rangle-\langle f,\Gamma g\rangle-\tau(x)\big) = F_0\big(\mu_0-\langle\alpha_0,f\rangle-\langle\beta_0,g\rangle-\langle f,\Gamma_0 g\rangle-\tau_0(x)\big)$$

for all $(f,g,x)\in\mathcal{H}\times\mathcal{H}\times\mathcal{X}$. Fix $x\in\mathcal{X}$ and define

$$G(z) = F\big(z+\mu-\tau(x)\big) \quad \text{and} \quad G_0(z) = F_0\big(z+\mu_0-\tau_0(x)\big).$$

It is easy to see that G and G_0 are also distribution functions, and

$$G(-\langle\alpha,f\rangle-\langle\beta,g\rangle-\langle f,\Gamma g\rangle) = G_0(-\langle\alpha_0,f\rangle-\langle\beta_0,g\rangle-\langle f,\Gamma_0 g\rangle)$$

for all $f,g\in\mathcal{H}$. Furthermore, by the condition (d) in Assumption 1, G and G_0 are strictly increasing and have unit variance.

Take $g=0$. If $\alpha_0=0$, we have $G(-\langle\alpha,f\rangle)=G_0(0)$ holds for all $f\in\mathcal{H}$, and therefore, $\alpha=0$. If $\alpha_0\neq 0$, fix an arbitrary f such that $\langle\alpha_0,f\rangle\neq 0$. Then for any $c\in\mathbb{R}$,

$$G\big(-\langle\alpha,cf\rangle\big) = G_0\big(-\langle\alpha_0,cf\rangle\big).$$

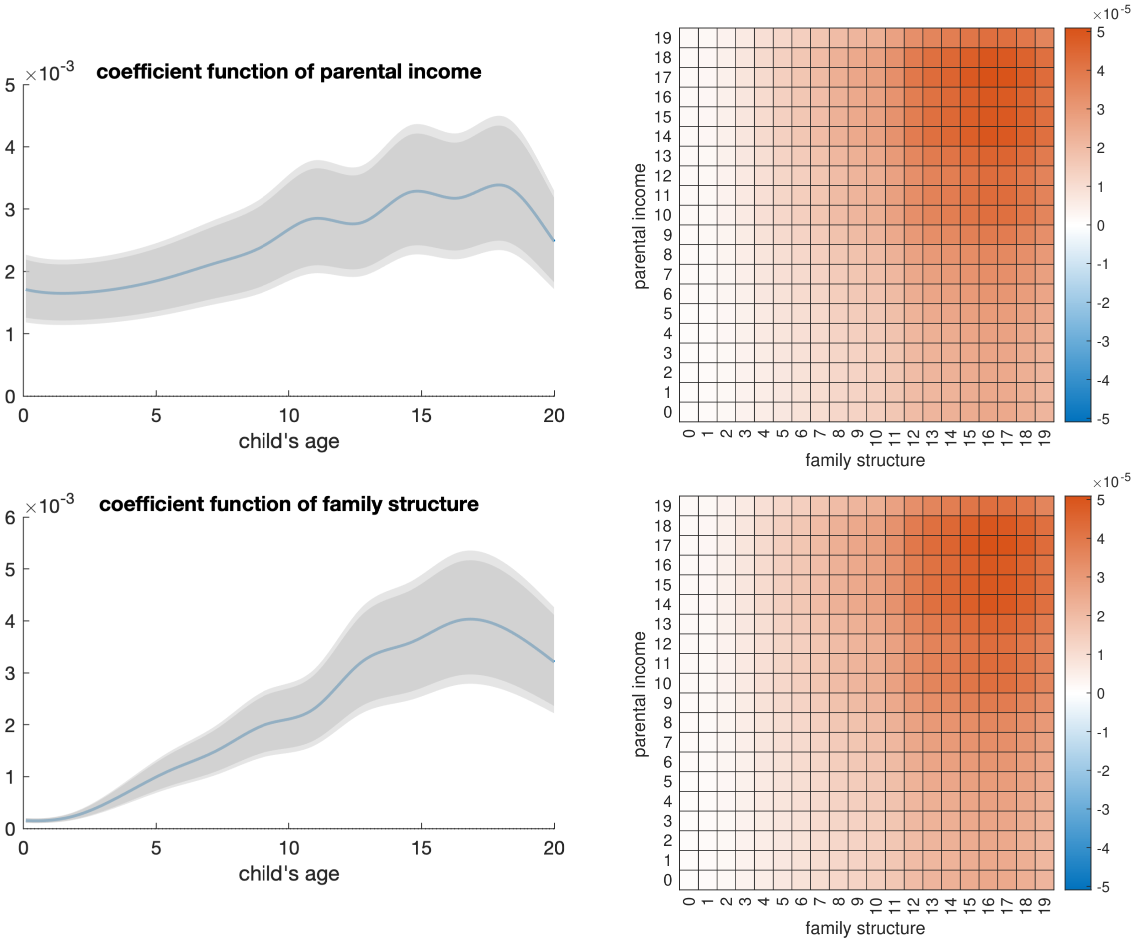
Since both G and G_0 have unit variance, this implies that $\langle\alpha,f\rangle=\langle\alpha_0,f\rangle$, and that $G=G_0$. Since f is arbitrary on $\mathcal{N}(\alpha_0)^\perp$, we conclude that $\alpha=\alpha_0$. We can similarly show that $\beta=\beta_0$ by taking $f=0$ instead. Now we have $G(-\langle\alpha_0,f\rangle-\langle\beta_0,g\rangle-\langle f,\Gamma g\rangle)=G(-\langle\alpha_0,f\rangle-\langle\beta_0,g\rangle-\langle f,\Gamma_0 g\rangle)$ for all $f,g\in\mathcal{H}$. Since G is strictly increasing, we have $\langle f,\Gamma g\rangle=\langle f,\Gamma_0 g\rangle$ for all $f,g\in\mathcal{H}$, which implies that $\Gamma=\Gamma_0$. This shows that α_0,β_0 and Γ_0 are identified.

Identification of μ,τ, and ρ. Once α_0,β_0 and Γ_0 are identified, the distribution of $\langle\alpha_0,f_i\rangle+\langle\beta_0,g_i\rangle+\langle f_i,\Gamma_0 g_i\rangle$ is identified. Also, this distribution has support $\mathbb{R}$. It then follows from Theorem 2.1 of Yan (2024) that τ_0 and ρ_0 are identified. Given that all other parameters are identified, the identification of μ_0 follows from the assumption that τ_0 has infimum 0 over its domain $\mathcal{X}$. □

2 Additional Results for Robustness Check

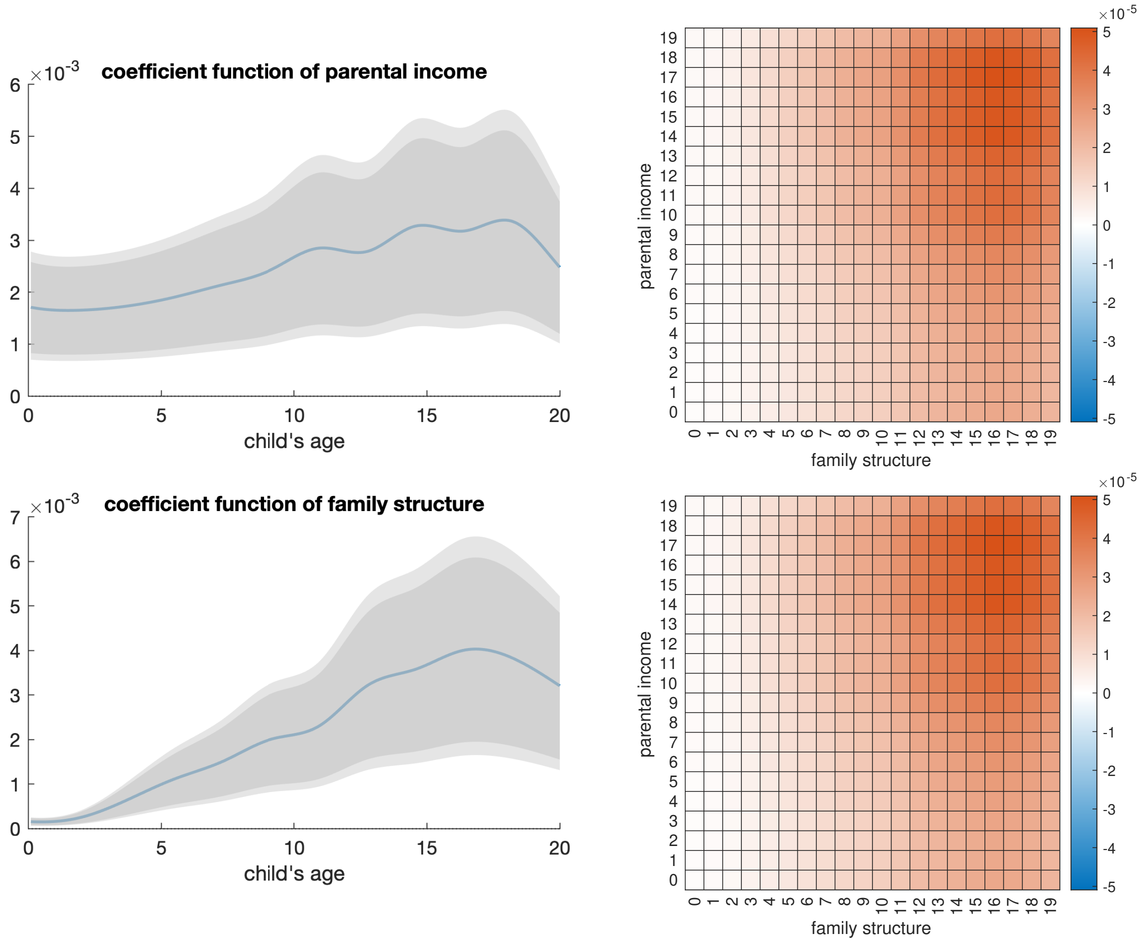
In this online appendix, we present additional estimation results for the model under alternative specifications of the error term distribution, different numbers of functional principal components used to represent family income and family structure trajectories, and the inclusion of survey weights.

Figure A1: Main and Interaction Effects of Parental Income and Family Structure (Probit)



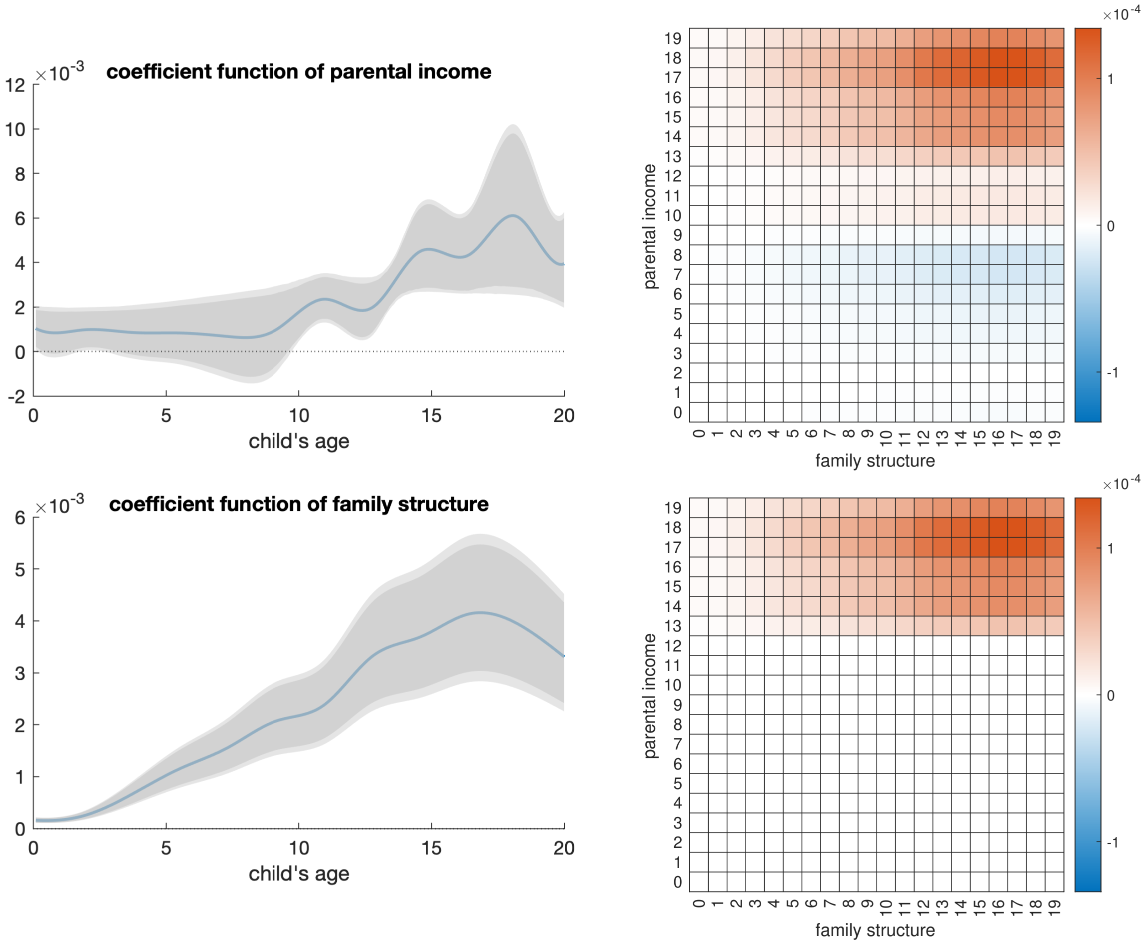
Notes: The upper-left panel shows the estimated functional coefficient $\widehat{\alpha}(t)$ for family income trajectories, while the lower-left panel shows the estimated functional coefficient $\widehat{\beta}(t)$ for family structure trajectories. The dark and light shaded areas represent the 90% and 95% confidence bands, respectively. The upper right panel displays the heat map of the estimated interaction function $\widehat{\Gamma}(t,s)$ between family income and family structure trajectories. The lower-right panel displays the corresponding heat map, retaining only the grid cells with values significantly different from zero at the 5% significance level.

Figure A2: Main and Interaction Effects of Parental Income and Family Structure (Logit)



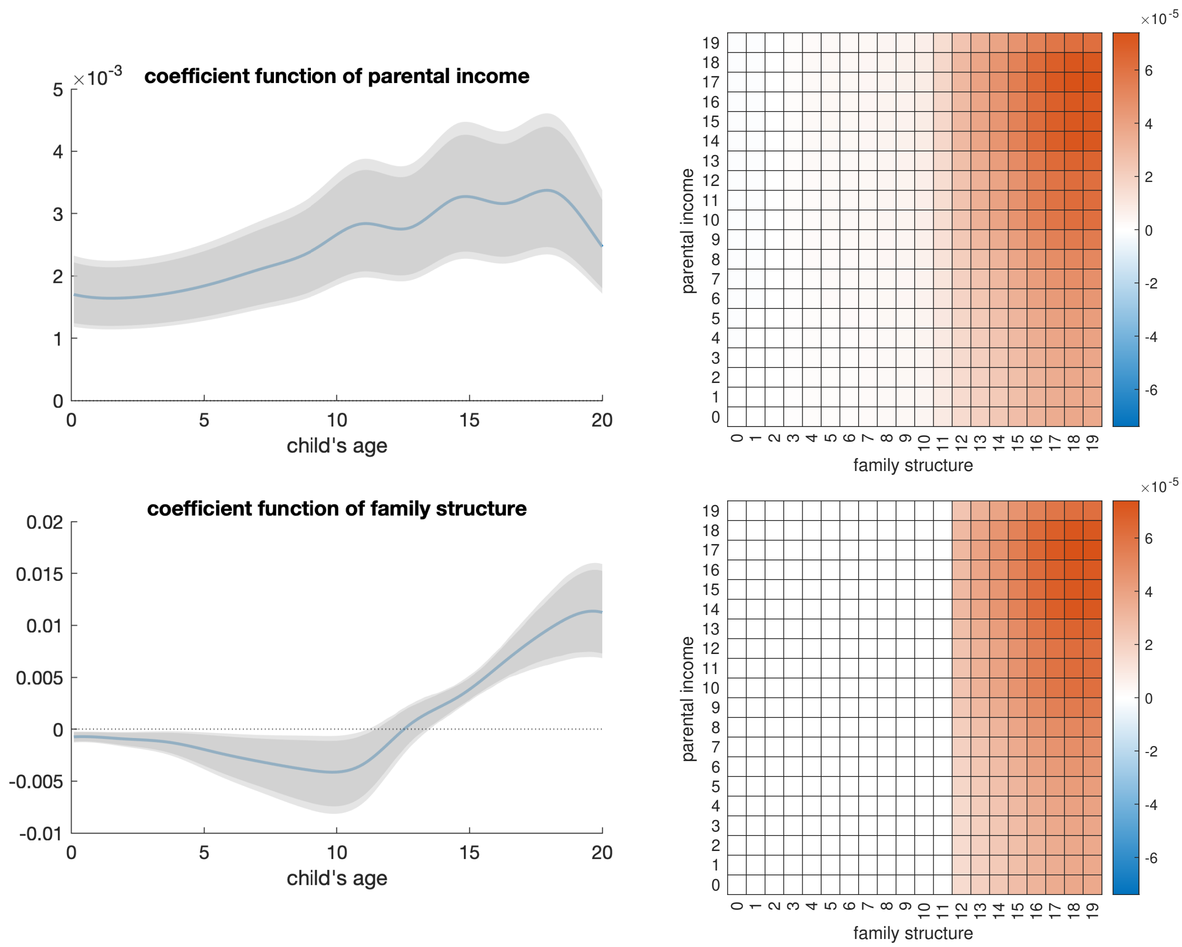
Notes: The upper-left panel shows the estimated functional coefficient $\widehat{\alpha}(t)$ for family income trajectories, while the lower-left panel shows the estimated functional coefficient $\widehat{\beta}(t)$ for family structure trajectories. The dark and light shaded areas represent the 90% and 95% confidence bands, respectively. The upper right panel displays the heat map of the estimated interaction function $\widehat{\Gamma}(t,s)$ between family income and family structure trajectories. The lower-right panel displays the corresponding heat map, retaining only the grid cells with values significantly different from zero at the 5% significance level.

Figure A3: Main and Interaction Effects of Parental Income and Family Structure (with $p_f = 2, p_g = 1$)



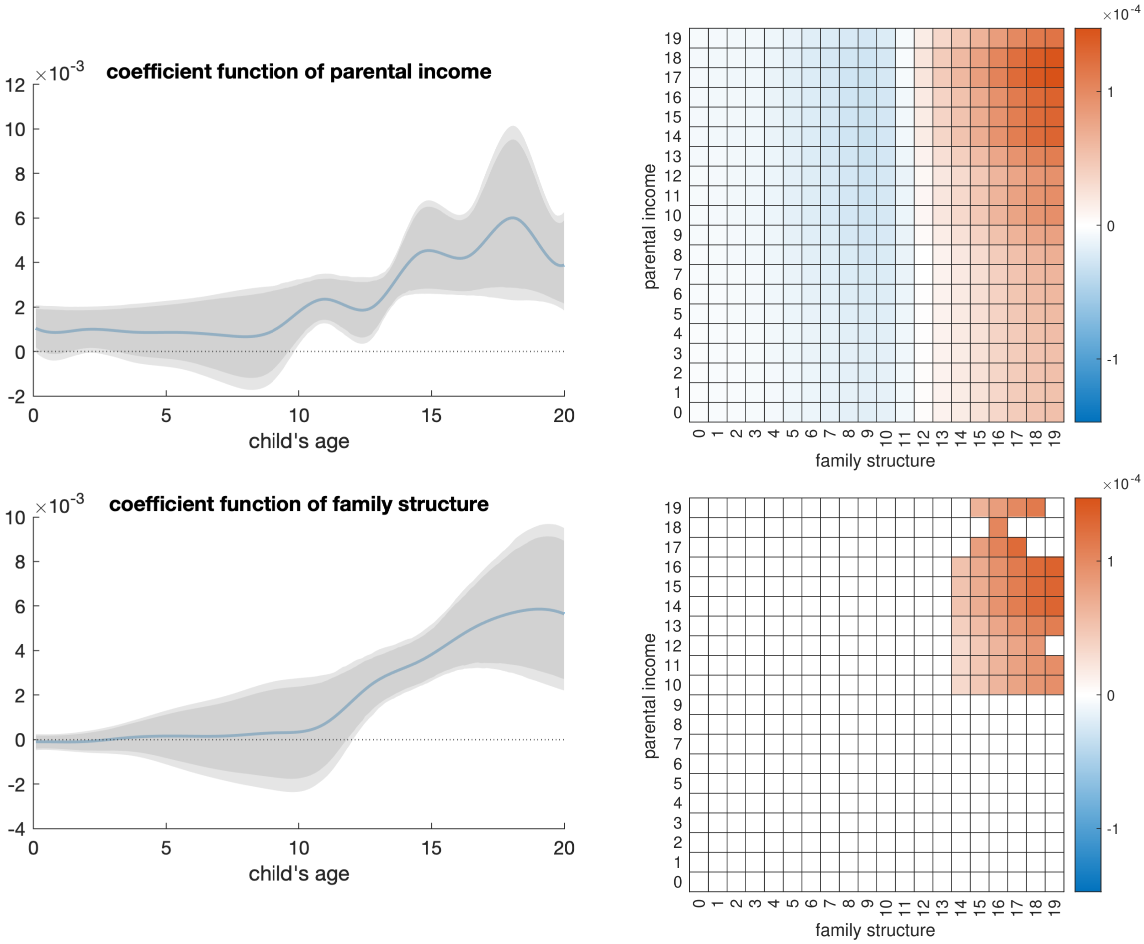
Notes: The upper-left panel shows the estimated functional coefficient $\widehat{\alpha}(t)$ for family income trajectories, while the lower-left panel shows the estimated functional coefficient $\widehat{\beta}(t)$ for family structure trajectories. The dark and light shaded areas represent the 90% and 95% confidence bands, respectively. The upper right panel displays the heat map of the estimated interaction function $\widehat{\Gamma}(t, s)$ between family income and family structure trajectories. The lower-right panel displays the corresponding heat map, retaining only the grid cells with values significantly different from zero at the 5% significance level.

Figure A4: Main and Interaction Effects of Parental Income and Family Structure (with $p_f = 1, p_g = 2$)



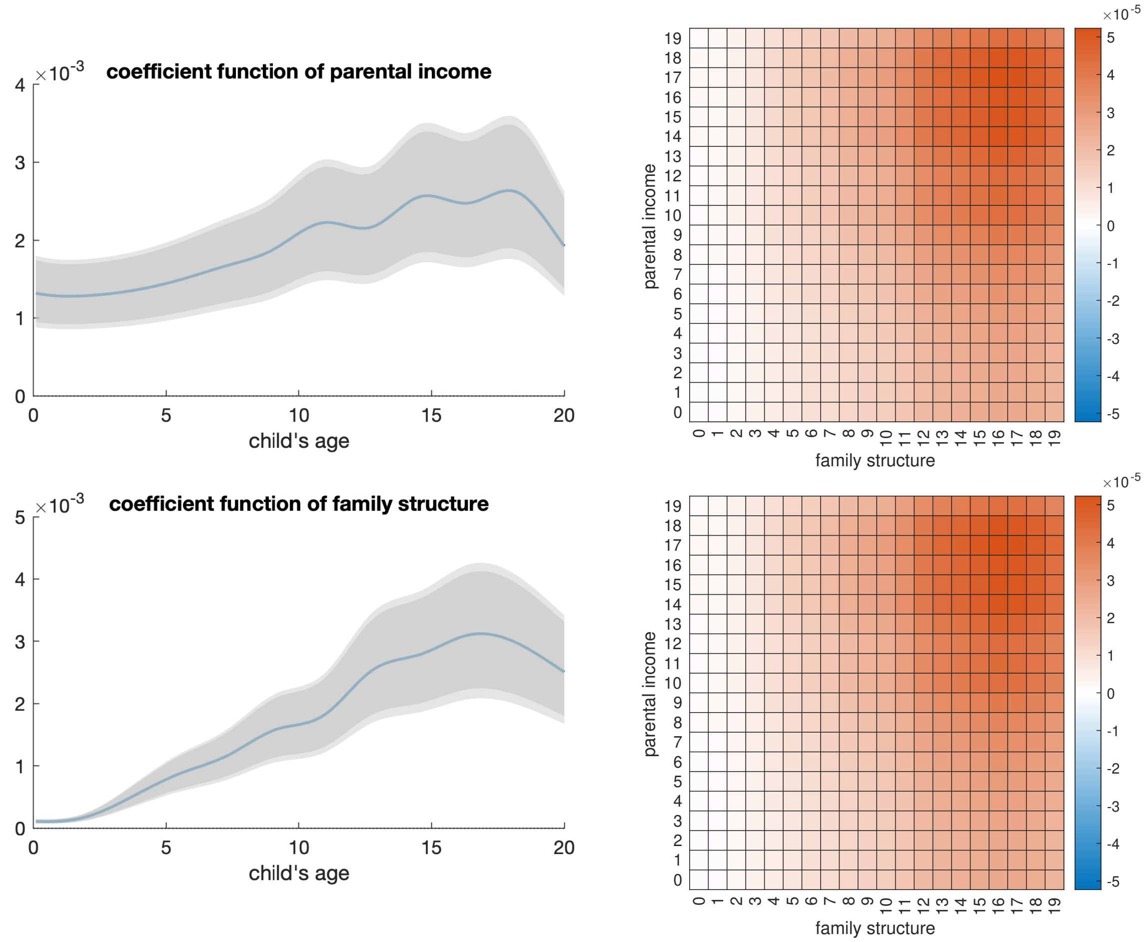
Notes: The upper-left panel shows the estimated functional coefficient $\widehat{\alpha}(t)$ for family income trajectories, while the lower-left panel shows the estimated functional coefficient $\widehat{\beta}(t)$ for family structure trajectories. The dark and light shaded areas represent the 90% and 95% confidence bands, respectively. The upper right panel displays the heat map of the estimated interaction function $\widehat{\Gamma}(t, s)$ between family income and family structure trajectories. The lower-right panel displays the corresponding heat map, retaining only the grid cells with values significantly different from zero at the 5% significance level.

Figure A5: Main and Interaction Effects of Parental Income and Family Structure (with $p_f = 2, p_g = 2$)



Notes: The upper-left panel shows the estimated functional coefficient $\widehat{\alpha}(t)$ for family income trajectories, while the lower-left panel shows the estimated functional coefficient $\widehat{\beta}(t)$ for family structure trajectories. The dark and light shaded areas represent the 90% and 95% confidence bands, respectively. The upper right panel displays the heat map of the estimated interaction function $\widehat{\Gamma}(t, s)$ between family income and family structure trajectories. The lower-right panel displays the corresponding heat map, retaining only the grid cells with values significantly different from zero at the 5% significance level.

Figure A6: Main and Interaction Effects of Parental Income and Family Structure (with Survey Weights)



Notes: The upper-left panel shows the estimated functional coefficient $\widehat{\alpha}(t)$ for family income trajectories, while the lower-left panel shows the estimated functional coefficient $\widehat{\beta}(t)$ for family structure trajectories. The dark and light shaded areas represent the 90% and 95% confidence bands, respectively. The upper right panel displays the heat map of the estimated interaction function $\widehat{\Gamma}(t, s)$ between family income and family structure trajectories. The lower-right panel displays the corresponding heat map, retaining only the grid cells with values significantly different from zero at the 5% significance level.

References

Matzkin, R. (1992). Nonparametric and distribution-free estimation of the binary threshold crossing and the binary choice models. *Econometrica*, 60(2):239–270.

Yan, G. (2024). A kernelization-based approach to nonparametric binary choice models. Manuscript.