Earnings Dynamics and Intergenerational Transmission of Skill*

Lance Lochner
University of Western Ontario

Youngmin Park
Bank of Canada

August 25, 2021

Abstract

This paper develops and estimates a two-factor model of intergenerational skill transmission when earnings inequality reflects differences in individual skills and other non-skill shocks. We consider heterogeneity in both initial skills and skill growth rates, allowing variation in skill growth to change over the lifecycle. Using administrative tax data on two linked generations of Canadians covering 37 years, we exploit covariances in log earnings (at different ages) both across and within generations to identify and estimate the intergenerational correlation structure for initial skills and skill growth rates, lifecycle skill growth profiles, and the dynamics of non-skill earnings shocks.

We estimate low intergenerational elasticities (IGEs) for earnings in Canada; however, skill IGEs are typically 2–3 times larger due to considerable (and persistent) variation in earnings conditional on skills. Both earnings and skill IGEs decline substantially for more recent cohorts and are lower for children born to younger fathers. Intergenerational transmission of both initial skills and skill growth rates explains up to 40% of children's skill variation. Skills become a more important determinant of earnings over the first part of workers' careers, while intergenerational transmission of skills becomes less important as children age. Parents' initial skills and skill growth rates are equally important determinants of children's skills, largely because both strongly influence children's initial skills.

Finally, we study intergenerational mobility for the 35 largest cities in Canada, documenting the extent to which considerable differences in earnings and skill IGEs vary with the extent of local heterogeneity in parental skills vs. earnings instability.

*For their valuable comments, we thank seminar/conference participants at *From Theory to Statistics to Empirics: An Econometrics Conference in Honor of James Heckman* at the University of Chicago, the Bank of Canada, Georgetown University, 2019 Annual Meeting of the Canadian Economic Association, 2019 Summer Meeting of the North American Econometric Society, Federal Reserve Bank of Minneapolis, 2019 Korean Economic Review International Conference, Statistics Canada, University of California–Riverside, Jinan University, and Seoul National University. We also thank Yuri Ostrovsky and Winnie Chan of Statistics Canada for their help with the data. The views expressed in this paper are those of the authors and do not necessarily represent those of the Bank of Canada.

1 Introduction

Intergenerational mobility is a critical component of social justice and an indicator of economic efficiency. As such, an extensive literature is devoted to the underlying theory of intergenerational transmission (e.g., Becker and Tomes, 1979, 1986; Loury, 1981; Cunha et al., 2006; Cunha and Heckman, 2007), its measurement (e.g., Jenkins, 1987; Solon, 1992; Chetty et al., 2014b; Landersø and Heckman, 2017), and the intersection between theory and empirics (e.g., Cunha, 2013; Del Boca, Flinn, and Wiswall, 2014; Cunha, Heckman, and Schennach, 2010; Gayle, Golan, and Soytas, 2015; Lee and Seshadri, 2019; Caucutt and Lochner, 2020).

Of course, it is no surprise that James Heckman has made pioneering contributions to all three of these literatures, connecting theory, statistics, and empirics as he has done throughout his career. His work has also contributed greatly to our understanding of skill formation over the lifecycle (e.g., Heckman, 1976; Heckman, Lochner, and Taber, 1998) and dynamic multi-factor models of earnings (e.g., Carneiro, Hansen, and Heckman, 2003; Cunha, Heckman, and Navarro, 2005; Heckman, Stixrud, and Urzua, 2006; Cunha and Heckman, 2008; Cunha, Heckman, and Schennach, 2010), two related literatures that intersect with the contributions of this paper.

This paper fits most tightly within the literature on measurement of intergenerational mobility, which has emphasized summary measures such as the intergenerational elasticity (IGE) of earnings or rank-rank slopes.[1] Motivated by the theoretical literature, this measurement literature has largely interpreted intergenerational earnings relationships through the lens of intergenerational skill transmission; yet, this link is rarely made explicit in the estimating equations or measurements. Instead, the literature has almost exclusively aimed to quantify intergenerational mobility in lifetime income, as a measure most closely related to economic welfare.[2] Early on, Solon (1992) and Zimmerman (1992) noted that year-to-year fluctuations in earnings would lead to attenuation bias (for the lifetime earnings IGE) in estimates based on annual measures of earnings. Jenkins (1987) and, more recently, Haider and Solon (2006) further argued that the correlation between earnings at any age and lifetime earnings varies systematically over the lifecycle, generating a different form of "lifecycle bias" (for the lifetime earnings IGE).[3] While Haider and Solon (2006) motivate this age-varying correlation between annual earnings and lifetime income based on a two-factor model of lifecycle earnings, the earnings IGE literature has remained focused on a single-factor model of skill transmission.

In this paper, we explicitly model the intergenerational transmission of two heterogeneous skill factors, *initial skills* and *skill growth rates*, while also accounting for both transitory and persistent non-skill earnings shocks (assumed to be uncorrelated across generations). These two forms of heterogeneity are motivated by human capital theory (Ben-Porath, 1967; Becker, 1975; Shaw, 1989)

[1]The earnings IGE is defined as the slope coefficient when regressing log child earnings on log parental earnings, while the rank-rank slope is defined as the slope coefficient when regressing the child's percentile (in their generation's earnings distribution) on the parent's percentile (within their generation's earnings distribution).

[2]See Gallipoli, Low, and Mitra (2020) for an analysis linking intergenerational earnings (and other income) relationships to intergenerational consumption persistence.

[3]It has become standard in the literature to address this concern by using measures of earnings at the age at which differences in annual earnings are thought to reflect differences in lifetime earnings. Because identifying this age requires earnings data over the full lifecycle, which is not widely available, IGE studies often rely on external evidence for this critical age. Unfortunately, the critical age may vary across time and populations in unknown ways.

and have found mixed support in the empirical literature on lifecycle skill formation and earnings dynamics (e.g., Lillard and Weiss, 1979; Hause, 1980; MaCurdy, 1982; Baker, 1997; Heckman, Lochner, and Taber, 1998; Haider, 2001; Guvenen, 2009; Hryshko, 2012; Lochner, Park, and Shin, 2018).

As in much of the human capital literature, we do not directly measure skills but use lifecycle earnings patterns to infer the evolution of skill distributions separately from the distributions of idiosyncratic labor market shocks. More precisely, our 2 skill factors imply heterogeneous lifecycle trajectories in expected earnings, which we loosely refer to as "skills", while we refer to deviations around these trajectories as "non-skill" shocks. This terminology is irrelevant for our model of intergenerational transmission and the importance of both skill factors for earnings mobility and inequality; however, it is important for our distinction between skill and earnings IGEs discussed further below. The simple dichotomy attributes to skills the roles of any personal characteristics determined early in workers' careers (e.g., education, abilities, personality traits), whose effects may vary across individuals and over the lifecycle, as well as other systematic differences across workers like local wage rates or persistent occupation/career choices. For reasons discussed in Section 2, it seems likely that most of the variation we attribute to skills reflects differences attributable to education and ability. Variability in earnings due to such factors as search and matching frictions, local economic shocks to wages, and sudden changes in health are, by assumption, considered unrelated to skills.[4]

Using 37 years (1978–2014) of administrative tax data from Canada on both fathers and sons, we show that intergenerational earnings covariance patterns (across different ages for fathers and sons) are consistent with a two-factor model of intergenerational skill transmission. Taking a minimum distance approach, we estimate our model by fitting both intergenerational covariances and intragenerational autocovariances for earnings. Our estimates reveal considerable heterogeneity in both initial skill levels and skill growth rates, with differences in skill growth rates declining with age (as human capital theory predicts). The lack of any covariance between the earnings growth of fathers and sons in our data implies no intergenerational correlation in skill growth rates; however, we estimate intergenerational correlations in initial skills as high as 0.3 for earlier cohorts in our sample, with that correlation declining by two-thirds for more recent cohorts. We also estimate correlations between the initial skills of one generation and the skill growth rates of the other that range between 0.14 and 0.26.

We distinguish between the IGE for skills and the more standard IGE for earnings, with the former always larger than the latter due to non-skill earnings shocks. We confirm the low earnings IGEs in Canada estimated by Corak and Heisz (1999) and Chen, Ostrovsky, and Piraino (2017), even when earnings are averaged over many years for both fathers and sons. We also estimate considerable variation in earnings conditional on skill levels, so skill IGEs are typically 2–3 times earnings IGEs, even when averaging across 5 or 9 years. Both earnings and skill IGEs have declined considerably over time, suggesting significant improvements in intergenerational mobility in skills and earnings in Canada. Among recent cohorts of sons, we also observe that both earnings and skill IGEs are

[4] See, e.g., Postel-Vinay and Turon (2010), Bowlus and Liu (2013), Bagger et al. (2014) and Karahan, Ozkan, and Song (2019) for the development and estimation of formal models including both human capital accumulation and search frictions.

significantly greater for those with younger fathers. This suggests that it is important to account for parental age differences (at child's birth) when comparing mobility across cohorts.

We next consider an alternative summary measure of intergenerational mobility that quantifies the extent to which intergenerational transmission explains variation in the skills or earnings of children. Specifically, we estimate the share of children's skill or earnings variance that is explained by their projected skills (i.e., the linear projection of their skills onto parental initial skills and skill growth factors). Among the earliest cohorts of children we study, projected skills explain roughly 40% of skill variation; however, this drops to around 20% for more recent cohorts. This share is also larger for children with younger parents. The cross-cohort patterns for this measure are qualitatively similar to those for earnings and skill IGEs.

Our two-factor model of skill transmission produces several novel lessons. For example, we show that skills become a more important component of earnings over the first 10–15 years of individuals' careers, but intergenerational transmission accounts for less and less of the variation in skills over these ages. While variation in projected skills grows with age, it grows at a slower rate than variation in skills themselves. Variance decompositions further show that while differences in initial skills are an important source of skill variation throughout the lifecycle, heterogeneity in skill growth becomes the dominating source of skill variation after only 5–10 years in the labor market. By contrast, the transmission of children's initial skills (compared to their skill growth rates) is a more important source of skill variation throughout life. Yet, skill growth heterogeneity is still an important component of intergenerational transmission, since parental skill growth is positively correlated with children's initial skill levels. Indeed, fathers' initial skills and skill growth rates contribute equally to the variance of sons' skills and earnings variation throughout their careers. Finally, we show that knowledge of fathers' average lifetime skill levels (but not their initial skills and growth rates) explains at most two-thirds of all projected skill variation among their sons, highlighting the importance of accounting for the transmission of both skill factors. Even a father's expected lifetime earnings — the emphasis of much of the literature — provides only a limited source of information about his son's earnings unless the lifecycle trajectory of the father's earnings is taken into account.

Given the recent interest in regional variation in intergenerational mobility (e.g., Chetty et al., 2014a; Connolly, Corak, and Haeck, 2019; Corak, 2019), we also use our data to study cross-city variation in earnings and skill IGEs for the 35 largest cities in Canada. We focus on the relationship between city-level measures of earnings/skill IGEs and parental earnings/skill inequality to shed new light on the underlying forces contributing to the so-called "Great Gatsby Curve" — the positive correlation between intergenerational persistence, typically measured by earnings IGEs or rank-rank slopes, and measures of cross-sectional inequality (Corak, 2013). We explore two potential explanations for the Great Gatsby Curve, based on our model of skill transmission and earnings. First, areas with greater earnings inequality may have stronger IGEs in skills and, therefore, in earnings. Second, areas with greater earnings inequality may have less earnings instability conditional on skills (i.e., skills may explain a larger share of the earnings variation). Much of the literature implicitly assumes the former, searching for explanations based on differences in schooling, families, or neighborhoods/peers that might explain stronger or weaker intergenerational skill transmission. The latter explanation might instead signal differences in the structure and flexibility of labor markets (e.g., minimum wages, unionization, unemployment policies). See Landersø and Heckman (2017) for a careful analysis of these issues in a comparison of Denmark and the U.S. showing that intergenerational income and earnings mobility is stronger in Denmark, while educational mobility rates are much more similar in the two countries.

Consistent with prior studies for Canada (Connolly, Corak, and Haeck, 2019; Corak, 2019) and the Great Gatsby Curve, we estimate higher earnings IGEs for cities with greater parental earnings inequality. More novel, we document substantial variation in the skill share of the earnings variance across cities. Because this share is (weakly) increasing in parental inequality, the relationship between skill IGEs and parental earnings inequality is weaker than that observed between earnings IGEs and inequality. Most interestingly, we find that the skill IGE is decreasing in the variance of parental skills but increasing in the variance of the non-skill component of earnings. Thus, cities with greater skill heterogeneity exhibit greater intergenerational skill mobility, while skill mobility is lower in cities with greater earnings instability.

The rest of this paper is organized as follows. Section 2 develops our model of lifecycle earnings and intergenerational skill transmission. Section 3 describes the data used in our empirical analysis and highlights key features of the data relevant to intergenerational transmission and the evolution of skill. Our estimation approach is described in Section 4. Section 5 reports our findings on intergenerational skill and earnings mobility for several different father-son cohort pairs, using data for all of Canada, while Section 6 discusses our analysis of city-level variation in earnings and skill IGEs. Section 7 concludes.

2 Earnings and Intergenerational Skill Transmission

In this section, we describe our two-factor intergenerational model of lifecycle skills and earnings. We begin by decomposing earnings into skill and non-skill components, followed by a discussion of their evolution over the lifecycle within each generation. Our approach considers two heterogeneous latent skill factors — initial skill level and a skill growth factor — that determine individual skill accumulation and the distribution of skills over the lifecycle. Transitory and persistent non-skill shocks also influence the dynamics of earnings. Finally, we discuss the intergenerational transmission of initial skills and skill growth, considering implications for intergenerational earnings and skill relationships over the lifecycle.

2.1 Skill and Earnings over the Lifecycle

Suppose that there are a large number of families (i.e., parent-child pairs) indexed by $i = 1, \ldots, N$. Let $j \in \{p, k\}$ be the index for family member (p for the parent and k for the kid), or "generation." Let $Y_{i,j,t}$ be the log earnings of family member j in family i at age t, and let $y_{i,j,t}$ be the log-earnings residual obtained by subtracting off average log earnings conditional on age, year, generation, and family cohort group (defined in Section 3). Although all of our analysis is conditional on birth cohorts of both parents and children, we keep the conditioning implicit here, essentially focusing on a particular cohort of parents and children. Our focus on within-cohort inequality and skill transmission abstracts from systematic differences in the level of skills across cohorts and across different ages.[5]

Consider a basic decomposition of the log-earnings residual:

$$y_{i,j,t} = \theta_{i,j,t} + \varepsilon_{i,j,t},$$

where $\theta_{i,j,t}$ reflects the individual's "unobserved skill" and $\varepsilon_{i,j,t}$ reflects the "non-skill" component at age t. We normalize both components so they have mean zero and are orthogonal to each other:[6]

$$\mathrm{Cov}(\theta_{j,t}, \varepsilon_{j',t'}) = 0, \quad \forall (j, t, j', t').$$

We assume that the non-skill components are uncorrelated across generations:

$$\mathrm{Cov}(\varepsilon_{j,t}, \varepsilon_{j',t'}) = 0, \quad \forall (j, t, j', t') \text{ such that } j \neq j'.$$

From a practical standpoint, this labels as skills any part of earnings that is (potentially) intergenerationally correlated. We assume independence of all components across families.

We consider a two-factor model for skill accumulation: $(\psi_{i,j}, \delta_{i,j})$. The first factor, $\psi_{i,j}$, reflects the *initial skill* for individual (i, j) at age $\underline{t}$, the beginning of his career. Thus, $\theta_{i,j,t} = \psi_{i,j}$ for $t = \underline{t}$. The second factor, $\delta_{i,j}$, represents a *skill growth factor*. For $t > \underline{t}$, we assume that skill growth is given by

$$\Delta\theta_{i,j,t} \coloneqq \theta_{i,j,t} - \theta_{i,j,t-1} = \lambda_{j,t}\delta_{i,j},$$

where $\delta_{i,j}$ determines the individual-specific skill growth rate, which can vary systematically with age according to $\lambda_{j,t}$.

We normalize $\lambda_{j,\underline{t}+1} = 1$, but otherwise require no restrictions (not even a sign restriction) on $\lambda_{j,t}$ for $t > \underline{t} + 1$. This normalization sets both the sign and scale of $\delta_{i,j}$ to that of the initial expected earnings growth rate for individual (i, j). For future reference, let $\Lambda_{j,t}$ be the cumulative sum of $\lambda_{j,t}$:

$$\Lambda_{j,t} \coloneqq \sum_{t'=\underline{t}+1}^{t} \lambda_{j,t'}, \qquad \text{for } t \geq \underline{t} + 1,$$

with $\Lambda_{j,\underline{t}} \coloneqq 0$. Then, an individual's skill level can be written in the following two-factor form:

$$\theta_{i,j,t} = \psi_{i,j} + \Lambda_{j,t}\delta_{i,j}, \qquad \text{for } t \geq \underline{t},$$

This general skill process nests the "heterogeneous income profiles" (HIP) framework, widely

[5]By using residual earnings, we remove average differences over time. These time effects would reflect both differences in average skill levels and aggregate skill price (or wage) movements, which cannot be separately identified without direct measures of skill.

[6]Let $x_{j,t}$ be a random variable and its realization for family i be $x_{i,j,t}$. Denote its cross-sectional second moments by $\mathrm{Var}(x_{j,t})$ and $\mathrm{Cov}(x_{j,t}, x_{j',t'})$.

used in the literature on earnings dynamics (e.g., Lillard and Weiss, 1979; Hause, 1980; Baker, 1997; Guvenen, 2007), which assumes $\lambda_{j,t} = 1$ for all (j, t). Our "generalized HIP" process allows for heterogeneous skill growth to vary over the lifecycle, accounting for the well-known decline in on-the-job skill investments with age that is emphasized in human capital theory (Ben-Porath, 1967; Mincer, 1974; Becker, 1975). We also allow skill growth heterogeneity to vary across generations, which may reflect different market incentives for skill accumulation over time as highlighted in Heckman, Lochner, and Taber (1998).[7]

Because we do not observe education in our data, heterogeneity in $(\psi_{i,j}, \delta_{i,j})$ and, consequently, $\theta_{i,j,t}$ will reflect any systematic effects of education on skill levels and growth rates, as well as variation in skills conditional on schooling (see, e.g., Heckman, Lochner, and Todd, 2006, 2008). Given our approach, differences in what we refer to as skills (and skill factors) will also include the influence of other lasting differences across individuals (e.g., choice of occupation or career path) and differences in local wage rates.[8] To this end, it would be more precise to refer to $\theta_{i,j,t}$ as expected age t earnings conditional on initial endowments $(\psi_{i,j}, \delta_{i,j})$. We, therefore, use the term *skills* loosely; however, there are good reasons to think that variation in $\theta_{i,j,t}$ primarily reflects variation in actual skills (and/or ability): (i) the fraction of log earnings explained by $\theta_{i,j,t}$ is broadly consistent with the share of variation in log earnings explained by schooling, experience, and various measures of skill/ability; and (ii) our estimates of the intergenerational skill elasticity are similar to its analogue based on educational attainment.[9]

We assume that the non-skill component of earnings consists of a persistent component that follows an AR(1) process and a transitory component that follows an MA(1) process:[10]

$$\begin{aligned}
\varepsilon_{i,j,t} &= \phi_{i,j,t} + (\xi_{i,j,t} + \kappa_j \xi_{i,j,t-1}), \\
\phi_{i,j,t} &= \rho_j \phi_{i,j,t-1} + \nu_{i,j,t}.
\end{aligned}$$

This general process is commonly employed in models of earnings dynamics (e.g., see the survey by Meghir and Pistaferri, 2011), where $\varepsilon_{i,j,t}$ reflects the influence of measurement error in earnings, shocks to earnings due to job-search or -matching frictions, or other shocks to earnings conditional

[7] We note that our factors $(\psi_{i,j}, \delta_{i,j})$ and skill growth sequence $\lambda_{j,t}$ may reflect the outcomes of endogenous human capital investments as functions of more primitive parameters like the "ability to learn" (Ben-Porath, 1967; Heckman, Lochner, and Taber, 1998), which may be heterogeneous within and across generations. Our skill measure $\theta_{i,j,t}$ would reflect skills devoted towards production in these on-the-job investment frameworks.

[8] Differences in local wage rates are unlikely to play an important role in our analysis, since we estimate a similar role for skills when examining intergenerational mobility within cities (see Section 6).

[9] Regarding (i), we show below that $\mathrm{Var}(\theta_{k,t})/\mathrm{Var}(y_{k,t})$ rises from about 25% at age 30 to roughly 40% over ages 35–45. Looking at American white males in their late-20s and early-30s, Cawley, Heckman, and Vytlacil (1999) obtain an R^2 of 0.25 when regressing log wages on potential experience, 10 principal components from a battery of cognitive test measures, and educational attainment. Deming (2017) estimates R^2 values of 0.30–0.36 (for all Americans) when controlling for measures of cognitive, non-cognitive, and social skills, as well as schooling attainment. Regarding (ii), Aydemir, Chen, and Corak (2013) estimate that a regression of Canadian men's schooling (from cohorts born 1964–1976) on their fathers' schooling yields a coefficient of about 0.33. If $\theta_{i,j,t}$ is linear in years of schooling (as implied by Mincerian log earnings regressions), then this should be (and is) similar to our estimated skill IGEs, which range from 0.25 to 0.45 for the same cohorts.

[10] Appendix A discusses identification for the more general case with an MA(q) transitory process; however, we assume the MA(1) transitory process in estimation, since it fits the data quite well.

on skills (e.g., illness, firm learning about worker productivity, or firm-specific shocks). We allow the variance of non-skill earnings innovations $(\xi_{j,t}, \nu_{j,t})$ to vary freely across cohorts and age/time and the dynamics of these components (ρ_j, κ_j) to differ across cohorts.

While the assumption that transitory earnings shocks are unrelated to worker skills is quite natural, more persistent shocks could be, at least partially, skill-related. Our estimates for ρ_j suggest that persistent shocks have a half-life of about 7 years for parents and 4 years for their children. To the extent that they do reflect random skill innovations, they fade out relatively quickly in the context of lifetime careers. Most important for our analysis is the assumption that non-skill shocks are uncorrelated with the skills and non-skill shocks of other generations. Consistent with this assumption, Gallipoli, Low, and Mitra (2020) estimate insignificant effects of permanent shocks to parental earnings on the earnings shocks of their children. As noted earlier, our notion of skills reflects the expected earnings of individuals conditional on their age and skill factors $(\psi_{i,j}, \delta_{i,j})$.

Empirically, we consider earnings profiles beginning at age $\underline{t} = 26$, so the person-specific skill factors may reflect the influence of family and educational investments up to that age. We next turn to the intergenerational transmission of these skill factors, reflecting both biological inheritance as well as endogenous family and school investments up to age $\underline{t}$.[11]

2.2 Intergenerational Transmission of Skill

In our framework, the intergenerational transmission of skill (and earnings) comes entirely from the intergenerational transmission of initial skill and the skill growth factor. We study intergenerational relationships without imposing restrictions on the joint distribution of parents' and children's skill factors $(\psi_p, \delta_p, \psi_k, \delta_k)$.

2.2.1 Intergenerational Elasticities

We begin by considering intergenerational elasticities (IGEs), the most common measure of intergenerational transmission in the literature. The earnings IGE is typically obtained from a regression of children's log earnings (at age t) on parental log earnings (at age t'):

$$\mathrm{IGE}_{y,t,t'} := \frac{\mathrm{Cov}(y_{k,t}, y_{p,t'})}{\mathrm{Var}(y_{p,t'})} = \frac{\mathrm{Cov}(\theta_{k,t}, \theta_{p,t'})}{\mathrm{Var}(y_{p,t'})} = s_{p,t'} \cdot \mathrm{IGE}_{\theta,t,t'}, \tag{1}$$

where

$$s_{j,t} := \frac{\mathrm{Var}(\theta_{j,t})}{\mathrm{Var}(y_{j,t})} = \frac{\mathrm{Var}(\theta_{j,t})}{\mathrm{Var}(\theta_{j,t}) + \mathrm{Var}(\varepsilon_{j,t})} \tag{2}$$

can be defined generally as the skill share of the earnings variance for generation j at age t and

$$\mathrm{IGE}_{\theta,t,t'} := \mathrm{Cov}(\theta_{k,t}, \theta_{p,t'})/\mathrm{Var}(\theta_{p,t'})$$

[11] We do not attempt to distinguish the roles of "nature" vs. "nurture". For fully specified models considering intergenerational transmission with endogenous skill investments in children, see Becker and Tomes (1986); Cunha et al. (2006); Cunha and Heckman (2007); Del Boca, Flinn, and Wiswall (2014); Gayle, Golan, and Soytas (2015); Lee and Seshadri (2019); Caucutt and Lochner (2020).

represents the analogous skill IGE. These earnings and skill IGEs, as well as the skill share of the earnings variance, can be similarly defined for averages of earnings or skills over any ages.

Equation (1) makes clear that the earnings IGE understates the skill IGE for the same ages due to variation in the non-skill component of parental earnings.[12] Intuitively, skills are more strongly correlated than earnings across generations, because earnings are like "noisy" measures of skills. The literature on earnings IGEs sometimes refers to this as a standard "attenuation bias" associated with mismeasured regressors (for early analyses, see Solon, 1992; Zimmerman, 1992); however, studies typically focus on the mismeasurement of "average lifetime earnings," which differs from skills at any given age.[13] It is worth further noting that average realized lifetime earnings also differ from average lifetime skills, since the average of all non-skill earnings shocks is not generally zero for finite careers. Indeed, this discrepancy could be sizeable when persistent shocks are an important determinant of earnings (e.g., one worker may luck out and find a high-paying career job right out of school, while an identical worker may end up cycling through a series of lower-paying temporary jobs before settling into something more stable and higher paying).[14] In our framework, the skill IGE reflects the intergenerational elasticity of the predictable components of earnings as determined by the transmission of skill factors (ψ_j, δ_j).

2.2.2 Projecting Children's Skill on Their Parents' Skill

We now consider an alternative way to quantify the importance of intergenerational skill transmission for skill (or earnings) inequality: measuring the share of children's skill (or earnings) variation that can be explained by variation in their skills as predicted by parental skill factors (ψ_p, δ_p).

Let $\hat{\psi}_k$ and $\hat{\delta}_k$ be "projected skill factors" — the children's skill factors that are linearly projected on their parents' skill factors:

$$\begin{aligned}\hat{\psi}_k &\coloneqq \mathrm{P}[\psi_k|\psi_p, \delta_p] = \alpha_{\psi,\psi}\psi_p + \alpha_{\delta,\psi}\delta_p \\ \hat{\delta}_k &\coloneqq \mathrm{P}[\delta_k|\psi_p, \delta_p] = \alpha_{\psi,\delta}\psi_p + \alpha_{\delta,\delta}\delta_p,\end{aligned}$$

where $\mathrm{P}[y|x]$ denotes the linear projection of y onto x and the linear-projection coefficients are given

[12]The same is true of standard intergenerational correlations, where the discrepancy depends on the skill share of the earnings variance for both generations.

[13]The distinction between expected earnings at a given age (like our measure of skill) and "average lifetime earnings" leads to what is sometimes referred to as "lifecycle bias" in the literature (Jenkins, 1987; Mazumder, 2005; Grawe, 2006; Haider and Solon, 2006; Nybom and Stuhler, 2016). As is further discussed below, this lifecycle bias has been characterized within a framework assuming that intergenerational relationships are explained by a single factor.

[14]Of course, if persistent shocks were thought to be skill-related, then the earnings IGE would more closely reflect the skill IGE (especially when averaged over several years), since the only discrepancy would be due to transitory shocks. Appendix Figure H13 shows that the annual share of the earnings variance explained by the transitory component alone is around one-third for parents in their 40s and early 50s. This means that even when treating persistent shocks as a component of skills, annual skill IGEs are about 50% larger than annual earnings IGEs.

by

$$\underbrace{\begin{bmatrix} \alpha_{\psi,\psi} & \alpha_{\psi,\delta} \\ \alpha_{\delta,\psi} & \alpha_{\delta,\delta} \end{bmatrix}}_{:=\mathbf{A}} := \underbrace{\begin{bmatrix} \mathrm{Var}(\psi_p) & \mathrm{Cov}(\psi_p, \delta_p) \\ \mathrm{Cov}(\psi_p, \delta_p) & \mathrm{Var}(\delta_p) \end{bmatrix}^{-1}}_{:=\mathbf{\Omega}_p} \underbrace{\begin{bmatrix} \mathrm{Cov}(\psi_p, \psi_k) & \mathrm{Cov}(\psi_p, \delta_k) \\ \mathrm{Cov}(\delta_p, \psi_k) & \mathrm{Cov}(\delta_p, \delta_k) \end{bmatrix}}_{:=\mathbf{\Omega}_{p,k}}. \tag{3}$$

While these linear projections are well-defined regardless of the distribution for $(\psi_p, \delta_p, \psi_k, \delta_k)$, if these skill factors are joint normally distributed, then the projections would reflect conditional expectations.

Next, we can define "projected skill" at age t:

$$\hat{\theta}_{k,t} := \hat{\psi}_k + \Lambda_{k,t}\hat{\delta}_k. \tag{4}$$

With this, we can define the fraction of skill variance explained by projected skills:

$$\frac{\mathrm{Var}(\hat{\theta}_{k,t})}{\mathrm{Var}(\theta_{k,t})} = \frac{\mathrm{Var}(\hat{\psi}_k + \Lambda_{k,t}\hat{\delta}_k)}{\mathrm{Var}(\psi_k + \Lambda_{k,t}\delta_k)}. \tag{5}$$

This provides a measure of the extent to which intergenerational skill transmission explains cross-sectional inequality in skills. One can similarly determine the importance of intergenerational skill transmission for earnings inequality from the share of earnings variance that is explained by projected skills:

$$\frac{\mathrm{Var}(\hat{\theta}_{k,t})}{\mathrm{Var}(y_{k,t})} = \frac{\mathrm{Var}(\hat{\theta}_{k,t})}{\mathrm{Var}(\theta_{k,t})} s_{k,t}, \tag{6}$$

where $s_{k,t}$ is the share of children's earnings explained by skills (see equation (2)). Intergenerational transmission explains a smaller share of earnings variation than skill variation due to the non-skill earnings shocks children experience throughout their lives.[15]

Finally, notice that we can substitute in for projected skill factors $(\hat{\psi}_k, \hat{\delta}_k)$ in equation (4) to write projected skills as linear functions of parental skill factors:

$$\hat{\theta}_{k,t} = \underbrace{(\alpha_{\psi,\psi} + \alpha_{\psi,\delta}\Lambda_{k,t})}_{:=\beta_{\psi,t}} \psi_p + \underbrace{(\alpha_{\delta,\psi} + \alpha_{\delta,\delta}\Lambda_{k,t})}_{:=\beta_{\delta,t}} \delta_p. \tag{7}$$

Thus, equations (4) and (7) can be used for two distinct projected skill variance decompositions:

$$\begin{aligned} \mathrm{Var}(\hat{\theta}_{k,t}) &= \mathrm{Var}(\hat{\psi}_k) + \Lambda_{k,t}^2 \mathrm{Var}(\hat{\delta}_k) + 2\Lambda_{k,t} \mathrm{Cov}(\hat{\psi}_k, \hat{\delta}_k) && (8) \\ &= \beta_{\psi,t}^2 \mathrm{Var}(\psi_p) + \beta_{\delta,t}^2 \mathrm{Var}(\delta_p) + 2\beta_{\psi,t}\beta_{\delta,t} \mathrm{Cov}(\psi_p, \delta_p). && (9) \end{aligned}$$

The first decomposition is informative about the relative importance of intergenerational transmission through children's projected initial skill vs. skill growth factors, while the second decomposition clarifies the relative importance of each parental skill factor for children's skills. Since parental

[15]We note that the share of earnings variance explained by skill transmission does not depend on whether persistent shocks, $\phi_{i,k,t}$, are considered part of skills, since these are not correlated with any components of parental earnings.

skill factors only influence children's earnings through their skills, these decompositions are also informative about the roles of intergenerational transmission in initial skills and skill growth for earnings inequality.

3 Canada's Intergenerational Income Database

We use tax records on the earnings of Canadian parents and their children from the Intergenerational Income Database (IID).[16] The IID consists of children ages 16–19, in one of the child "cohort years"—1982, 1984, 1986, 1991, 1996, and 2001—who are linked to their parents. Therefore, the children are born in years 1963–1985, excluding 1971, 1976, and 1981. The child-parent linkage is created based on the names and addresses found on individual T1 tax files within four years of the cohort years. For example, those who were 19 years old in 1982 are linked to their parents if they and their parents filed their taxes between 1982 and 1986 under the same last name and address. Therefore, the IID does not include children who did not file their taxes or did not live with their parents during their late teens or early 20s. The IID covers about 70% of the population, and those who are not included in the data tend to be children from low-income families (Statistics Canada, 2017). We use sample weights constructed by Statistics Canada to address sample representation issues.

For linked child-parent pairs, each of their annual earnings are obtained from tax files. Earnings are measured as a sum of the employment income reported on T4 slips (issued by employers) and other employment income. For children ages 16–19 in the first three cohort years and their parents, called "Panel A" families, earnings are observed from 1978 to 2014. The rest of the families are called "Panel B" families, and their earnings are observed from 1981 to 2014.

We study the earnings of eldest sons and their fathers over ages 26–55, restricting our sample to fathers ages 16–50 when the eldest son was born.[17] Because we do not observe earnings before 1978, fathers older than 55 in 1978 (born before 1923) are excluded from our sample. Birth years are more dispersed among parents than among children due to variation in parents' ages when their children were born. Because we are interested in the trends of intergenerational mobility across birth cohorts, we split the sample based on birth years of both parents and children.

First, as Table 1 shows, each generation of parents and children is divided into 5 "cohort groups." The cohort grouping is based on roughly 5-year birth cohorts for children and 10-year birth cohorts for parents. Next, by combining the cohort groups of both parents and children, we create ten "family cohort groups," as illustrated in Table 2. Due to the limited variation in parental age at birth conditional on the birth year of children, each of the children's cohort groups is linked to only two cohort groups of parents. Because the range of years covered is fixed, we do not observe earnings during the early part of the lifecycle for older cohorts, while the opposite is true for more recent cohorts. This means that we generally observe children's earnings when they are young and parents' earnings when they are old.

[16]The IID was previously used by Corak and Heisz (1999) and Chen, Ostrovsky, and Piraino (2017), and a detailed description can be found in Statistics Canada (2017).

[17]More precisely, we use the oldest son (for each family) that is observed in our data, which does not include children born prior to 1963.

Panel	Observed Years	Children			Parents		
		Cohort Group	Birth Years	Observed Ages	Cohort Group	Birth Years	Observed Ages
A	1978–2014	1965	1963–1966	26–48	1930	1923–1933	55
		1970	1967–1970	26–44	1940	1934–1944	44–55
B	1981–2014	1975	1972–1975	26–39	1940	1934–1944	47–55
		1980	1977–1980	26–34	1950	1945–1955	36–55
		1985	1982–1985	26–29	1960	1956–1966	26–48

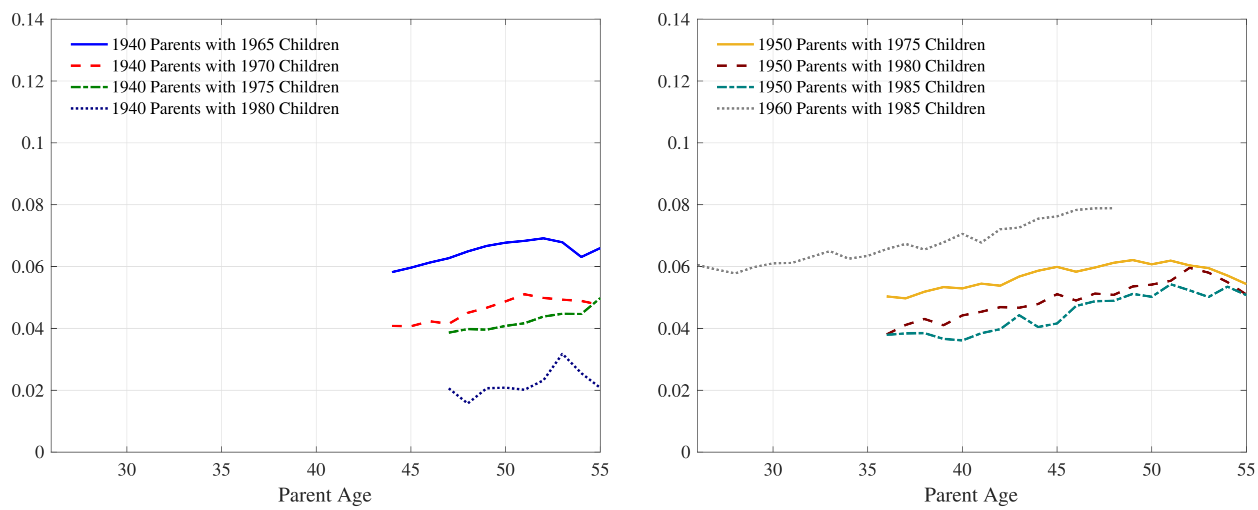
Table 1: Cohort Grouping For Each Generation

Family Cohort Group		
Family	Children	Parents
1	1965	1930
2		1940
3	1970	1930
4		1940
5	1975	1940
6		1950
7	1980	1940
8		1950
9	1985	1950
10		1960

Table 2: Family Cohort Grouping

Our empirical analysis focuses on log earnings residuals $y_{i,j,t}$ and their various covariances. Within generation, age, year, and family cohort group, we first drop the top and bottom 1% of strictly positive earnings and then subtract average log earnings (within generation, age, year, and family cohort group) in order to calculate log earnings residuals. We calculate two sets of covariances between log earnings residuals: (i) "intragenerational" covariances between one's own log earnings residuals from different ages (i.e., autocovariances) and (ii) "intergenerational" covariances between parents' and children's log earnings residuals. Across all family cohort groups, there are 1,154 autocovariances for parents, 1,252 autocovariances for children, and 1,385 intergenerational covariances. The weighted number of families contributing to each covariance ranges from 35,650 to 272,290, with an average of 153,880.[18]

Autocovariances are commonly exploited to identify the earnings process for a single generation (e.g., Lillard and Weiss, 1979).[19] Although the intergenerational mobility literature often focuses directly on the earnings IGE, it rarely discusses intergenerational covariances themselves. Yet, these covariances are a transparent source of identification in our two-factor model of intergenerational transmission.



(a) 1940 Parents

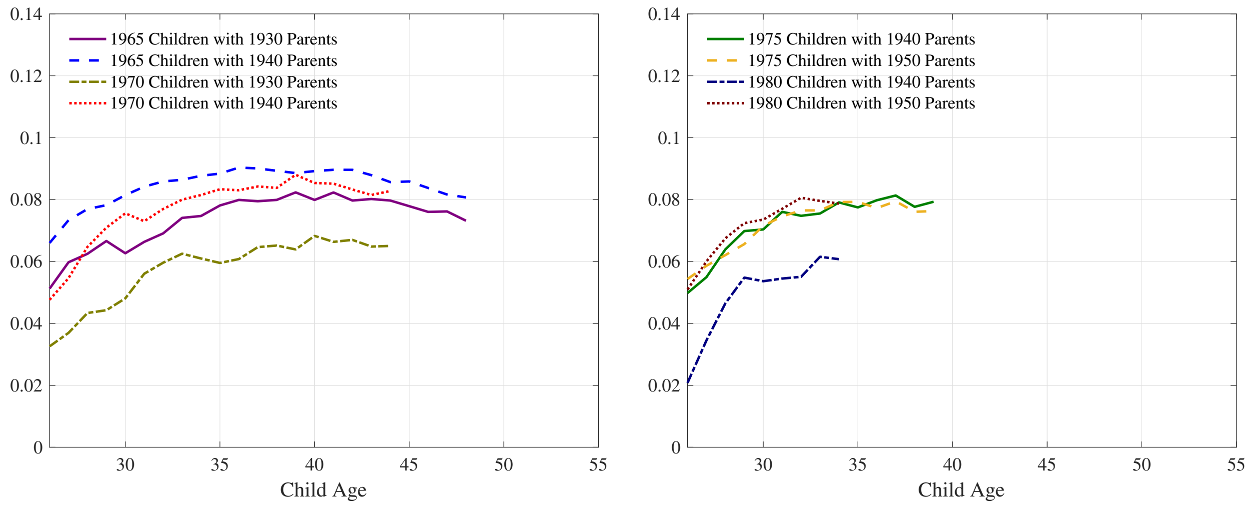
(b) 1950 and 1960 Parents

Figure 1: Intergenerational Covariances of Log Earnings by Parents' Age (at Children's Age 26)

Figure 1 reports the covariances between parents' and children's log earnings when the child's age is held constant at age 26, where each line corresponds to a different family cohort group. We see that within each family cohort group, the intergenerational covariances are positive and increase with the parents' age until they flatten out and eventually decrease after age 50. Moreover, the covariances are lower for families with more recent cohorts of children conditional on the parents' cohort group, while they are higher for families with the more recent cohort of parents within the children's cohort

[18] We only consider families that contribute to at least one intergenerational covariance. The weighted number of observations must be rounded to the nearest 10 due to confidentiality. Although the number of unweighted observations cannot be revealed, it is not far from the number of weighted observations due to the high coverage rate.

[19] See Appendix Figures B1 and B2 for selected intragenerational variances and autocovariances. The sharp decline in autocovariances for one lag and the roughly exponential decline for longer lags motivates our assumed process for earnings shocks with AR(1) and MA(1) components.



(a) 1965 and 1970 Children

(b) 1975 and 1980 Children

Figure 2: Intergenerational Covariances of Log Earnings by Children's Age (at Parents' Age 55)

group. The latter pattern implies that the intergenerational covariances are higher for children born to younger parents.

Lifecycle and cohort patterns also emerge when we hold parents' age (instead of children's age) constant at 55, as presented in Figure 2. The intergenerational covariances are positive and hump-shaped in children's age, but the peak is reached at earlier ages—around ages 35–45. The covariances are also decreasing in parental age at birth, although they are nearly identical among three of the family cohort groups reported in Figure 2b.[20]

To more formally examine whether intergenerational log earnings covariances vary with the ages of parents and children, as suggested by Figures 1 and 2, we separately test whether (i) fathers' log earnings levels are uncorrelated with their sons' log earnings growth and (ii) fathers' log earnings growth is uncorrelated with their sons' log earnings levels.[21] We strongly reject both hypotheses, confirming the presence of lifecycle effects (for both fathers and sons) on intergenerational log earnings covariances.

We can also examine how log earnings growth, $\Delta y_{i,j,t} := y_{i,j,t} - y_{i,j,t-1}$, is correlated across generations. Figures 3 and 4 show the intergenerational covariances in log earnings growth by parents' and children's ages, respectively. In contrast to the covariances between log earnings, the covariances between log earnings growth are centered around zero and exhibit no clear lifecycle or cohort patterns. Using a bootstrap Wald test, we cannot reject that all of the age- and cohort-specific intergenerational covariances of log earnings growth jointly equal zero at the 5% significance level.[22] Performing the

[20]One might be concerned that the lifecycle patterns for different cohorts shown in Figure 2 are driven by gradual entry into the labor market by sons, which may differ across cohorts. To examine this issue, we re-estimate the reported covariances over ages 26–34 using only families in which the sons worked at least 7 of the 9 years in that age range. The estimated lifecycle patterns (across cohorts) are quite similar; however, all covariances are slightly larger.

[21]Specifically, we jointly test whether all relevant covariances, separately for cases (i) and (ii), equal zero using a bootstrap Wald test. See Appendix C for further details.

[22]For some perspective, the largest correlation (in absolute value) is less than 0.017. When separately testing whether each of the 1,130 sample covariances between the earnings growth of fathers and their sons equals zero, we reject that 57 (or 5%) equal zero at the 5% significance level. See Appendix C for details.

same joint test separately for each of our 10 cohort groups, we fail to reject zero intergenerational covariance in log earnings growth in all cases.

The covariance patterns in log earnings, as well as log earnings growth, can be interpreted through the lens of our intergenerational model of earnings presented in Section 2, which implies

$$\begin{aligned}
\mathrm{Cov}(y_{p,t}, y_{k,t'}) =& \mathrm{Cov}(\psi_p, \psi_k) + \Lambda_{p,t}\Lambda_{k,t'}\,\mathrm{Cov}(\delta_p, \delta_k) + \Lambda_{k,t'}\,\mathrm{Cov}(\psi_p, \delta_k) + \Lambda_{p,t}\,\mathrm{Cov}(\delta_p, \psi_k),\\
\mathrm{Cov}(y_{j,t}, \Delta y_{j',t'}) =& \lambda_{j',t'}[\mathrm{Cov}(\psi_j, \delta_{j'}) + \Lambda_{j,t}\,\mathrm{Cov}(\delta_j, \delta_{j'})], \quad \text{for } j \neq j',\\
\mathrm{Cov}(\Delta y_{p,t}, \Delta y_{k,t'}) =& \lambda_{p,t}\lambda_{k,t'}\,\mathrm{Cov}(\delta_p, \delta_k).
\end{aligned}$$

Figures 1 to 4 are revealing about the intergenerational covariances in initial skills and skill growth factors. The finding that log earnings covariances vary with both parents' and children's ages (equivalently, that $\mathrm{Cov}(y_{j,t}, \Delta y_{j',t'}) \neq 0$) means that $\lambda_{p,t}$ and $\lambda_{k,t}$ are non-zero over those ages. This, coupled with the result that $\mathrm{Cov}(\Delta y_{p,t}, \Delta y_{k,t'}) = 0$ for all ages of parents and children, indicates that the skill growth factor is uncorrelated across generations (for all family cohort groups): $\mathrm{Cov}(\delta_p, \delta_k) = 0$. However, the skill growth factor still plays an important role in intergenerational transmission. Because $\mathrm{Cov}(\delta_p, \delta_k) = 0$, the positively sloped lines in Figures 1 and 2 (at least over younger ages) indicate that both $\mathrm{Cov}(\delta_p, \psi_k)$ and $\mathrm{Cov}(\psi_p, \delta_k)$ are positive.

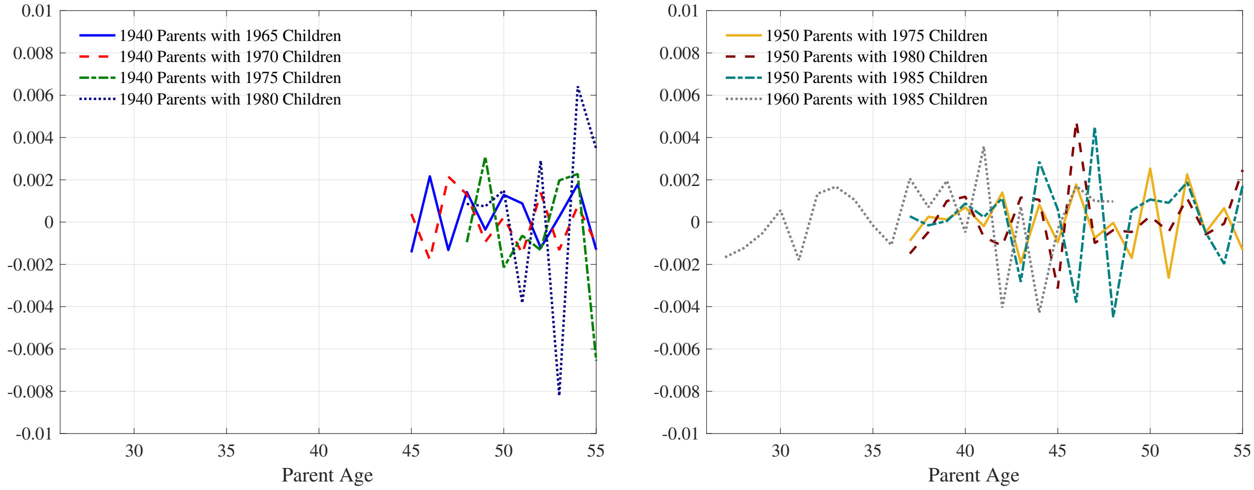
Figures 1 and 2 also shed important light on cohort differences in intergenerational transmission. The fact that log earnings covariances tend to be decreasing in children's birth year (conditional on parent's birth year) and increasing in parental birth year (conditional on child's birth year) suggests that $\mathrm{Cov}(\psi_p, \psi_k)$ is greater for children born to younger parents. The roughly parallel lifecycle trajectories across cohorts in Figure 1 suggest that $\mathrm{Cov}(\delta_p, \psi_k)$ are quite similar across cohorts, while the steeper lifecycle trajectories for children born to older parents (e.g., children from the 1980 cohort born to parents from the 1940 cohort) shown in Figure 2 suggest that $\mathrm{Cov}(\psi_p, \delta_k)$ is higher for these families relative to those in which parents were younger when their child was born.[23]

These observed intergenerational covariance patterns are inconsistent with a single-factor model of skill transmission. To see this, suppose that $\theta_{i,j,t} = \chi_{j,t}\overline{\theta}_{i,j}$ as assumed by much of the recent literature (Jenkins, 1987; Grawe, 2006; Haider and Solon, 2006; Nybom and Stuhler, 2016; Deutscher and Mazumder, 2020).[24] This single-factor model implies that $\mathrm{Cov}(y_{p,t}, y_{k,t'}) = \chi_{p,t}\chi_{k,t'}\,\mathrm{Cov}(\overline{\theta}_p, \overline{\theta}_k)$, so differences in these covariances across parents' or children's ages implies that $\Delta\chi_{p,t}$ and $\Delta\chi_{k,t'}$ are non-zero for those ages (see Figures 1 and 2). Yet, this is inconsistent with zero intergenerational covariances in log earnings growth (i.e., $\mathrm{Cov}(\Delta y_{p,t}, \Delta y_{k,t'}) = \Delta\chi_{p,t}\Delta\chi_{k,t'}\,\mathrm{Cov}(\overline{\theta}_p, \overline{\theta}_k)$) for all ages (see Figure 4).[25]

[23] Our model estimates for $\mathrm{Cov}(\psi_p, \psi_k)$, $\mathrm{Cov}(\delta_p, \psi_k)$, and $\mathrm{Cov}(\psi_p, \delta_k)$, discussed in Section 5.1, are broadly consistent with these conclusions (also see Appendix Figure G6).

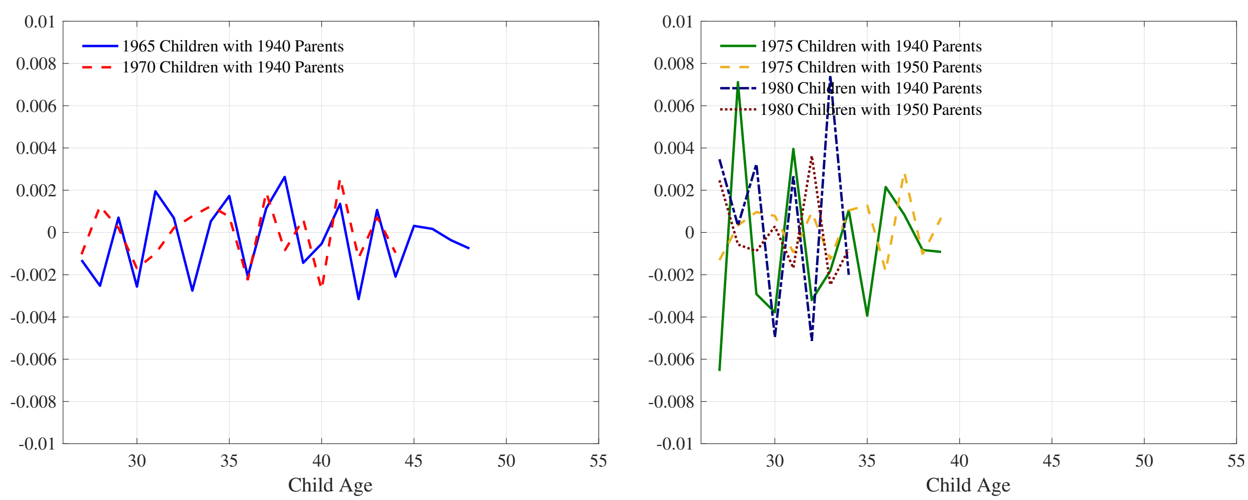
[24] If $\overline{\theta}_{i,j}$ reflects expected average lifetime income, then the lifecycle variation in $\chi_{j,t}$ reflects the changing relationship between annual and lifetime earnings, as emphasized in Haider and Solon (2006).

[25] Haider and Solon (2006) motivate their analysis of measurement error and lifecycle bias using a two-factor model of lifecycle earnings analogous to our model of skills. In this case, one can define $u_{i,j,t} := y_{i,j,t} - \chi_{j,t}\overline{\theta}_{i,j}$ where $\chi_{j,t}$ is the linear-projection coefficient and $u_{i,j,t}$ is uncorrelated with $\overline{\theta}_{i,j}$. However, the $u_{i,j,t}$ will depend on both factors ($\psi_{i,j}$ and $\delta_{i,j}$ in our framework) and will, therefore, be intergenerationally correlated. Thus, the lifecycle biases Haider and Solon (2006) define when considering only one generation at a time in a regression of $y_{i,j,t}$ on some exogenous variable $x_{i,j,t}$ (or the



(a) 1940 Parents (b) 1950 and 1960 Parents

Figure 3: Intergenerational Covariance of Log Earnings Growth by Parents' Age (at Children's Age 27)



(a) 1965 and 1970 Children (b) 1975 and 1980 Children

Figure 4: Intergenerational Covariance of Log Earnings Growth by Children's Age (at Parents' Age 55)

The discussion above suggests that intergenerational covariances in log earnings are also helpful for understanding the evolution of earnings and skills over the lifecycle. Most notably, the literature on earnings dynamics has struggled to identify the extent of heterogeneity in lifecycle earnings profiles (i.e., HIP process) relative to the importance of persistent earnings shocks using autocovariances from a single generation alone (see, e.g., Lillard and Weiss, 1979; Hause, 1980; MaCurdy, 1982; Baker, 1997; Haider, 2001; Guvenen, 2009; Hryshko, 2012). In our framework, the age-varying intergenerational covariances shown in Figures 1 and 2 provide transparent evidence of individual heterogeneity in lifecycle skill growth. Moreover, the hump-shaped lifecycle patterns for intergenerational covariances imply that the standard HIP process (assuming $\lambda_{j,t} = 1$) is too restrictive. Instead, they suggest that $\lambda_{j,t}$ is declining with age, consistent with the diminishing role of human capital investment over the lifecycle as predicted by human capital theory (Ben-Porath, 1967; Mincer, 1974; Becker, 1975).

We next exploit these covariance patterns along with intragenerational autocovariances to estimate our full intergenerational model (as described in Section 2).

4 Minimum Distance Estimation

We estimate our model by minimizing the distance between the sample covariances for log earnings and the theoretical covariances implied by the model. The parameters related to the dynamic process and variances of non-skill shocks, the cross-sectional distributions of initial skills and the skill growth factor, and the trajectory for skill growth over the lifecycle can be identified from the lifecycle autocovariances within generations, while the intergenerational transmission process for initial skills and the skill growth factor are identified from intergenerational covariances at different ages of parents and their children. See Appendix A for a detailed discussion of identification.

While we allow for differences in parameters across family cohort groups, we impose some restrictions to ensure that parameters are identified and can be easily compared across family cohort groups. As Table 1 shows, the earnings of individuals in the 1930 parent cohort group and 1985 child cohort group are observed for only a few years. Since identification of the earnings process for a single generation requires a sufficient number of years for which earnings are observed, we do not estimate variances of the skill factors and earnings shocks for these cohorts and, thus, do not target their autocovariances. Excluding these 22 autocovariances results in 3,769 targeted covariances.

The covariances of skill factors $(\psi_p, \delta_p, \psi_k, \delta_k)$ are composed of three covariance matrices: (i) the covariance matrix for parental skills $\mathbf{\Omega}_p$, (ii) the intergenerational skill covariance matrix $\mathbf{\Omega}_{p,k}$, and (iii) the covariance matrix for children's skills

$$\mathbf{\Omega}_k := \begin{bmatrix} \mathrm{Var}(\psi_k) & \mathrm{Cov}(\psi_k, \delta_k) \\ \mathrm{Cov}(\psi_k, \delta_k) & \mathrm{Var}(\delta_k) \end{bmatrix}.$$

reverse regression) provide an incomplete characterization of the lifecycle bias for IGEs in $\overline{\theta}_{i,j}$ if skills are characterized by our two-factor model. Nybom and Stuhler (2016) account for differences in lifecycle earnings profiles by post-secondary education attendance, estimating important differences between the actual lifecycle bias and that implied by the terms emphasized in Haider and Solon (2006).

From these, we can obtain estimates of the covariance matrix for projected skill factors $(\hat{\psi}_k, \hat{\delta}_k)$ and the linear-projection coefficients $\mathbf{A}$ (defined earlier in Section 2.2.2). We estimate the covariance matrices $\mathbf{\Omega}_p$, $\mathbf{\Omega}_{p,k}$, and $\mathbf{\Omega}_k$ separately for each family cohort group. Based on our failure to reject $\text{Cov}(\Delta y_p, \Delta y_k) = 0$ for all cohorts (as discussed in Section 3), we assume $\text{Cov}(\delta_p, \delta_k) = 0$.[26]

The variances of persistent and transitory shocks are estimated separately for each generation, age, and family cohort group; however, some normalizations are needed. Persistent shock variances are held constant for the last 2 years, because they are not separately identified from transitory shock variances. We assume all transitory shock variances prior to the initial period (age $\underline{t} = 26$ or the first year in the sample for the cohort) are the same as that of the initial period.

The remaining parameters (coefficients on the skill growth factor for each age, $\Lambda_{j,t}$, and the AR(1) and MA(1) coefficients) are estimated separately for each generation, assuming that they are identical across family cohort groups within generations. Recall that $\Lambda_{j,t}$ is normalized to 1 at age 27 ($\underline{t} + 1$), which sets the scale of the skill growth factor. Although we observe all child cohorts as early as age 26, this is not the case for most parent cohorts. By assuming that $\Lambda_{p,t}$ does not vary across parent cohorts, we avoid setting the scale of the skill growth factor based on different ages for different cohorts. This facilitates cross-cohort comparisons of estimated variances/covariances for the skill factors.

We use the equally weighted minimum distance (MD) estimator $\hat{\mathbf{\Theta}}$, which solves

$$\text{SSR} := \min_{\mathbf{\Theta}} \sum_{g=1}^{10} \sum_{j,j',t,t'} \left\{ \widehat{\text{Cov}}(y_{j,t}, y_{j',t'}|g) - \text{Cov}(y_{j,t}, y_{j',t'}|g, \mathbf{\Theta}) \right\}^2,$$

where $\widehat{\text{Cov}}(y_{j,t}, y_{j',t'}|g)$ is the sample covariance for family cohort group g, and $\text{Cov}(y_{j,t}, y_{j',t'}|g, \mathbf{\Theta})$ is the corresponding theoretical covariance for a given parameter $\mathbf{\Theta}$ defined earlier. This approach is widely applied in the literature on lifecycle earnings dynamics but less so in the literature on intergenerational transmission (notable exceptions include Altonji and Dunn, 1991; Zimmerman, 1992; Gallipoli, Low, and Mitra, 2020). We calculate standard errors by repeating the MD estimation over 100 bootstrap samples.

5 Intergenerational Mobility in Canada

5.1 Model Parameter Estimates

Our model has hundreds of parameters, including those related to the distribution of skill factors (ψ_j, δ_j) and their intergenerational transmission, the process of lifecycle skill accumulation, and the dynamics of non-skill earnings shocks. We only briefly summarize these estimates here, providing detailed results in Appendix G.

Our estimates suggest that the variance of initial skills, ψ_j, is substantially greater for fathers (ranging from 0.3 to 0.55) than for their sons (generally less than 0.1), while the variance of skill

[26]As shown in Appendix A, $\text{Cov}(\delta_p, \delta_k)$ is identified, along with all other parameters. Since we do not observe parents or children for many years in some cohorts, results obtained when relaxing this assumption (i.e., when estimating $\text{Cov}(\delta_p, \delta_k)$) are less precise for these cohorts; however, they are generally quite similar to those reported in the paper.

growth, δ_j, ranges from 0.0004 to 0.0024 for both parent and child cohorts. Focusing on child cohorts, we generally observe a declining variance in initial skills and a growing variance in skill growth over time. The most notable difference across the two generations is in the correlation between initial skills and skill growth, which is greater than 0.8 for all child cohorts but always negative and less than -0.6 for all parent cohorts. Early career skill differences tend to grow over the lifecycle for the child cohorts, while the opposite is observed for their parents, at least early in their careers.

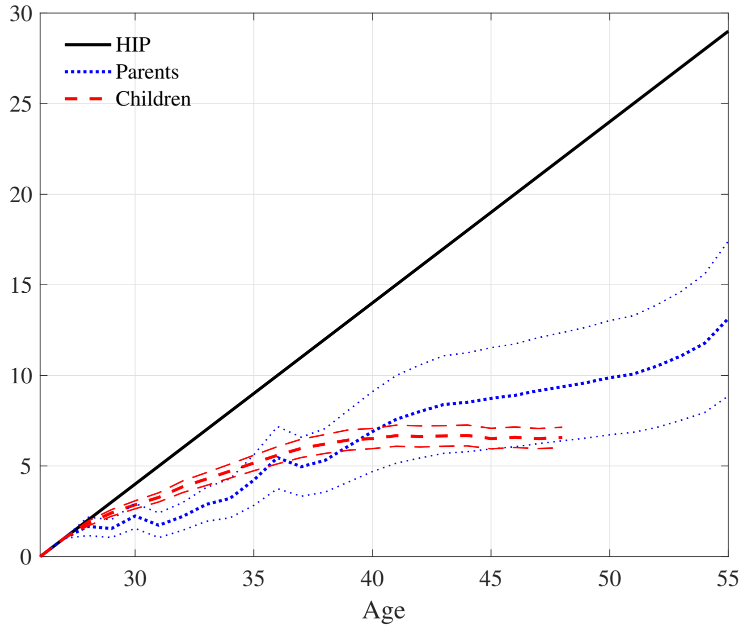


Figure 5: Coefficient on the Skill Growth Factor ($\Lambda_{j,t}$, Thick Lines) and 95% Confidence Interval (Thin Lines)

Figure 5 shows the estimated lifecycle skill variance trajectory, Λ_j, for both sons and fathers, compared to that implied by the standard HIP model. As human capital theory predicts (Ben-Porath, 1967; Becker, 1975), variation in skill growth declines smoothly over the lifecycle for the child's generation, with no further growth in heterogeneity beyond age 40. Variation in skill growth for the generation of parents is more constant over the lifecycle; although, there is a notable decline after the first couple years.[27]

Skills are positively correlated across generations. The intergenerational correlation in initial skills is as high as 0.3 for the 1965 cohort of children (with fathers from the 1940 cohort) and declines over time to about 0.11 for the 1980 cohort of children (with fathers from either the 1940 or 1950 cohorts). For families with fathers from the 1940 cohort, we observe a similar time pattern for Corr(δ_p, ψ_k) but the opposite for Corr(ψ_p, δ_k), with both of these correlations ranging from 0.14 to 0.26. See Appendix G for other cohorts and estimates of (standardized) linear-projection coefficients **A**.

[27]The stronger skill growth after age 40 among fathers (relative to sons) likely reflects the fact that many fathers in our sample would have been in their 40s during the 1980s when the returns to skill rose sharply, while even the oldest sons in our sample did not reach age 40 until after 2000 when returns to skill are generally thought to have risen more slowly or even fallen. See, e.g., Acemoglu and Autor (2011) for a survey of mostly US-based studies or Boudarbat, Lemieux, and Riddell (2010), Green and Sand (2015), and Fortin and Lemieux (2015) for changes in the Canadian wage structure.

Finally, Appendix Figures G9 and G10 show the estimated variances of persistent ($\nu_{j,t}$) and transitory ($\xi_{j,t}$) shocks, respectively, for all cohorts at each age. The estimated autocorrelation for persistent shocks, ρ_j, is higher for the generation of fathers than sons (0.91 vs. 0.83), but both suggest that earnings shocks are moderately persistent with half lives of about 7 and 4 years, respectively. The estimated MA(1) correlation parameters for transitory shocks, κ_j, are very similar and modest for fathers and sons (0.12 and 0.11, respectively).

5.2 Intergenerational Earnings and Skill Elasticities

Before we discuss the intergenerational transmission of skills, we document basic earnings IGEs for each family cohort pair in Table 3. We provide two sets of earnings IGE estimates: one set that is estimated directly from our data (denoted as "Data" in the table) and another that is computed from the estimated model parameters discussed in the previous subsection (denoted as "Model" in the table), where the close agreement of the two provides confidence in our intergenerational framework and MD estimation approach. The table reports estimated IGEs for annual measures of earnings (at ages 30 for children and 50 for parents), 5-year average earnings (centered around ages 30 and 50), and 9-year averages (at ages 26–34 for children and 47–55 for parents). These ages are broadly consistent with those used in previous research and enable us to exploit sufficiently large sample sizes for both generations.[28] Figure 6 displays the "Model" earnings IGE estimates for all cohorts of children with fathers from the 1940 or 1950 cohorts.

Four points stand out from Table 3 and Figure 6. First, these earnings IGEs are quite low, with none exceeding 0.2 and annual earnings IGEs as low as 0.06. As previously noted by Corak and Heisz (1999) and Chen, Ostrovsky, and Piraino (2017), intergenerational earnings mobility is much stronger in Canada than in the U.S. Second, earnings IGEs are declining for more recent cohorts of children. Among children with fathers from the 1940 cohort, the annual IGE falls by about half from 0.13 for the earliest (1965) cohort of sons to 0.06 for the most recent (1980) cohort. Third, IGEs are higher for the two most recent cohorts of sons with younger (1950 cohort) relative to older (1940 cohort) fathers. This could reflect differences over time in the environments in which these fathers grew up, or it could signal a role for parental age in the intergenerational transmission process. We show in Appendix E that these patterns do not simply reflect differences in family structure or birth order that are correlated with parental age. Among single or two-parent families or among families with 1, 2, or 3+ children, we observe remarkably similar declines in earnings IGEs with father's age (at son's birth). This same pattern is also observed when broadening our sample to include all sons and looking across birth order. Fourth, 5-year average earnings IGEs are about 40% larger than annual IGEs, while 9-year averages are another 5–10% larger. Using the same data, Chen, Ostrovsky, and Piraino (2017) estimate earnings IGEs as high as 0.32 (for a similar cohort to our 1965 children with 1940 parents) when using parental earnings averaged over 21 years.[29] This pattern reflects the diminished

[28] We allow for at most 1 missing observation for 5-year averages and 2 missing observations for 9-year averages.

[29] Chen, Ostrovsky, and Piraino (2017) estimate the IGE for parental earnings averaged over ages 35–55 and children's earnings averaged over ages 38–42. Appendix D shows that nearly all of the difference between their estimate and our 5-year average IGE of 0.19 for 1965 children with 1940 parents (Table 3, Panel B) is explained by the different ages over which

	Cohort Group					
Parents	1940				1950	
Children	1965	1970	1975	1980	1975	1980
			A. Annual			
Earnings IGE (Data)	0.145	0.124	0.090	0.059	0.112	0.100
	(0.003)	(0.003)	(0.003)	(0.005)	(0.002)	(0.002)
Earnings IGE (Model)	0.132	0.108	0.089	0.060	0.099	0.094
	(0.002)	(0.001)	(0.002)	(0.003)	(0.002)	(0.001)
Skill IGE	0.422	0.316	0.253	0.184	0.378	0.330
	(0.010)	(0.007)	(0.012)	(0.017)	(0.007)	(0.006)
			B. 5-Year Average			
Earnings IGE (Data)	0.189	0.156	0.122	0.081	0.140	0.127
	(0.003)	(0.003)	(0.003)	(0.006)	(0.003)	(0.002)
Earnings IGE (Model)	0.186	0.153	0.127	0.087	0.141	0.131
	(0.002)	(0.002)	(0.003)	(0.005)	(0.002)	(0.002)
Skill IGE	0.422	0.316	0.252	0.183	0.379	0.330
	(0.010)	(0.007)	(0.012)	(0.017)	(0.007)	(0.006)
			C. 9-Year Average			
Earnings IGE (Data)	0.196	0.160	0.128	0.084	0.145	0.135
	(0.003)	(0.003)	(0.003)	(0.006)	(0.002)	(0.002)
Earnings IGE (Model)	0.201	0.163	0.136	0.092	0.151	0.142
	(0.002)	(0.002)	(0.003)	(0.005)	(0.002)	(0.002)
Skill IGE	0.436	0.326	0.263	0.185	0.399	0.342
	(0.009)	(0.007)	(0.011)	(0.016)	(0.009)	(0.007)

Notes: Annual measures are based on ages 30 for children and 50 for fathers, 5-year averages based on ages 28–32 for children and 48–52 for fathers, and 9-year averages based on ages 26–34 for children and 47–55 for parents. Standard errors (in parentheses) for "Earnings IGE (Data)" are based on standard asymptotic ordinary least squares (OLS) formulas, while all other standard errors are based on 100 bootstrap samples.

Table 3: Earnings and Skill IGEs by Cohort

influence of year-to-year fluctuations in earnings when averaging across several years of data (Solon, 1992; Jenkins, 1987).

One potential concern with the results reported in Table 3 and Figure 6 is the roughly 20-year age difference at which earnings for fathers and sons are measured. We report IGEs at these ages to enable comparisons across many different cohorts; however, earnings IGEs generally depend on the ages at which earnings are measured (Jenkins, 1987; Haider and Solon, 2006). Indeed, our two-factor model of skills is largely motivated by these differences and an effort to better understand them. Importantly, estimation of our model takes advantage of intergenerational relationships at all available ages for each family cohort. To explore the extent to which ages at which earnings are measured affects earnings IGEs (especially across cohorts), Appendix Figure F4 documents annual earnings IGEs separately for cohorts of children born from 1968 to 1974 (whose parents were no older than 30 years old at their birth) for whom we observe earnings at age 40 for both children and their parents. While earnings IGEs can differ by as much as 0.06 (never exceeding 0.20) depending on the ages at which earnings are measured, we generally observe lower IGEs for more recent cohorts of children. Indeed, the decline over time appears to be slightly stronger when earnings are measured at age 40 for both generations compared to ages 30 for children and 50 for parents.

These seemingly low IGEs, even when averaging earnings across several years, seem to suggest that intergenerational skill transmission is relatively weak in Canada. Yet, as discussed in Section 2.2.1, these low estimates could simply reflect considerable earnings instability, since the earnings IGE equals the skill IGE scaled by the fraction of parental log earnings attributed to skills. Appendix Table H6 shows that skills account for one-quarter to one-third of the annual log earnings variance (see rows labeled $\text{Var}(\theta_{p,t})/\text{Var}(y_{p,t})$), rising to about 40–50% of the variance for 5- and 9-year average log earnings. Looking at 13-year average log earnings (ages 36–48), we see that skills explain as much as 70% of the variance for the 1950 cohort of parents. Among the youngest (1960) cohort of fathers, for whom we see earnings over much of their lives, skills account for only 50% of the 13-year average log earnings and 56% of their average earnings over 23 years (ages 26–48). While we cannot estimate IGEs for 13- or 23-year averages, since we do not observe the same set of fathers and sons for those lengths of time, it is clear that even long-run averages of earnings provide only a murky picture of skill levels. These statistics suggest that skill IGEs are 2–3 times greater than earnings IGEs.

Skill IGEs are reported in Table 3 and graphed in Figure 7.[30] These skill IGEs show similar cross-cohort patterns to earnings IGEs, except that they are noticeably larger, ranging from 0.18 to 0.44. Interestingly, taking 5- or 9-year averages of skills does not appreciably impact the skill IGEs for two reasons: (i) transitory earnings shocks do not affect skill IGEs, and (ii) $\Lambda_{j,t}$ is roughly linear (i.e., $\lambda_{j,t}$ is almost constant) over the ages we average across. This implies that the increase in the earnings IGEs associated with averaging earnings across longer periods comes entirely from the reduction in

fathers' and sons' earnings are averaged: 56% of the difference is explained by our use of fathers' earnings averaged over 5 rather than 21 years (centered around the midpoint of ages used in Chen, Ostrovsky, and Piraino (2017)), while 37% of the difference is due to the ages around which both sons' and fathers' earnings are measured.

[30] We report skill IGEs based on our MD model earnings IGE estimates (divided by the estimated share of parental earnings from skills), but it is also feasible to use the "Data" earnings IGE estimates instead. Since these are quite similar, we do not report them.

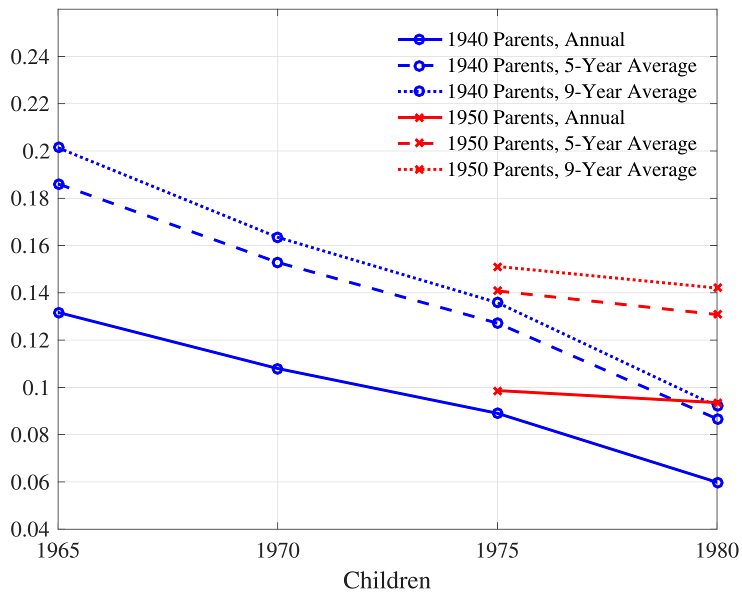


Figure 6: Earnings IGEs by Cohort

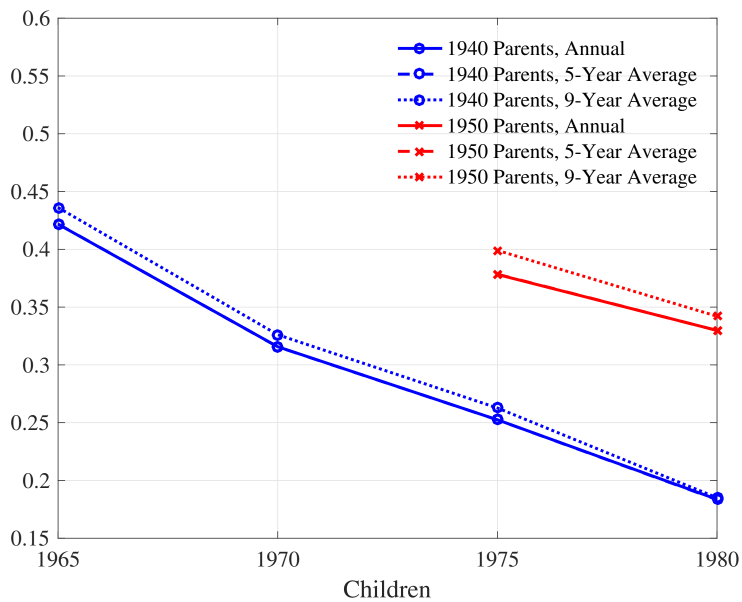


Figure 7: Skill IGEs by Cohort

variation from averaging earnings shocks. We also see a much larger increase in skill IGEs (compared to earnings IGEs) associated with the younger 1950 cohort of parents. This directly reflects the lower share of their earnings variation that is attributed to variation in skills. While children from the 1970 cohort of children with fathers from the 1940 cohort have higher earnings IGEs than children from the 1975 cohort with parents from the 1950 cohort, the reverse is true when we look at skill IGEs. This demonstrates the importance of accounting for earnings instability if interested in understanding the intergenerational transmission of skills themselves.

5.3 Projected Skill Variation

While much of the literature has focused on earnings IGEs (and, more recently, rank-rank slopes), we consider an alternative measure of intergenerational transmission: the share of children's earnings or skill variance that can be explained by projected skills (i.e., skill levels predicted by parental skill factors ψ_p and δ_p). This measure is informative about the extent of inequality across children that can be explained by intergenerational transmission, while both the IGE and rank-rank slope are informative about the rate of regression to the mean in earnings or skills. A nice feature of the projected skill variance is that it can be easily decomposed into variation related to initial skills vs. skill growth factors, which can shed light on the relative importance of the intergenerational transmission of these distinct factors over the lifecycle.

It is useful to begin by first documenting the extent to which children's skills account for variation in their earnings, much as we already did for parents. Table 4 reports the estimated share of earnings variation explained by skills for annual measures (at age 30), 5-year averages (ages 28–32), and 9-year averages (ages 26–34). Appendix Table H7 reports shares based on 13-year averages (ages 26–38) and 23-year averages (ages 26–48). Consistent with our findings for the fathers in our sample, these shares are as low as 25% for annual averages, rising to about 40% for 9-year averages (and as high as 70% for 23-year averages). Thus, skills are an important component of earnings, but other idiosyncratic factors also play an important role in determining not only earnings at a point in time but also a career's worth of earnings.

Our main interest is in the transmission of skills. To that end, Table 4 also reports the shares of children's earnings and skill variation that are explained by variation in projected skill levels, $\hat{\theta}_k$, for annual measures as well as averages over several years (see equations (5) and (6)). Figure 8 displays these estimated shares of skill variation for all cohorts of children with fathers from the 1940 and 1950 cohorts. These show qualitatively similar cross-cohort patterns as the IGEs discussed earlier: there is a clear downward trend in the importance of intergenerational transmission of skills. However, this measure shows stability between the 1970 and 1975 cohorts of children (with 1940 parents), while the skill IGEs showed a nearly linear decline across the 1965–1980 cohorts. Looking at the children of fathers from the 1940 cohort, the variation in skills predicted by parental skill factors (ψ_p and δ_p) explains about 40% of the skill variation for the 1965 child cohort, dropping to 20–23% for the 1970 and 1975 cohorts, and further down to slightly less than 10% for the 1980 cohort. Among the two most recent cohorts with younger fathers (from the 1950 cohort), the share of skill variation explained by

Parents	1930		1940				1950	
Children	1965	1970	1965	1970	1975	1980	1975	1980
			A. Annual					
$\mathrm{Var}(\theta_{k,t})/\mathrm{Var}(y_{k,t})$	0.245	0.246	0.250	0.241	0.246	0.239	0.250	0.250
	(0.003)	(0.005)	(0.003)	(0.003)	(0.004)	(0.007)	(0.005)	(0.003)
$\mathrm{Var}(\hat{\theta}_{k,t})/\mathrm{Var}(y_{k,t})$			0.097	0.051	0.055	0.020	0.066	0.051
			(0.014)	(0.009)	(0.012)	(0.004)	(0.005)	(0.003)
$\mathrm{Var}(\hat{\theta}_{k,t})/\mathrm{Var}(\theta_{k,t})$			0.388	0.211	0.225	0.085	0.263	0.202
			(0.056)	(0.036)	(0.050)	(0.017)	(0.021)	(0.012)
			B. 5-Year Average					
$\mathrm{Var}(\overline{\theta}_k)/\mathrm{Var}(\overline{y}_k)$	0.376	0.366	0.385	0.366	0.363	0.357	0.379	0.377
	(0.004)	(0.006)	(0.003)	(0.004)	(0.006)	(0.010)	(0.007)	(0.005)
$\mathrm{Var}(\overline{\hat{\theta}}_k)/\mathrm{Var}(\overline{y}_k)$			0.150	0.077	0.082	0.030	0.100	0.076
			(0.022)	(0.013)	(0.018)	(0.006)	(0.008)	(0.004)
$\mathrm{Var}(\overline{\hat{\theta}}_k)/\mathrm{Var}(\overline{\theta}_k)$			0.389	0.212	0.225	0.085	0.264	0.202
			(0.056)	(0.036)	(0.050)	(0.017)	(0.021)	(0.012)
			C. 9-Year Average					
$\mathrm{Var}(\overline{\theta}_k)/\mathrm{Var}(\overline{y}_k)$	0.432	0.406	0.441	0.405	0.408	0.396	0.425	0.423
	(0.004)	(0.007)	(0.004)	(0.004)	(0.007)	(0.010)	(0.008)	(0.005)
$\mathrm{Var}(\overline{\hat{\theta}}_k)/\mathrm{Var}(\overline{y}_k)$			0.174	0.087	0.093	0.034	0.114	0.086
			(0.026)	(0.015)	(0.021)	(0.007)	(0.009)	(0.005)
$\mathrm{Var}(\overline{\hat{\theta}}_k)/\mathrm{Var}(\overline{\theta}_k)$			0.395	0.214	0.228	0.086	0.268	0.204
			(0.057)	(0.037)	(0.051)	(0.017)	(0.022)	(0.012)

Cohort Group spans all columns: Parents 1930 (children 1965, 1970); 1940 (children 1965, 1970, 1975, 1980); 1950 (children 1975, 1980).

Notes: Annual measures are based on ages 30 for children and 50 for fathers, 5-year averages based on ages 28–32 for children and 48–52 for fathers, and 9-year averages based on ages 26–34 for children and 47–55 for parents. Standard errors (in parentheses) are based on 100 bootstrap samples.

Table 4: Children's Skill and Earnings Variances by Cohort

projected skills is about 5–10 percentage points higher than among those children with older fathers. Clearly, intergenerational transmission has become a less important determinant of skill inequality for more recent cohorts.

We next consider the importance of skill transmission over the lifecycle, continuing to focus on the share of earnings or skill variation explained by projected skills. As equation (6) makes clear, intergenerational transmission will explain a greater share of earnings variation when skills are a more important component of children's earnings.[31] Figure 9 shows that the fraction of earnings variation explained by skills grows steadily over the first half of children's careers due to declining variation in non-skill shocks and the rising influence of skill growth heterogeneity.[32] Figure 10 documents a very similar lifecycle pattern for the share of earnings variation explained by projected skills. For example, consider the 1965 cohort of children with fathers from the 1940 cohort: projected skills explain about 8% of the total variance in log earnings at age 26, rising to about 12% by ages 35-40 when it stabilizes. Other cohorts show similar growth over the lifecycle, but much lower levels of intergenerational transmission (consistent with the cohort differences reported earlier).

Although the growing share of earnings variation explained by projected skills (at least over early parts of the lifecycle) can be understood with the traditional single-factor model, we next show that accounting for two skill factors (initial skills and skill growth) provides novel insights. First, we note that the share of skill variation explained by projected skills—equivalently, the share of skill variation explained by parental skill factors—is not necessarily constant over the lifecycle. Figure 11 shows that for most cohorts, projected skills explain a declining share of the variance in skills early in their careers, because heterogeneity in skill growth rates is not well-predicted by parental skill factors. Most dramatically, the 1965 cohort of children (with fathers from the 1940 cohort) sees this share decline from about 0.5 in their mid-twenties to just under 0.3 by their late-thirties. While skills become a more important driver of earnings inequality over the first 10–15 years of workers' careers, intergenerational transmission becomes a less important determinant of skills themselves. The former highlights the importance of heterogeneity in the ability to acquire new skills, while the latter indicates that individual differences in this ability are not strongly determined by one's parents.

To better understand these patterns, we consider variance decompositions that shed light on the roles of initial skill and skill growth heterogeneity. Figure 12 shows the contribution of children's initial skills and skill growth rates (as well as their covariance) to the variance of their skills over the lifecycle for the 1965 and 1975 cohorts (both with fathers from the 1940 cohort).[33] While, by definition, variation in skills is entirely determined by initial skills at the beginning of workers' careers, the influence of skill growth heterogeneity (including the covariance term) begins to dominate after only 5–10 years. An analogous decomposition for projected skills (i.e., separating the variation in $\hat{\theta}$ into components related to predicted initial skills, $\hat{\psi}_k$, and predicted skill growth, $\hat{\delta}_k$, as well as their covariance) yields very similar lifecycle trajectories (Figure 13); however, predicted initial skill levels

[31] By contrast, the discrepancy between earnings and skill IGEs depends directly on the share of *parental* earnings variation explained by skills, which may vary with the age at which parental earnings are measured.

[32] By contrast, Appendix Figures H12 and H13 show that the share of earnings variation explained by skills declines over the later part of the lifecycle, at least for the generation of fathers.

[33] Results in Figures 12–14 are quite similar for other cohorts.

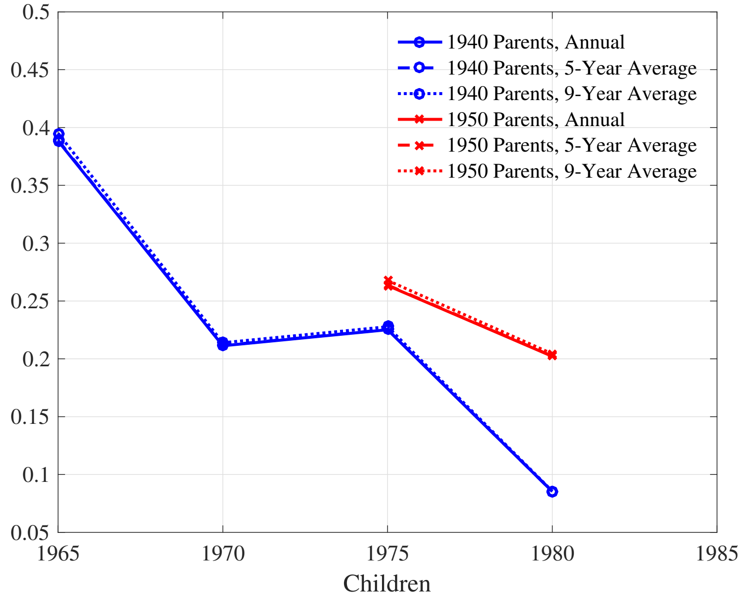
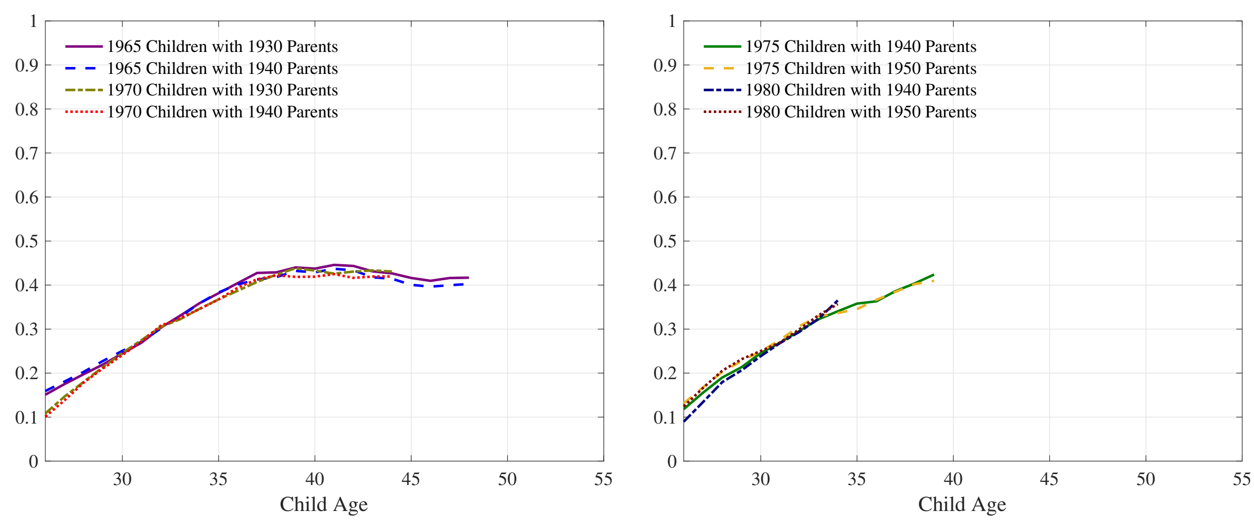


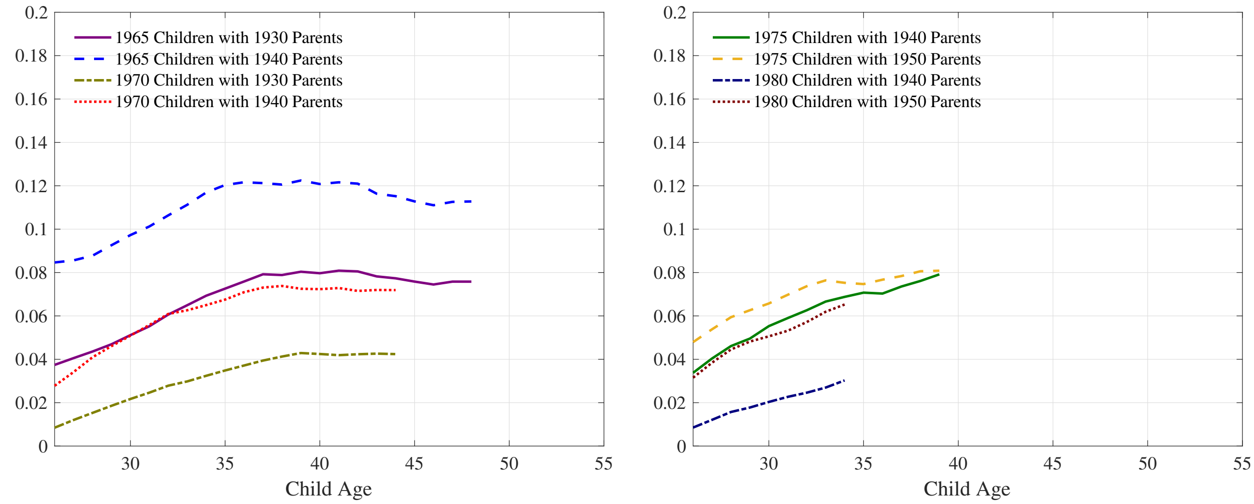
Figure 8: $\mathrm{Var}(\overline{\hat{\theta}}_k)/\mathrm{Var}(\overline{\theta}_k)$ by Cohort



(a) 1965 and 1970 Children

(b) 1975 and 1980 Children

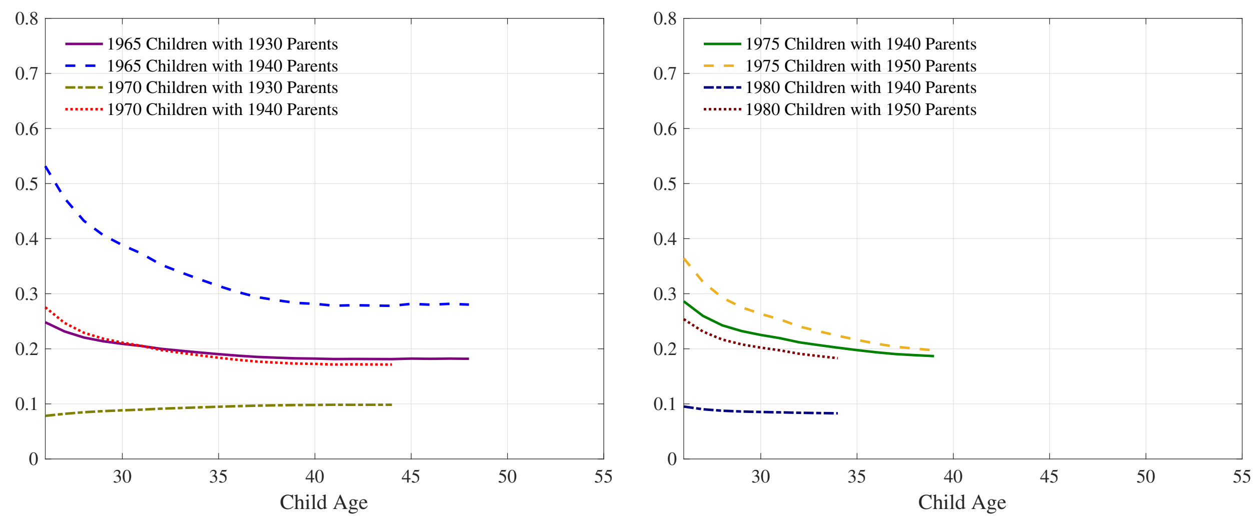
Figure 9: $\mathrm{Var}(\theta_{k,t})/\mathrm{Var}(y_{k,t})$ by Cohort



(a) 1965 and 1970 Children

(b) 1975 and 1980 Children

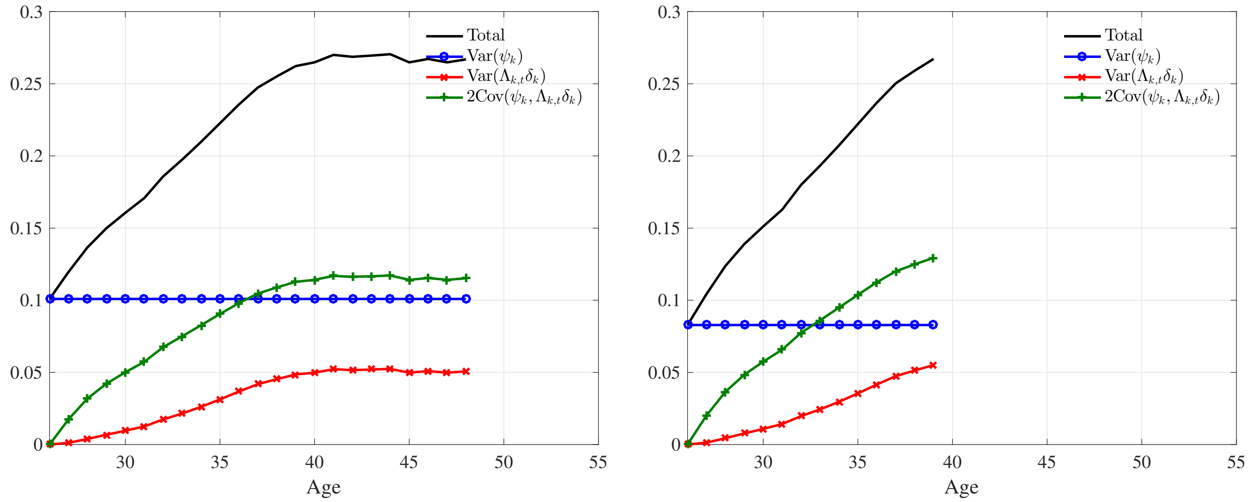
Figure 10: $\text{Var}(\hat{\theta}_{k,t})/\text{Var}(y_{k,t})$ by Age and Cohort



(a) 1965 and 1970 Children

(b) 1975 and 1980 Children

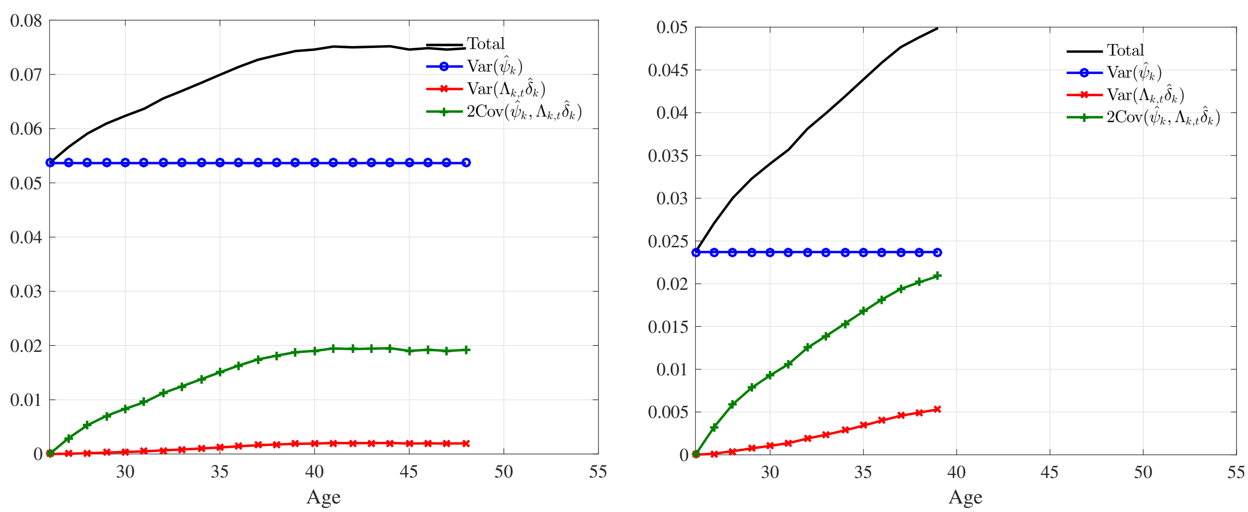
Figure 11: $\text{Var}(\hat{\theta}_{k,t})/\text{Var}(\theta_{k,t})$ by Age and Cohort



(a) 1965 Children, 1940 Parents

(b) 1975 Children, 1940 Parents

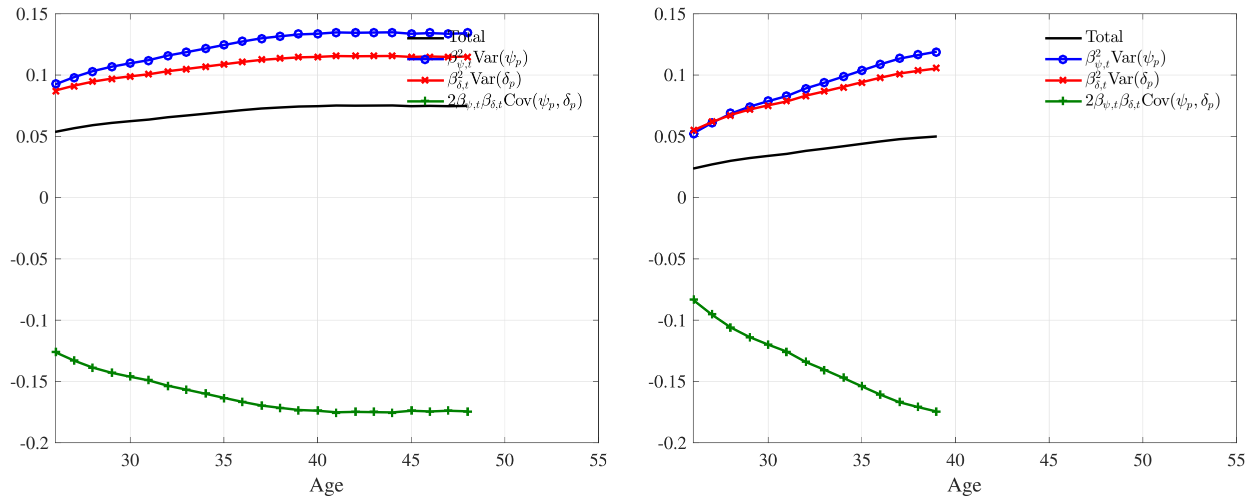
Figure 12: Skill Variance Decomposition by Age for Children



(a) 1965 Children, 1940 Parents

(b) 1975 Children, 1940 Parents

Figure 13: Projected Skill Variance Decomposition by Age for Children



(a) 1965 Children, 1940 Parents

(b) 1975 Children, 1940 Parents

Figure 14: Alternative Projected Skill Variance Decomposition by Age for Children

explain much more of the variation in projected skills. The intergenerational transmission of initial skills is a much more important determinant of skill variation, even at older ages.

Because predicted child skill factors $(\hat{\psi}_k, \hat{\delta}_k)$ are both linear functions of parental skill factors (ψ_p, δ_p), we can also decompose the variance of children's projected skills, $\hat{\theta}_k$, into components related to the two parental skill factors as shown in equation (9). Figure 14 shows that the variation in parental initial skills and skill growth play roughly equal roles in explaining the level and lifecycle growth in children's projected skill variation, where the influence of both parental skill factors comes primarily through children's initial skill levels. Because ψ_p and δ_p are (strongly) negatively correlated, their covariation offsets much of their influence on children's projected skill variation.

These decomposition results establish that both skill factors play important roles in the intergenerational transmission of skills. Another way to see this is to consider the extent to which the variation in children's projected skills can be explained by variation in parental skills alone. In a single-factor model of intergenerational transmission, variation in parental skills fully explains all of the variation in projected skills (i.e., $\mathrm{Var}(\hat{\theta}_{k,t}|\theta_{p,t'}) = 0$); however, this is not the case in our two-factor model.[34] For example, if (ψ_p, δ_p) is joint normally distributed, then $(\hat{\theta}_{k,t}, \theta_{p,t'})$ is also jointly normal, which implies that

$$\mathrm{Var}(\hat{\theta}_{k,t}|\theta_{p,t'}) = \left[1 - \mathrm{Corr}(\hat{\theta}_{k,t}, \theta_{p,t'})^2\right] \mathrm{Var}(\hat{\theta}_{k,t}) = \mathrm{E}\left[\mathrm{Var}(\hat{\theta}_{k,t}|\theta_{p,t'})\right],$$

where the second equality follows from the fact that $\mathrm{Var}(\hat{\theta}_{k,t}|\theta_{p,t'})$ does not depend on the value of $\theta_{p,t'}$. Consequently, assuming joint normality for (ψ_p, δ_p), the fraction of the children's projected skill variance (due to variation in ψ_p and δ_p) conditional on their parents' skill can be easily calculated

[34]The variance of the projected skill can be generally decomposed as $\mathrm{Var}(\hat{\theta}_{k,t}) = \mathrm{Var}(\mathrm{E}[\hat{\theta}_{k,t}|\theta_{p,t'}]) + \mathrm{E}[\mathrm{Var}(\hat{\theta}_{k,t}|\theta_{p,t'})]$. Although $\mathrm{Var}(\hat{\theta}_{k,t}|\psi_p, \delta_p) = 0$ in our two-factor model, $\mathrm{Var}(\hat{\theta}_{k,t}|\theta_{p,t'})$ will not generally equal zero due to the variation in ψ_p and δ_p conditional on $\theta_{p,t'}$.

as

$$\frac{\mathrm{E}\left[\mathrm{Var}(\hat{\theta}_{k,t}|\theta_{p,t'})\right]}{\mathrm{Var}(\hat{\theta}_{k,t})} = 1 - \mathrm{Corr}(\hat{\theta}_{k,t}, \theta_{p,t'})^2.$$

Appendix Figure H11 shows that this fraction, based on average lifetime skills for both parents and children, ranges from 0.47 to 0.67 across different cohorts.[35] This indicates that knowledge of parents' lifetime skill levels explains no more than two-thirds of the variation in children's projected lifetime skills. Put another way, even with perfect information about fathers' expected lifetime earnings — the "Holy Grail" in most of the intergenerational transmission literature — additional information about the trajectory of their lifetime earnings profiles can substantially improve the predictability of children's earnings.

6 Intergenerational Mobility Across Major Canadian Cities

Recent studies have documented considerable variation in intergenerational earnings mobility across countries as well as across geographic regions within a country (Chetty et al., 2014a; Connolly, Corak, and Haeck, 2019; Corak, 2019). This literature has further emphasized the positive correlation between intergenerational earnings persistence (typically measured as IGEs or rank-rank slopes) and cross-sectional earnings inequality across geographic regions/countries (see, e.g., Corak, 2013), often referring to this as the Great Gatsby Curve.

Our analysis points to two potential explanations for variation in intergenerational persistence (or mobility), each with very different economic and policy implications. On one hand, intergenerational earnings persistence may be higher in areas with greater intergenerational skill persistence—an interpretation implicit in much of the literature. On the other hand, intergenerational earnings persistence may be higher in some areas, simply because skills are a more important determinant of earnings in those areas (i.e., these areas may have less earnings instability).[36] The first possibility draws attention to policies related to early childhood, education systems, or neighborhoods and peer effects, while the second focuses attention more on labor markets themselves (e.g., minimum wages, unionization, market flexibility). To better understand these issues, we estimate earnings and skill IGEs across major metropolitan areas in Canada and explore the cross-city relationships between these measures of mobility and cross-sectional measures of earnings and skill inequality.

We assign families to different "cities," as defined by census metropolitan area (CMA) or census agglomeration (CA), based on the family address at the time children and parents were first linked (i.e., when children were ages 16–23).[37] To ensure sufficient sample sizes, we consider families living in

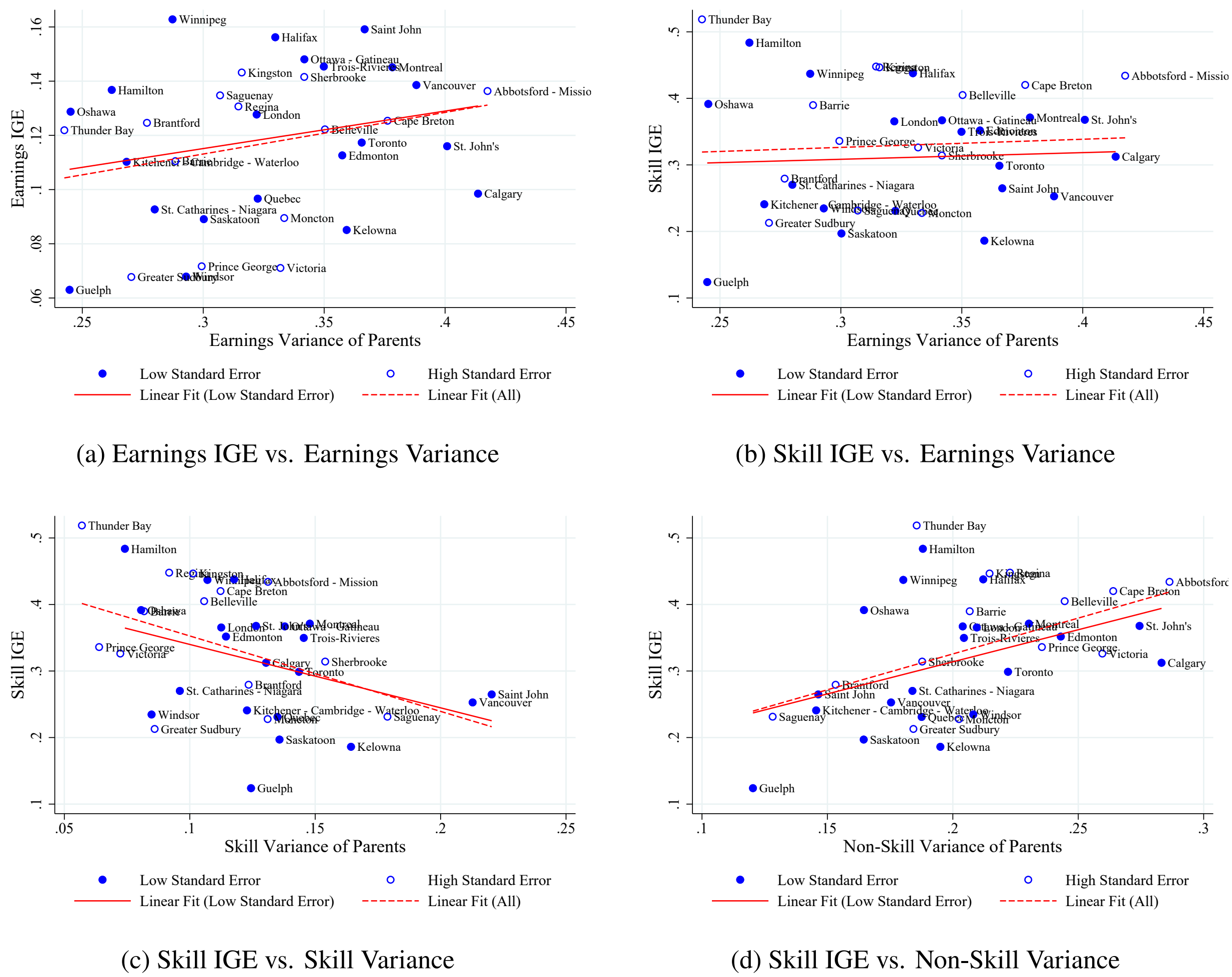
[35]Specifically, we report results using parental skills averaged over ages 26–55 and children's skills averaged over ages 26–48. If we instead consider skills at specific ages, the fraction of child's projected skills (at age 30) explained by parental skills alone (at age 50) is 10–20 percentage points lower, ranging from 0.28 to 0.55 across cohorts.

[36]Based on the single-factor error components model of Haider and Solon (2006), Deutscher and Mazumder (2020) correct regional estimates of rank–rank correlations for lifetime earnings when there are location-specific year-to-year fluctuations in earnings; however, they do not consider the implications of these fluctuations for the Great Gatsby Curve.

[37]A CMA or a CA is a collection of adjacent municipalities around a population center. A CMA must have a total population of at least 100,000 of which 50,000 or more live in the population center. A CA must have a population of at least 10,000 in

the 35 most populous cities. Given the cross-cohort variation in intergenerational mobility identified in Section 5, we focus on children from the 1975 and 1980 cohorts with parents from the 1950 cohort (i.e., family cohort groups 6 and 8) for whom we observe both parents and children for many years.

For each city, we estimate the earnings IGE using OLS. We also use autocovariances for fathers' earnings to estimate their earnings process separately for each city, from which we calculate the skill share of parental earnings variance, $s_{p,t}$.[38] Then, as Equation (1) suggests, we can estimate city-specific skill IGEs by dividing each city's earnings IGE by its skill share of the earnings variance.



(a) Earnings IGE vs. Earnings Variance

(b) Skill IGE vs. Earnings Variance

(c) Skill IGE vs. Skill Variance

(d) Skill IGE vs. Non-Skill Variance

Figure 15: Intergenerational Mobility and Inequality across Canadian Cities: 5-Year Averages

Figure 15 shows several scatterplots characterizing the relationships between earnings/skill IGEs and different measures of cross-sectional inequality. Each dot represents estimates for separate cities and all are based on 5-year average earnings/skill measures for children ages 28–32 and parents ages 48–52. The fact that estimated earnings and skill IGEs in Figure 15 are scattered around their

the population center. Because individuals may move to different regions later in life, there could be regional differences in earnings levels even among those living in the same city during the parent–child linkage year. In Appendix J, we show that controlling for current residential location does not alter our main results.

[38] We use the same autocovariance moments (and equally weighted MD estimation) as in Section 5. Across all cities, we target 7,350 parental autocovariances (210 for each city).

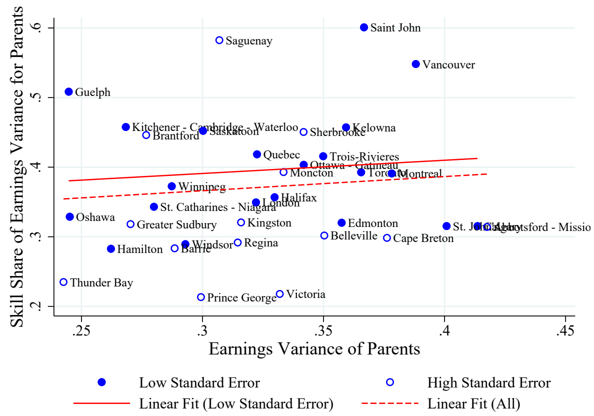


Figure 16: Skill Share of Earnings Variance vs. Earnings Variance: 5-Year Averages

counterparts from the national analysis reported in Table 3 suggests that differences in earnings levels across cities do not play an important role in our analysis of intergenerational earnings and skill mobility at the national level. (See Appendix Figures I15 and I16 for analogous results for annual and 9-year averages.) Because sample sizes are small for some of the smaller cities, some of the estimates (especially estimated skill IGEs) are imprecise.[39] Therefore, we distinguish between those 21 cities with skill IGE standard errors less than 0.12 — "Low Standard Error" cities — represented by filled circles and the remaining "High Standard Error" cities represented by empty circles.[40] The solid lines reflect the best linear fit based on all "Low Standard Error" cities, while the dashed lines reflect the best linear fit for all cities. The similarity of the lines suggests that the relationships between mobility and inequality may not be too distorted by sampling variation of the estimators; however, we explore this issue more formally below.

Variables of Regression		(1) Uncorrected	(2)	(3) Bias-Corrected	(4)		
Dependent	Independent	Estimate	90% CI		Estimate	90% CI	
Earnings IGE	Earnings Variance	0.138	-0.021	0.298	0.158	-0.002	0.318
Skill IGE	Earnings Variance	0.100	-0.468	0.669	0.025	-0.544	0.593
Skill IGE	Skill Variance	-0.953	-1.659	-0.247	-0.715	-1.421	-0.009
Skill IGE	Non-Skill Variance	0.965	0.368	1.562	0.796	0.199	1.393
Skill Share of Earnings Variance	Earnings Variance	0.190	-0.289	0.670	0.251	-0.229	0.730

Notes: Confidence intervals and bias-corrected estimates are based on 200 bootstrap samples.

Table 5: Slope Coefficients Among "Low Standard Error" Cities: 5-Year Averages

Figure 15a shows that cities with higher inequality in parental earnings tend to have less intergenerational earnings mobility as characterized by a higher earnings IGE. This is the relationship depicted

[39]For the largest cities (Toronto and Montreal), the weighted number of families that contributed to each parental autocovariance ranges from 33,950 to 49,940. For the smallest cities (Barrie and Prince George), this range is 1,250 to 1,710.

[40]We use 200 bootstrap samples separately for each city to calculate standard errors and conduct inference. "Low Standard Error" cities include the 15 largest cities, plus 6 other cities spread throughout the next 20 largest cities.

by the Great Gatsby Curve and is largely consistent with the conclusions of Corak (2019), who examines a variety of measures for intergenerational income mobility and inequality across 266 Canadian regions. Figure 15b reveals much greater variation in skill IGEs across cities due to substantial variation in the skill share of the earnings variance for parents, $s_{p,t}$, as observed in Figure 16; however, the relationship between skill IGEs and parental earnings inequality is only slightly weaker than the relationship between earnings IGEs and parental earnings inequality. This can be seen in column (1) of Table 5, which reports the estimated slope coefficients for regressions of earnings or skill IGEs on the variance of parental log earnings for the "Low Standard Error" sample of cities (i.e., the solid lines in Figures 15a and 15b). Based on the 90% confidence intervals (CI's) reported in column (2), we cannot reject that earnings or skill IGEs are uncorrelated with parental earnings inequality. Unfortunately, the relationship between skill IGEs and earnings inequality is very imprecisely estimated.

Figures 15c and 15d examine the relationship between skill IGEs and the variance of the skill and non-skill components of earnings. Both figures suggest much stronger relationships, with skill IGEs declining in the level of skill heterogeneity but increasing just as strongly in variation from the non-skill component of earnings. An important concern, however, is that any estimation error in the decomposition of earnings into skill vs. non-skill components could strongly distort these empirical relationships, since the skill IGE depends on the skill share of the earnings variance. Specifically, any estimation error in $s_{p,t}$ will mechanically generate a negatively biased relationship between the skill IGE and skill variance, while producing a positively biased relationship between the skill IGE and non-skill variance. More generally, estimation errors for both the dependent and independent variables may induce finite-sample bias for all of the coefficients reported in column (1) of Table 5.[41]

Following MacKinnon (2002), we use the bootstrap approach to correct for any finite-sample bias (due to estimation error in both dependent and independent variables) in point estimates and confidence intervals for the slope estimates in Figures 15 and 16.[42] Columns (3) and (4) of Table 5 show the bias-corrected slope estimates and 90% CI's. Because the estimated relationships between the skill IGE and variance of skill and non-skill components of earnings are mechanically affected by correlated sampling errors, it is not surprising that these estimated relationships are quite biased. In both cases, the bias-corrected slope estimates are notably closer to zero but still suggest a strong negative (positive) relationship between the skill IGE and variance of skill (non-skill) earnings component. Both estimates are significantly different from zero (at 0.1 significance level).[43] The estimated biases for all other slope coefficients are quite modest; however, the corrected estimates suggest that the earnings IGE and earnings variance are positively related — consistent with the Great Gatsby Curve — at approximately the 0.1 significance level.

[41]All slope coefficients in Figure 15 and column (1) of Table 5 may be biased for two reasons: (i) the slope coefficient is a non-linear function of the sampling/estimation errors in both the dependent and independent variables and (ii) the sampling/estimation errors are not necessarily mean zero, because the skill-related variables are obtained from non-linear estimation (MD).

[42]Let $\hat{\beta}$ be a slope estimate. For each bootstrap sample $b = 1, 2, \dots, 200$, we calculate the bootstrap version of the slope estimate, β^*_b. Let μ^* and σ^* be the mean and standard deviation of β^*_b over 200 bootstrap samples. Then the bias-corrected point estimate is $2\hat{\beta} - \mu^*$ and the 90% confidence interval is $[2\hat{\beta} - \mu^* - 1.645 \times \sigma^*, 2\hat{\beta} - \mu^* + 1.645 \times \sigma^*]$.

[43]The positive relationship between the skill IGE and variance of the non-skill component of earnings is also significant at the 0.05 level.

Altogether, these results suggest that while intergenerational earnings mobility is (weakly) decreasing in local parental earnings inequality, intergenerational skill mobility is increasing in local skill heterogeneity. Put another way, variation in community skill levels may promote (or reflect) intergenerational mobility, whereas inequality in earnings conditional on skills (i.e., earnings instability) appears to be the driving force for the Great Gatsby phenomenon.

The negative relationship between earnings instability and intergenerational skill mobility could reflect imperfect insurance and/or credit constraints, as uninsured earnings shocks discourage human capital investment in children whose parents are lower income/skill relative to those with higher income/skills (e.g., Carneiro and Ginja, 2016; Caucutt and Lochner, 2020). The influence of credit/insurance market frictions on the relationship between local skill variation and intergenerational skill mobility is less clear.[44] One potential explanation for the positive skill mobility – inequality relationship is that greater inequality in school peers may benefit lower ability children relative to higher ability children as suggested by the estimates of Imberman, Kugler, and Sacerdote (2012); however, there are very few estimates of and no consensus on the nonlinear structure of peer effects (see Sacerdote, 2014, for a review). Additionally, greater inequality at the city level need not imply greater inequality within schools due to local geographic income segregation. There is recent evidence that local spending on public schools appears to increase with levels of income inequality at the city level (Corcoran and Evans, 2010; Boustan et al., 2013); yet, these studies do not distinguish between inequality driven by skill differences vs. other idiosyncratic shocks. More generally, the economics literature has made little effort to separately analyze the distinct relationships between intergenerational mobility and different sources of income inequality.

7 Conclusions

The empirical literature on intergenerational earnings mobility often interprets its findings in the context of intergenerational skill transmission; yet, it rarely makes this connection explicit. This paper develops a two-factor model of the intergenerational transmission of initial skills and skill growth rates to explicitly study the intergenerational mobility of skills, as distinct from earnings. We show that skills, loosely defined, are an important component of earnings but that other idiosyncratic factors (some moderately persistent) also play an important role in determining earnings not only at a single point in time but also over a worker's entire career. This creates an important distinction between IGEs for earnings and skills, where the latter are always larger due to variability in non-skill components of earnings.

Estimating our model using 37 years of administrative tax data on fathers and sons from Canada, we show that both factors of intergenerational skill transmission are needed to explain several prominent patterns in intergenerational covariances of earnings—the traditional single-factor model is inadequate. Based on these patterns, we estimate important heterogeneity in both initial skill levels and skill growth

[44] While differences in these frictions across communities could lead to a positive relationship between skill mobility and inequality (e.g., better credit markets could lead to a more efficient allocation of skill investments across children, which could improve intergenerational mobility while increasing inequality in skills), differences in credit/insurance are likely to be quite small across Canadian cities. That said, such differences could play a more important role in cross-country comparisons.

rates, where interpersonal differences in skill growth decline over the careers of children in our sample (as human capital theory predicts). As a simple metric for characterizing the importance of both skill factors, we show that knowledge of father's expected lifetime earnings explains no more than two-thirds of the variation in children's projected lifetime skills. Consequently, while knowledge of parents' initial skill and skill growth rates together explain 20–40% of the variation in children's skills (10–20% of the variation in their earnings) for most cohorts, knowledge of average lifetime parental skill levels alone would explain substantially less.

Our two-factor model produces several novel insights about the evolution of skills over the lifecycle and the influence of intergenerational skill transmission for skill and earnings trajectories. For example, we show that while skills become a more important driver of earnings inequality over the first 10–15 years of workers' careers, the intergenerational transmission of skills becomes a less important determinant of skills themselves. Our estimates also reveal that intergenerational transmission of children's initial skills (compared to skill growth) is a more important determinant of skill variation, even at older ages when differences in skill growth rates have their greatest influence. Thus, what parents pass onto their children, through nature or nurture, seems to have its greatest impact on their children's early-career skills more than through their ability to accumulate additional skills in the labor market. This would not be surprising if learning abilities are primarily determined by nature, whereas early-career skill levels are influenced by both nature and nurture through more than two decades of family and educational investments. Still, we find that parents' initial skills and skill growth rates are equally important for children's skill levels and lifecycle trajectories, largely because both factors are important determinants of children's initial skills.

Like others (Corak and Heisz, 1999; Chen, Ostrovsky, and Piraino, 2017), we estimate low earnings IGEs in Canada relative to the United States, with IGEs based on 5- or 9-year average earnings less than 0.2 for all cohorts we study. Due to considerable variation in non-skill earnings components, skill IGEs are generally higher by a factor of 2 to 3, with skill IGEs exceeding 0.4 for the earliest cohort of sons we study. Our estimates suggest that intergenerational mobility has improved in Canada with skill IGEs falling by more than half between the 1965 and 1980 cohorts of sons (with parents from the 1940 birth cohort). We also find that skill (and earnings) IGEs are larger for sons with older fathers, suggesting that parental age at birth may play an important role in skill transmission. At a minimum, researchers comparing IGEs across cohorts should be careful to account for differences in parents' ages when their children are born. We obtain similar cross-cohort patterns for our measure of intergenerational transmission based on the fraction of the children's earnings variation that can be explained by their projected skills.

Finally, we use our data to study IGEs across large Canadian cities. Given our sample of only 35 cities, this analysis is exploratory in nature and provides a guide for future work on the geography of intergenerational mobility. Consistent with the well-documented Great Gatsby Curve, we estimate a positive relationship between earnings IGEs and the variance of parental earnings across cities: earnings mobility is decreasing in earnings inequality. Digging deeper, we find that cities with greater parental skill heterogeneity exhibit higher levels of intergenerational mobility, while cities characterized by greater earnings instability exhibit less mobility. The latter raises natural concerns about distortions due to imperfect credit and insurance markets. The former is not so readily understood, in part because the literature has not generally considered the distinct relationships between intergenerational mobility and different sources of income inequality. Our findings, coupled with those of Landersø and Heckman (2017), suggest that a deeper exploration of these differences (both theoretically and empirically) is warranted.

References

Acemoglu, Daron and David Autor. 2011. "Skills, tasks and technologies: Implications for employment and earnings." In *Handbook of Labor Economics*, vol. 4B. Elsevier, 1043–1171.

Altonji, Joseph G. and Thomas A. Dunn. 1991. "Relationships among the Family Incomes and Labor Market Outcomes of Relatives." *Research in Labor Economics* 12:269–310.

Aydemir, Abdurrahman, Wen-Hao Chen, and Miles Corak. 2013. "Intergenerational Education Mobility among the Children of Canadian Immigrants." *Canadian Public Policy* 39 (Supplement 1):S107–S122.

Bagger, Jesper, Fran Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin. 2014. "Tenure, experience, human capital, and wages: A tractable equilibrium search model of wage dynamics." *American Economic Review* 104 (6):1551–96.

Baker, Michael. 1997. "Growth-rate heterogeneity and the covariance structure of life-cycle earnings." *Journal of Labor Economics* 15 (2):338–375.

Becker, Gary S. 1975. *Human Capital*. New York, NY: Columbia University Press.

Becker, Gary S. and Nigel Tomes. 1979. "An equilibrium theory of the distribution of income and intergenerational mobility." *Journal of Political Economy* 87 (6):1153–1189.

———. 1986. "Human capital and the rise and fall of families." *Journal of Labor Economics* 4 (3, Part 2):S1–S39.

Ben-Porath, Yoram. 1967. "The Production of Human Capital and the Life Cycle of Earnings." *Journal of Political Economy* 75 (4, Part 1):352–365.

Boudarbat, Brahim, Thomas Lemieux, and W. Craig Riddell. 2010. "The evolution of the returns to human capital in Canada, 1980–2005." *Canadian Public Policy* 36 (1):63–89.

Boustan, Leah, Fernando Ferreira, Hernan Winkler, and Eric M Zolt. 2013. "The effect of rising income inequality on taxation and public expenditures: Evidence from US municipalities and school districts, 1970–2000." *Review of Economics and Statistics* 95 (4):1291–1302.

Bowlus, Audra J and Huju Liu. 2013. "The contributions of search and human capital to earnings growth over the life cycle." *European Economic Review* 64:305–331.

Carneiro, Pedro and Rita Ginja. 2016. "Partial Insurance and Investments in Children." *The Economic Journal* 126 (596):F66–F95.

Carneiro, Pedro, Karsten T. Hansen, and James J. Heckman. 2003. "Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice." *International Economic Review* 44 (2):361–422.

Caucutt, Elizabeth M. and Lance Lochner. 2020. "Early and late human capital investments, borrowing constraints, and the family." *Journal of Political Economy* 128 (3):1065–1147.

Cawley, John, James Heckman, and Edward Vytlacil. 1999. "Meritocracy in America: Wages within and across Occupations." *Industrial Relations* 38 (3):250–296.

Chen, Wen-Hao, Yuri Ostrovsky, and Patrizio Piraino. 2017. "Lifecycle variation, errors-in-variables bias and nonlinearities in intergenerational income transmission: New evidence from Canada." *Labour Economics* 44 (Supplement C):1–12.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014a. "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics* 129 (4):1553–1623.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. 2014b. "Is the United States Still a Land of Opportunity? Recent Trends in Intergenerational Mobility." *American Economic Review* 104 (5):141–47.

Connolly, Marie, Miles Corak, and Catherine Haeck. 2019. "Intergenerational Mobility between and within Canada and the United States." *Journal of Labor Economics* 37 (S2):S595–S641.

Corak, Miles. 2013. "Income inequality, equality of opportunity, and intergenerational mobility." *Journal of Economic Perspectives* 27 (3):79–102.

———. 2019. "The Canadian Geography of Intergenerational Income Mobility." *The Economic Journal* 130 (631):2134–2174.

Corak, Miles and Andrew Heisz. 1999. "The Intergenerational Earnings and Income Mobility of Canadian Men: Evidence from Longitudinal Income Tax Data." *The Journal of Human Resources* 34 (3):504–533.

Corcoran, Sean and William N Evans. 2010. "Income inequality, the median voter, and the support for public education." NBER Working Paper No. 16097.

Cunha, Flavio. 2013. "Investments in Children When Markets are Incomplete." Working Paper, University of Pennsylvania.

Cunha, Flavio and James Heckman. 2007. "The technology of skill formation." *American Economic Review* 97 (2):31–47.

Cunha, Flavio, James Heckman, and Salvador Navarro. 2005. "Separating uncertainty from heterogeneity in life cycle earnings." *Oxford Economic papers* 57 (2):191–261.

Cunha, Flavio and James J. Heckman. 2008. "Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation." *Journal of Human Resources* 43 (4):738–782.

Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2006. "Interpreting the evidence on life cycle skill formation." In *Handbook of the Economics of Education*, vol. 1, edited by Eric A Hanushek and Finis Welch, chap. 12. Amsterdam: North Holland: Elsevier, 697–812.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3):883–931.

Del Boca, Daniela, Christopher Flinn, and Matthew Wiswall. 2014. "Household Choices and Child Development." *Review of Economic Studies* 81 (1):137–185.

Deming, David. 2017. "The growing importance of social skills in the labor market." *Quarterly Journal of Economics* 132 (4):1593–1640.

Deutscher, Nathan and Bhashkar Mazumder. 2020. "Intergenerational mobility across Australia and the stability of regional estimates." *Labour Economics* 66:101861.

Fortin, Nicole and Thomas Lemieux. 2015. "Changes in wage inequality in Canada: An interprovincial perspective." *Canadian Journal of Economics* 48 (2):682–713.

Gallipoli, Giovanni, Hamish Low, and Aruni Mitra. 2020. "Consumption and Income Inequality across Generations." Working Paper.

Gayle, George-Levi, Limor Golan, and Mehmet A. Soytas. 2015. "What Accounts for the Racial Gap in Time Allocation and Intergenerational Transmission of Human Capital?" Working Paper.

Grawe, Nathan D. 2006. "Lifecycle bias in estimates of intergenerational earnings persistence." *Labour Economics* 13 (5):551–570.

Green, David A and Benjamin M Sand. 2015. "Has the Canadian Labour Market Polarized?" *Canadian Journal of Economics* 48 (2):612–646.

Guvenen, Fatih. 2007. "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?" *American Economic Review* 97 (3):687–712.

———. 2009. "An empirical investigation of labor income processes." *Review of Economic Dynamics* 12 (1):58–79.

Haider, Steven and Gary Solon. 2006. "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review* 96 (4):1308–1320.

Haider, Steven J. 2001. "Earnings Instability and Earnings Inequality of Males in the United States: 1967–1991." *Journal of Labor Economics* 19 (4):799–836.

Hansen, Bruce. 2021. "Econometrics." https://www.ssc.wisc.edu/~bhansen/econometrics/Econometrics.pdf. (accessed: May 12, 2021).

Hause, John C. 1980. "The fine structure of earnings and the on-the-job training hypothesis." *Econometrica* 48 (4):1013–1029.

Heckman, James J. 1976. "A life-cycle model of earnings, learning, and consumption." *Journal of Political Economy* 84 (4, Part 2):S9–S44.

Heckman, James J., Lance Lochner, and Christopher Taber. 1998. "Explaining rising wage inequality: Explorations with a dynamic general equilibrium model of labor earnings with heterogeneous agents." *Review of Economic Dynamics* 1 (1):1–58.

Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond." In *Handbook of the Economics of Education*, vol. 1, edited by Eric Hanushek and Finis Welch, chap. 7. Elsevier, 307–458.

———. 2008. "Earnings functions and rates of return." *Journal of Human Capital* 2 (1):1–31.

Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior." *Journal of Labor Economics* 24 (3):411–482.

Hryshko, Dmytro. 2012. "Labor income profiles are not heterogeneous: Evidence from income growth rates." *Quantitative Economics* 3 (2):177–209.

Imberman, Scott A, Adriana D Kugler, and Bruce I Sacerdote. 2012. "Katrina's children: Evidence on the structure of peer effects from hurricane evacuees." *American Economic Review* 102 (5):2048–82.

Jenkins, Stephen. 1987. "Snapshots versus movies: 'Lifecycle biases' and the estimation of intergenerational earnings inheritance." *European Economic Review* 31 (5):1149–1158.

Karahan, Fatih, Serdar Ozkan, and Jae Song. 2019. "Anatomy of lifetime earnings inequality: Heterogeneity in job ladder risk vs. human capital." *FRB of New York Staff Report* (908).

Landersø, Rasmus and James J. Heckman. 2017. "The Scandinavian fantasy: The sources of intergenerational mobility in Denmark and the US." *The Scandinavian Journal of Economics* 119 (1):178–230.

Lee, Sang Yoon and Ananth Seshadri. 2019. "On the intergenerational transmission of economic status." *Journal of Political Economy* 127 (2):855–921.

Lillard, Lee A. and Yoram Weiss. 1979. "Components of variation in panel earnings data: American scientists 1960-70." *Econometrica* 47 (2):437–454.

Lochner, Lance, Youngmin Park, and Youngki Shin. 2018. "Wage Dynamics and Returns to Unobserved Skill." NBER Working Paper No. 24220.

Loury, Glenn C. 1981. "Intergenerational transfers and the distribution of earnings." *Econometrica* 49 (4):843–867.

MacKinnon, James G. 2002. "Bootstrap inference in econometrics." *Canadian Journal of Economics* 35 (4):615–645.

MaCurdy, Thomas E. 1982. "The use of time series processes to model the error structure of earnings in a longitudinal data analysis." *Journal of Econometrics* 18 (1):83–114.

Mazumder, Bhashkar. 2005. "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data." *The Review of Economics and Statistics* 87 (2):235–255.

Meghir, Costas and Luigi Pistaferri. 2011. "Earnings, consumption and life cycle choices." In *Handbook of Labor Economics*, vol. 4, edited by David Card and Orley Ashenfelter, chap. 9. Elsevier, 773–854.

Mincer, Jacob. 1974. *Schooling, Experience and Earnings*. New York: Columbia University Press.

Nybom, Martin and Jan Stuhler. 2016. "Heterogeneous Income Profiles and Lifecycle Bias in Intergenerational Mobility Estimation." *Journal of Human Resources* 51 (1):239–268.

Postel-Vinay, Fabien and Helene Turon. 2010. "On-the-job search, productivity shocks, and the individual earnings process." *International Economic Review* 51 (3):599–629.

Sacerdote, Bruce. 2014. "Experimental and quasi-experimental analysis of peer effects: Two steps forward?" *Annual Review of Economics* 6 (1):253–272.

Shaw, Kathryn L. 1989. "Life-cycle labor supply with human capital accumulation." *International Economic Review* 30 (2):431–456.

Solon, Gary. 1992. "Intergenerational Income Mobility in the United States." *American Economic Review* 82 (3):393–408.

Statistics Canada. 2017. *Intergenerational Income Database*. Ottawa, ON: Social Analysis and Modelling Division, Statistics Canada.

Zimmerman, David J. 1992. "Regression Toward Mediocrity in Economic Stature." *American Economic Review* 82 (3):409–429.

Appendix

A Identification

In this appendix, we consider a more general MA(q) process for the transitory component of earnings:

$$\varepsilon_{i,j,t} = \phi_{i,j,t} + \sum_{l=0}^{\min\{q,t-\underline{t}\}} \kappa_{j,l}\xi_{i,j,t-l},$$

where we normalize $\kappa_{j,0} = 1$ for $j \in \{p, k\}$, and all other model features are the same as in the text.

We now make the conditioning on parents' and children's cohorts explicit because we exploit cross-cohort variation in covariances. Let $c_{i,j}$ be the birth year, or "cohort" of the individual (i, j). Individuals are observed from age $t = \underline{t}$ until age $t = \bar{t}$.

Notice that

$$\begin{aligned}
y_{i,j,t} =& \psi_{i,j} + \Lambda_{j,t}\delta_{i,j} + \phi_{i,j,t} + \sum_{l=0}^{\min\{q,t-\underline{t}\}} \kappa_{j,l}\xi_{i,j,t-l}, \\
\Delta y_{i,j,t} =& \lambda_{j,t}\delta_{i,j} + \Delta\phi_{i,j,t} + \sum_{l=0}^{\min\{q,t-\underline{t}\}} \kappa_{j,l}\Delta\xi_{i,j,t-l}, \\
\Delta y_{i,j,t} - \rho_j \Delta y_{i,j,t-1} =& (\lambda_{j,t} - \rho_j\lambda_{j,t-1})\delta_{i,j} + \Delta\nu_{i,j,t} + \sum_{l=0}^{\min\{q,t-\underline{t}\}} \kappa_{j,l}(\Delta\xi_{i,j,t-l} - \rho_j\Delta\xi_{i,j,t-1-l}), \\
y_{i,j,t} - \rho_j y_{i,j,t-1} =& (1-\rho_j)\psi_{i,j} + (\lambda_{j,t} - \rho_j\lambda_{j,t-1})\delta_{i,j} + \nu_{i,j,t} + \sum_{l=0}^{\min\{q,t-\underline{t}\}} \kappa_{j,l}(\xi_{i,j,t-l} - \rho_j\xi_{i,j,t-1-l}).
\end{aligned}$$

Identification of ρ_j For (t, t') such that $t' - t > q + 2$,

$$\mathrm{Cov}\left(\Delta y_{j,t}, \Delta y_{j,t'} - \rho_j \Delta y_{j,t'-1} | c_j\right) = \lambda_{j,t}(\lambda_{j,t'} - \rho_j\lambda_{j,t'-1})\,\mathrm{Var}(\delta_j|c_j).$$

For any l,

$$\mathrm{Cov}\left(\Delta y_{j,t+l}, \Delta y_{j,t'+l} - \rho_j \Delta y_{j,t'+l-1} | c_j + l\right) = \lambda_{j,t}(\lambda_{j,t'} - \rho_j\lambda_{j,t'-1})\,\mathrm{Var}(\delta_j|c_j + l).$$

Therefore,

$$\frac{\mathrm{Cov}\left(\Delta y_{j,t}, \Delta y_{j,t'} - \rho_j \Delta y_{j,t'-1} | c_j\right)}{\mathrm{Cov}\left(\Delta y_{j,t+l}, \Delta y_{j,t'+l} - \rho_j \Delta y_{j,t'+l-1} | c_j + l\right)} = \frac{\mathrm{Var}(\delta_j|c_j)}{\mathrm{Var}(\delta_j|c_j + l)}$$

Similarly, for (t'', t''') and l such that $t''' - t'' > q + 2$,

$$\frac{\mathrm{Cov}\left(\Delta y_{j,t''}, \Delta y_{j,t'''} - \rho_j \Delta y_{j,t'''-1} | c_j\right)}{\mathrm{Cov}\left(\Delta y_{j,t'+l}, \Delta y_{j,t'''+l} - \rho_j \Delta y_{j,t'''+l-1} | c_j + l\right)} = \frac{\mathrm{Var}(\delta_j|c_j)}{\mathrm{Var}(\delta_j|c_j + l)}.$$

Therefore, ρ_j is identified from the following:

$$\frac{\text{Cov}\left(\Delta y_{j,t}, \Delta y_{j,t'} - \rho_j \Delta y_{j,t'-1} | c_j\right)}{\text{Cov}\left(\Delta y_{j,t+l}, \Delta y_{j,t'+l} - \rho_j \Delta y_{j,t'+l-1} | c_j + l\right)} = \frac{\text{Cov}\left(\Delta y_{j,t''}, \Delta y_{j,t'''} - \rho_j \Delta y_{j,t'''-1} | c_j\right)}{\text{Cov}\left(\Delta y_{j,t'+l}, \Delta y_{j,t'''+l} - \rho_j \Delta y_{j,t'''+l-1} | c_j + l\right)}.$$

Identification of $\lambda_{j,t}$ For (t, t') such that $t' - t > q + 2$,

$$\text{Cov}\left(\Delta y_{j,t}, \Delta y_{j,t'} - \rho_j \Delta y_{j,t'-1} | c_j\right) = \lambda_{j,t} (\lambda_{j,t'} - \rho_j \lambda_{j,t'-1}) \text{Var}(\delta_j | c_j), \tag{10}$$

$$\text{Cov}\left(\Delta y_{j,t-1}, \Delta y_{j,t'} - \rho_j \Delta y_{j,t'-1} | c_j\right) = \lambda_{j,t-1} (\lambda_{j,t'} - \rho_j \lambda_{j,t'-1}) \text{Var}(\delta_j | c_j), \tag{11}$$

$$\text{Cov}\left(\Delta y_{j,t}, \Delta y_{j,t'-1} - \rho_j \Delta y_{j,t'-2} | c_j\right) = \lambda_{j,t} (\lambda_{j,t'-1} - \rho_j \lambda_{j,t'-2}) \text{Var}(\delta_j | c_j). \tag{12}$$

From (10) and (11), we have

$$\frac{\text{Cov}\left(\Delta y_{j,t}, \Delta y_{j,t'} - \rho_j \Delta y_{j,t'-1} | c_j\right)}{\text{Cov}\left(\Delta y_{j,t-1}, \Delta y_{j,t'} - \rho_j \Delta y_{j,t'-1} | c_j\right)} = \frac{\lambda_{j,t}}{\lambda_{j,t-1}},$$

from which we can identify $\lambda_{j,t}$ for all $t = \underline{t}+2, \underline{t}+3, \ldots, \bar{t} - q - 3$, given the normalization $\lambda_{j,\underline{t}+1} = 1$.

Next, from (10) and (12), we have

$$\frac{\text{Cov}\left(\Delta y_{j,t}, \Delta y_{j,t'} - \rho_j \Delta y_{j,t'-1} | c_j\right)}{\text{Cov}\left(\Delta y_{j,t}, \Delta y_{j,t'-1} - \rho_j \Delta y_{j,t'-2} | c_j\right)} = \frac{\lambda_{j,t'} - \rho_j \lambda_{j,t'-1}}{\lambda_{j,t'-1} - \rho_j \lambda_{j,t'-2}},$$

which identifies $\lambda_{j,t}$, given that $\lambda_{j,t-1}$ and $\lambda_{j,t-2}$ are already identified. Therefore, $\lambda_{j,t}$ for $t = \bar{t} - q - 2, \bar{t} - q - 1, \ldots, \bar{t}$ can be identified if $\lambda_{j,t}$ for $t = \underline{t}, \underline{t}+1, \ldots, \bar{t} - q - 4, \bar{t} - q - 3$ are identified. This requires that $\bar{t} - \underline{t} \geq q + 4$.

Identification of Var$(\delta_j | c_j)$ Once ρ_j and $\lambda_{j,t}$ are identified, Var$(\delta_j | c_j)$ for all j and c_j can be identified from (10).

Identification of Cov$(\psi_j, \delta_j | c_j)$ For (t, t') such that $t' - t > q + 2$,

$$\text{Cov}\left(y_{j,t}, \Delta y_{j,t'} - \rho_j \Delta y_{j,t'-1} | c_j\right) = (\lambda_{j,t} - \rho_j \lambda_{j,t-1}) \left[\text{Cov}(\psi_j, \delta_j | c_j) + \Lambda_{j,t} \text{Var}(\delta_j | c_j)\right].$$

Identification of Var$(\psi_j | c_j)$ For (t, t') such that $t' - t > q + 1$,

$$\begin{aligned}\text{Cov}\left(y_{j,t}, y_{j,t'} - \rho_j y_{j,t'-1} | c_j\right) =& (1 - \rho_j) \text{Var}(\psi_j | c_j) + \Lambda_{j,t} (\Lambda_{j,t'} - \rho_j \Lambda_{j,t'-1}) \text{Var}(\delta_j | c_j) \\ &+ \left[(1 - \rho_j)\Lambda_{j,t} + \Lambda_{j,t'} - \rho_j \Lambda_{j,t'-1}\right] \text{Cov}(\psi_j, \delta_j | c_j).\end{aligned}$$

Identification of Var$(\phi_j | c_j)$ Notice that, for $t' > t$,

$$\phi_{i,j,t'} = \rho_j^{t'-t} \phi_{i,j,t} + \sum_{l=0}^{t'-t-1} \rho_j^l \nu_{i,j,t'-l}$$

Then, for $t' - t > q$,

$$\begin{aligned}\operatorname{Cov}(y_{j,t}, y_{j,t'}|c_j) = &\operatorname{Var}(\psi_j|c_j) + \Lambda_{j,t}\Lambda_{j,t'} \operatorname{Var}(\delta_j|c_j) \\ &+ (\Lambda_{j,t} + \Lambda_{j,t'}) \operatorname{Cov}(\psi_j, \delta_j|c_j) + \rho_j^{t'-t} \operatorname{Var}(\phi_{j,t}|c_j).\end{aligned}$$

Thus, $\operatorname{Var}(\phi_{j,t}|c_j)$ is identified for all but the last $q+1$ periods of data.

Identification of Var$(\nu_{j,t}|c_j)$ If $\operatorname{Var}(\phi_{j,t}|c_j)$ and $\operatorname{Var}(\phi_{j,t-1}|c_j)$ are identified,

$$\operatorname{Var}(\nu_{j,t}|c_j) = \operatorname{Var}(\phi_{j,t}|c_j) - \rho_j^2 \operatorname{Var}(\phi_{j,t-1}|c_j).$$

Identification of $\kappa_{j,l}$ **and Var**$(\xi_{j,t}|c_j)$ Consider (j, c_j) for which $\operatorname{Var}(\phi_{j,\underline{t}}|c_j)$ is identified. Then,

$$\operatorname{Var}(y_{j,\underline{t}}|c_j) = \operatorname{Var}(\psi_j|c_j) + \operatorname{Var}(\phi_{j,\underline{t}}|c_j) + \operatorname{Var}(\xi_{j,\underline{t}}|c_j),$$

from which $\operatorname{Var}(\xi_{j,\underline{t}}|c_j)$ is identified.

Moreover, for $l \in \{1, 2, \dots, q\}$,

$$\operatorname{Cov}(y_{j,\underline{t}}, y_{j,\underline{t}+l}|c_j) = \operatorname{Var}(\psi_j|c_j) + l \operatorname{Cov}(\psi_j, \delta_j|c_j) + \rho_j^l \operatorname{Var}(\phi_{j,\underline{t}}|c_j) + \kappa_{j,l} \operatorname{Var}(\xi_{j,\underline{t}}|c_j),$$

which gives $\kappa_{j,l}$.

Finally, for $t > \underline{t}$,

$$\begin{aligned}\operatorname{Var}(y_{j,t}|c_j) = &\operatorname{Var}(\psi_j|c_j) + \Lambda_{j,t} \operatorname{Var}(\delta_j|c_j) + 2\Lambda_{j,t} \operatorname{Cov}(\psi_j, \delta_j|c_j) + \operatorname{Var}(\phi_{j,t}|c_j) \\ &+ \sum_{l=0}^{\min\{q, t-\underline{t}\}} \kappa_{j,l}^2 \operatorname{Var}(\xi_{j,t-l}|c_j).\end{aligned}$$

Thus, $\operatorname{Var}(\xi_{j,t}|c_j)$ is identified as long as $\operatorname{Var}(\phi_{j,t}|c_j)$ is identified.

From now on, consider (j, j') such that $j \neq j'$.

Identification of Cov$(\delta_j, \delta_{j'}|c_j, c_{j'})$

$$\operatorname{Cov}(\Delta y_{j,t}, \Delta y_{j',t'}|c_j, c_{j'}) = \lambda_{j,t}\lambda_{j',t'} \operatorname{Cov}(\delta_j, \delta_{j'}|c_j, c_{j'}).$$

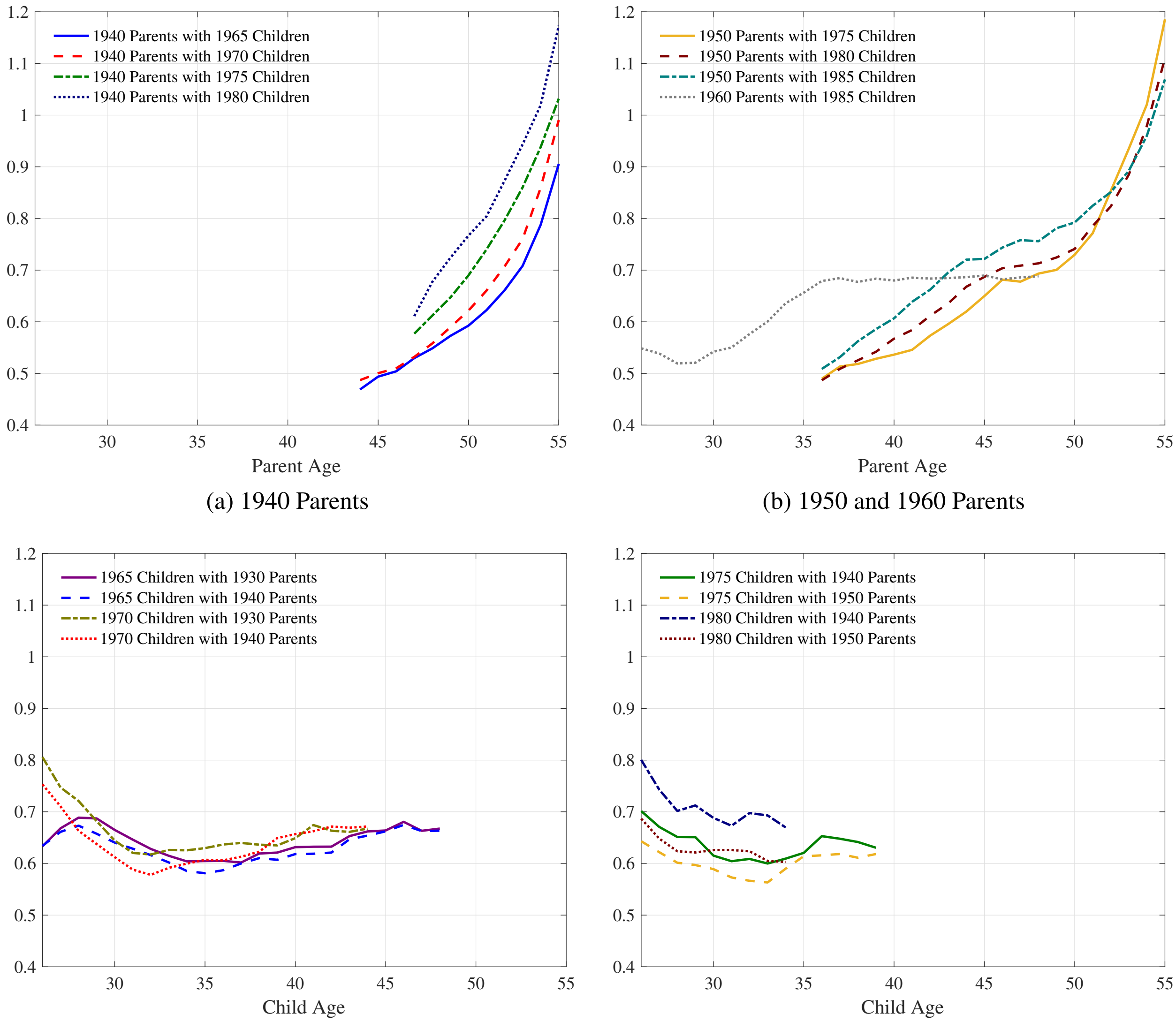
Identification of Cov$(\psi_j, \delta_{j'}|c_j, c_{j'})$

$$\operatorname{Cov}(y_{j,t}, \Delta y_{j',t'}|c_j, c_{j'}) = \lambda_{j',t'} \big[\operatorname{Cov}(\psi_j, \delta_{j'}|c_j, c_{j'}) + \Lambda_{j,t} \operatorname{Cov}(\delta_j, \delta_{j'}|c_j, c_{j'})\big].$$

Identification of Cov$(\psi_j, \psi_{j'} | c_j, c_{j'})$

$$\begin{aligned}\mathrm{Cov}(y_{j,t}, y_{j',t'} | c_j, c_{j'}) = &\mathrm{Cov}(\psi_j, \psi_{j'} | c_j, c_{j'}) + \Lambda_{j,t} \Lambda_{j',t'} \mathrm{Cov}(\delta_j, \delta_{j'} | c_j, c_{j'}) \\ &+ \Lambda_{j,t} \mathrm{Cov}(\delta_j, \psi_{j'} | c_j, c_{j'}) + \Lambda_{j',t'} \mathrm{Cov}(\psi_j, \delta_{j'} | c_j, c_{j'}).\end{aligned}$$

B Additional Covariances (National Sample)



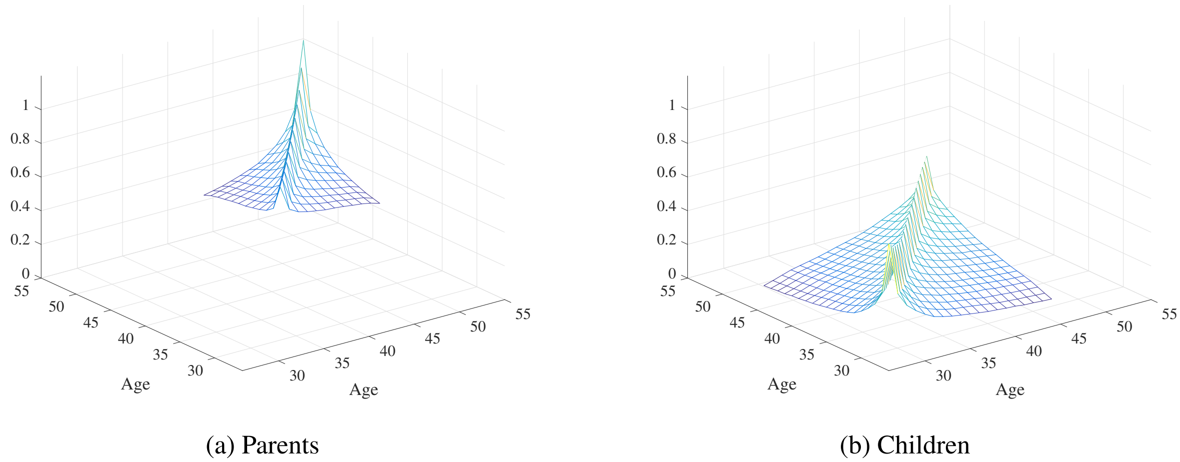
(a) 1940 Parents

(b) 1950 and 1960 Parents

(c) 1965 and 1970 Children

(d) 1975 and 1980 Children

Figure B1: Variance of Earnings by Cohort



(a) Parents

(b) Children

Figure B2: Autocovariances for 1970 Children and 1940 Parents

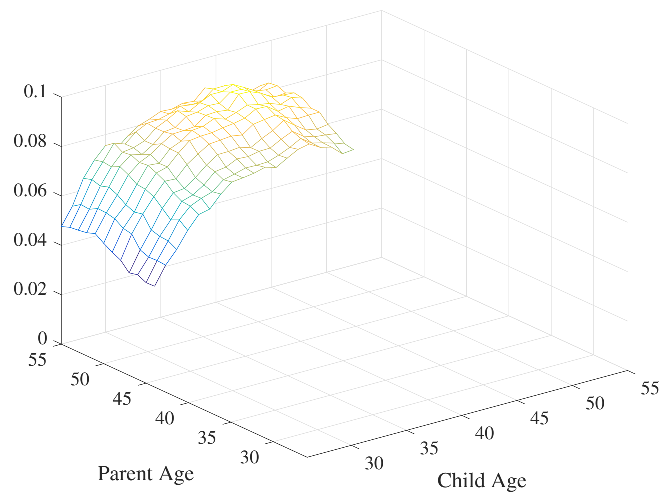


Figure B3: Intergenerational Covariances for 1970 Children and 1940 Parents

C Testing for Zero Intergenerational Covariances

In Section 3, we show patterns for intergenerational covariances of earnings that are inconsistent with a one-factor model of skill: earnings growth is uncorrelated across generations, while the earnings level of each generation is correlated with the earnings growth of the other. In this appendix, we formally test these patterns based on 100 bootstrap samples. Specifically, we conduct bootstrap tests for the following 3 types of hypotheses:

$$\begin{aligned}
\mathrm{Cov}(y_{p,t}, \Delta y_{k,t'}|g) =& 0, \quad \forall (g, t, t'), \\
\mathrm{Cov}(\Delta y_{p,t}, y_{k,t'}|g) =& 0, \quad \forall (g, t, t'), \\
\mathrm{Cov}(\Delta y_{p,t}, \Delta y_{k,t'}|g) =& 0, \quad \forall (g, t, t'),
\end{aligned}$$

where g is the index for family cohort group (defined in Section 4).

Let $\boldsymbol{\beta} = (\beta_l)_{l=1}^{L}$ be a vector of covariances that we wish to test. For example, for the first type of hypotheses, each element of $\boldsymbol{\beta}$ is $\mathrm{Cov}(y_{p,t}, \Delta y_{k,t'}|g)$ for some (g, t, t'). Let $\hat{\boldsymbol{\beta}}$ be the point estimate of $\boldsymbol{\beta}$ (i.e., sample covariances) and $\boldsymbol{\beta}_b^* = (\beta_{b,l}^*)_{l=1}^{L}$ be the corresponding estimate from a bootstrap sample $b = 1, 2, \ldots, B$.

We calculate boostrap p-values for individual and joint hypotheses as explained in Chapter 10 of Hansen (2021). For a single hypothesis $H_0 : \beta_l = 0$, the bootstrap p-value is calculated based on the non-studentized t-statistic:

$$\frac{1}{B}\sum_{b=1}^{B}\mathbb{I}\Big(\big|\beta_{b,l}^* - \hat{\beta}_l\big| > \big|\hat{\beta}_l\big|\Big),$$

where $\mathbb{I}(x)$ is the indicator function that equals 1 if the statement x is true and zero otherwise.

For joint hypotheses $H_0 : \boldsymbol{\beta} = \mathbf{0}$, we conduct a Wald-type bootstrap test based on the statistic

$$W = \hat{\boldsymbol{\beta}}^\top \boldsymbol{V}^{-1} \hat{\boldsymbol{\beta}}.$$

The weighting matrix $\boldsymbol{V}$ can be any positive-definite matrix, and we use an identity matrix. For each bootstrap sample b, the Wald statistic is

$$W_b^* = (\boldsymbol{\beta}_b^* - \hat{\boldsymbol{\beta}})^\top \boldsymbol{V}^{-1} (\boldsymbol{\beta}_b^* - \hat{\boldsymbol{\beta}}).$$

Then the bootstrap p-value is

$$\frac{1}{B}\sum_{b=1}^{B}\mathbb{I}(W_b^* > W).$$

Table C1 reports individual and joint test statistics. As shown in columns (1) and (2), we reject that many — far more than 5% — of the individual covariances between earnings levels and growth are zero at the 5% significance level. We also reject that roughly 5% of the individual covariances between earnings growth are zero at the 5% significance level: this is precisely what we would expect

Covariance between		(1) Number of Individual Tests: Total	(2) Number of Individual Tests: p-value<0.05	(3) Joint Test: W	(4) Joint Test: Range of W_b^*	(5) Joint Test: Range of W_b^*	(6) Joint Test: p-value
$y_{p,t}$	$\Delta y_{k,t'}$	1,258	405	0.0175	0.0039	0.0081	0
$\Delta y_{p,t}$	$y_{k,t'}$	1,247	153	0.0066	0.0038	0.0058	0
$\Delta y_{p,t}$	$\Delta y_{k,t'}$	1,130	57	0.0035	0.0028	0.0040	0.48

Table C1: Zero Covariance Test Statistics

when all earnings growth covariances are zero. This is further confirmed by the high p-value for the joint hypothesis that all earnings growth covariances are zero, reported in column (6). The joint tests also confirm that covariances between earnings levels and growth are significantly different from zero.

D Comparing IGEs with Previous Estimates

Our estimated IGEs are smaller than those reported in other studies using the same Canadian data. In particular, the most recent study by Chen, Ostrovsky, and Piraino (2017) reports an IGE of 0.32 as its preferred estimate (row (4) of Table 1). In this appendix, we show that the difference between their and our estimates reflect differences in sample selection and the calculation of average earnings.

Chen, Ostrovsky, and Piraino (2017) use earnings data over years 1978–2008 for eldest sons born from 1963 to 1966 and their fathers born from 1932 to 1955. Sons must have annual earnings of at least \$500 (in 2010 Canadian dollars) in at least 3 years during ages 38–42, and fathers' annual earnings must be at least \$500 in at least 10 years during ages 35–55. Using only earnings of at least \$500, earnings are averaged over ages 38–42 for sons and ages 35–55 for fathers. The IGE is estimated by regressing log average earnings of sons on log average earnings of fathers.

In contrast, we use earnings data from 1978 to 2014, and our sample includes sons born in later years. Our cohort of 1965 children (born 1963–1966) with 1940 parents (born 1934–1944) most closely matches the cohorts used in Chen, Ostrovsky, and Piraino (2017); however, we do not select individuals based on the number of years with earnings above a certain threshold. In addition to these differences in sample selection, we calculate average earnings differently. First, fathers' earnings are averaged over 5 (and 9) years. Second, fathers' earnings are measured around age 50 and sons earnings are measured around age 30. Third, as earnings measures, we use averages of log earnings residuals, where extreme earnings observations (top and bottom 1% within age, year, and family cohort group) are dropped before residualizing earnings. In calculating average log earnings residuals over 5 years, we allow for 1 year of missing observation.

To isolate the impact of differences in earnings measures, we use the sample of Chen, Ostrovsky, and Piraino (2017) and estimate IGEs based on different earnings measures. These results are shown in Table D2. In Column (1), we follow the same procedure as Chen, Ostrovsky, and Piraino (2017), reproducing their preferred estimate. As Column (2) shows, the IGE decreases by about 0.04 when fathers' earnings are averaged over 9 years (instead of 21 years) around age 45 (the midpoint of their age range). Averaging fathers' earnings over 5 years further reduces the IGE by around 0.03 (Column (3)), and averaging fathers' earnings around age 50 (instead of age 45) lowers the IGE by another 0.02 (Column (4)). In Column (5), sons' earnings are also averaged around age 30 (instead of age 40, the midpoint of their age range), which produces an IGE of 0.20. Finally, Column (6) reports an IGE of 0.21 when we use average log earnings residuals (with trimming) as earnings measures. This is slightly higher than our estimate for 1965 children with 1940 parents, 0.19, reported in Table 3.

Summarizing, our IGE estimate (using five-year average earnings) of 0.19 for the 1965 children with 1940s parents is 0.13 lower than the preferred estimate of Chen, Ostrovsky, and Piraino (2017). While 13% of this difference is due to the narrower set of father birth-year cohorts included in our "1940s" parent cohort, most of the difference is explained by our use of fathers' earnings averaged over 5 rather than 21 years (56%) and the ages around which both sons' and fathers' earnings are measured (37%).

	(1)	(2)	(3)	(4)	(5)	(6)
Fathers' Ages	35–55	41–49	43–47	48–52		
Sons' Ages	38–42			28–32		
Earnings Measures	Log Average Earnings Conditional on ≥$500					Average Log Residual
IGE	0.316 (0.003)	0.277 (0.003)	0.244 (0.003)	0.225 (0.002)	0.197 (0.002)	0.205 (0.003)

Table D2: IGE Estimated with Different Earnings Measures

E Earnings IGEs by Family Structure

In Section 5.2, we show that earnings IGEs are lower for children born to older parents, suggesting a potential role for parental age at birth in intergenerational mobility. In this appendix, we show that this pattern is not driven by differences in family structure that are correlated with parental age at birth. Specifically, we estimate the earnings IGE separately by the number of children, number of parents, and birth order of children.

We use information on the number of children and parents for each family based on the year children were linked to their parents. We construct the birth order for each child within family using all children (not just sons) included in the IID data. Unfortunately, we may not observe older (younger) siblings of older (younger) cohorts of IID children. For this reason, we restrict the sample to the middle cohort of IID children—those born 1968–1980—whose siblings are most likely to be included in the data.

Tables E3 and E4 show the distribution of children and 5-year earnings IGEs by family structure among eldest sons (Panel A) and all sons (Panel B). Sons born to older parents are more likely to have older and fewer siblings, while having a lone father appears to be largely unrelated to fathers' age at birth. Table E4 shows that these differences in family structure are only very weakly related to earnings IGEs. Parental age at child's birth is still negatively correlated with the earnings IGE when we condition on family structure.

Father's Age at Birth	Birth Order			Number of Children			Single Fathers	Total
	First	Second	Third+	One	Two	Three+		
				A. Eldest Sons				
16–25	89.9	9.6	0.5	17.3	44.3	38.5	7.3	22.0
26–35	73.9	22.3	3.7	23.1	47.3	29.6	6.0	71.3
36–50	62.6	28.1	9.3	39.3	41.4	19.3	6.9	6.7
				B. All Sons				
16–25	81.4	16.8	1.8	16.6	43.2	40.1	7.2	18.4
26–35	54.6	33.9	11.5	21.1	45.1	33.8	5.5	73.0
36–50	37.4	35.7	27.0	34.6	40.3	25.1	6.0	8.6

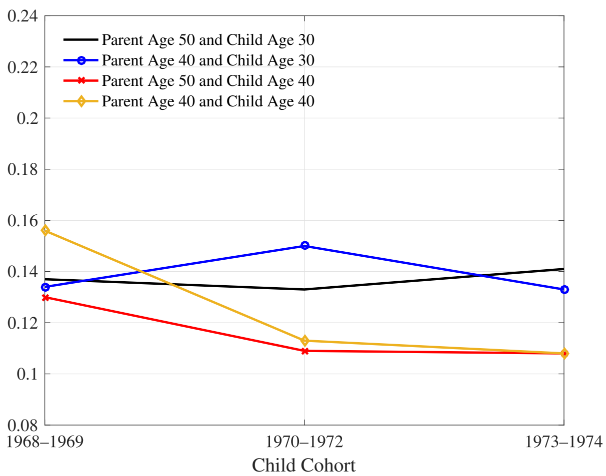
Table E3: Fraction of Children (%) by Family Structure

Father's Age at Birth	Birth Order			Number of Children			Parents	
	First	Second	Third+	One	Two	Three+	Single	Two
				A. Eldest Sons				
16–25	0.155	0.147	0.175	0.152	0.156	0.151	0.156	0.153
	(0.003)	(0.009)	(0.039)	(0.007)	(0.004)	(0.004)	(0.012)	(0.003)
26–35	0.128	0.122	0.126	0.135	0.119	0.126	0.132	0.125
	(0.002)	(0.003)	(0.008)	(0.003)	(0.002)	(0.003)	(0.007)	(0.002)
36–50	0.099	0.096	0.090	0.097	0.095	0.105	0.106	0.096
	(0.007)	(0.010)	(0.017)	(0.009)	(0.009)	(0.012)	(0.024)	(0.006)
				B. All Sons				
16–25	0.155	0.153	0.169	0.152	0.157	0.153	0.154	0.154
	(0.003)	(0.007)	(0.022)	(0.008)	(0.005)	(0.005)	(0.013)	(0.003)
26–35	0.128	0.119	0.140	0.129	0.120	0.128	0.127	0.125
	(0.002)	(0.002)	(0.004)	(0.003)	(0.002)	(0.003)	(0.007)	(0.002)
36–50	0.099	0.101	0.089	0.094	0.095	0.099	0.113	0.094
	(0.007)	(0.008)	(0.009)	(0.008)	(0.007)	(0.009)	(0.022)	(0.005)

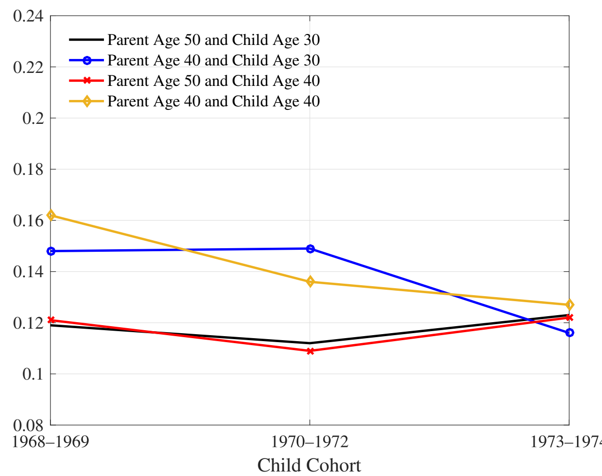
Table E4: 5-Year Average Earnings IGE by Family Structure

F Earnings IGE Measured at Different Ages

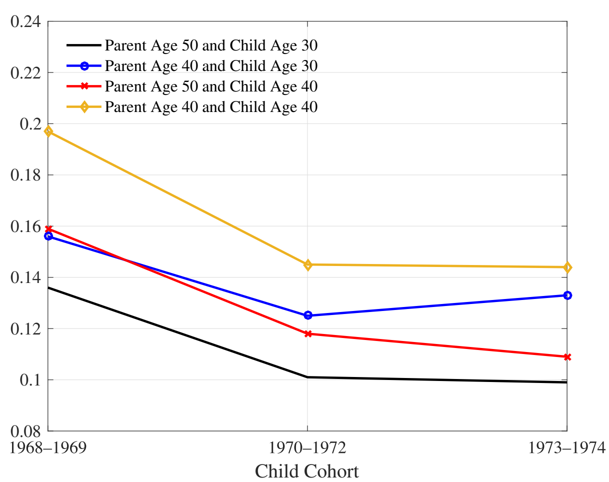
As explained in Section 3, we generally observe (in the IID data) children's earnings when they are young and parents' earnings when they are old. In particular, we report earnings IGEs based on earnings around ages 50 for parents and 30 for children. In this appendix, we show how the earnings IGE cohort patterns (e.g., see Table 3 and Figure 6) are affected when earnings are measured at more similar ages for parents and children for a very limited set of cohorts. Specifically, we examine children born from 1968 to 1974 whose fathers were ages 16–30 when they were born. For this sample, we can observe earnings at age 40 for both parents and children.



(a) Parental Age at Child's Birth: 16–20



(b) Parental Age at Child's Birth: 21–25



(c) Parental Age at Child's Birth: 26–30

Figure F4: Cohort Patterns for Annual Earnings IGEs Measured at Different Ages

G MD Parameter Estimates (National Analysis)

This appendix reports key parameter estimates from our MD approach using the Canadian national sample.[45] Table G5 shows estimated AR(1) and MA(1) coefficients and their standard errors for the persistent and transitory components. Figure G5 displays estimated covariances of skill factors (within generation) for fathers and sons. Figure G6 displays intergenerational covariance estimates for the skill factors, while Figure G7 reports the corresponding intergenerational correlations.

Figure G8 shows "standardized" linear-projection coefficients (i.e., linear-projection coefficients when all variables are standardized to have a unit standard deviation), defined as follows:

$$\underbrace{\begin{bmatrix} \alpha^*_{\psi,\psi} & \alpha^*_{\psi,\delta} \\ \alpha^*_{\delta,\psi} & \alpha^*_{\delta,\delta} \end{bmatrix}}_{:=\mathbf{A}^*} := \begin{bmatrix} 1 & \mathrm{Corr}(\psi_p,\delta_p) \\ \mathrm{Corr}(\psi_p,\delta_p) & 1 \end{bmatrix}^{-1} \begin{bmatrix} \mathrm{Corr}(\psi_p,\psi_k) & \mathrm{Corr}(\psi_p,\delta_k) \\ \mathrm{Corr}(\delta_p,\psi_k) & \mathrm{Corr}(\delta_p,\delta_k) \end{bmatrix}$$

$$= \begin{bmatrix} \alpha_{\psi,\psi}\frac{\sigma(\psi_p)}{\sigma(\psi_k)} & \alpha_{\psi,\delta}\frac{\sigma(\psi_p)}{\sigma(\delta_k)} \\ \alpha_{\delta,\psi}\frac{\sigma(\delta_p)}{\sigma(\psi_k)} & \alpha_{\delta,\delta}\frac{\sigma(\delta_p)}{\sigma(\delta_k)} \end{bmatrix},$$

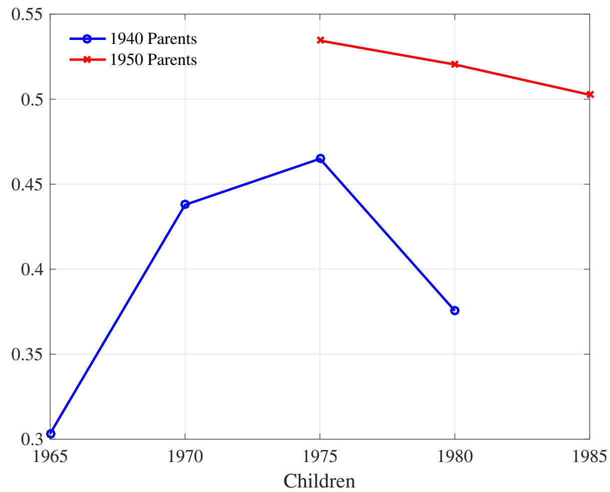
where $\sigma(x) := \sqrt{\mathrm{Var}(x)}$ is the standard deviation of x.

Finally, Figures G9 and G10 report estimated variances for the persistent and transitory shocks by age and cohort.

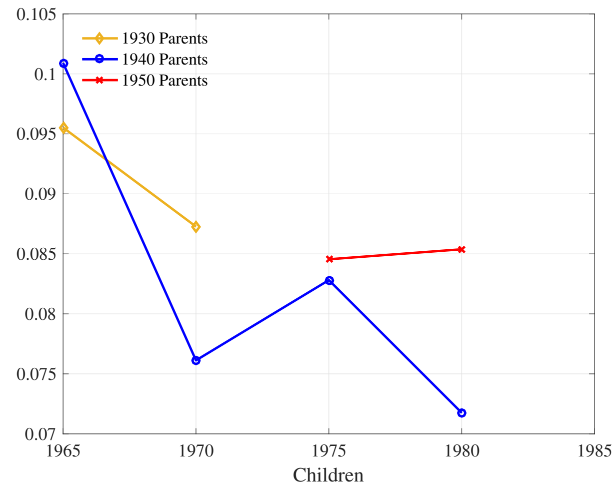
	Parents	Children
ρ_j	0.908	0.833
	(0.002)	(0.001)
κ_j	0.120	0.106
	(0.002)	(0.002)

Table G5: AR(1) and MA(1) Coefficients

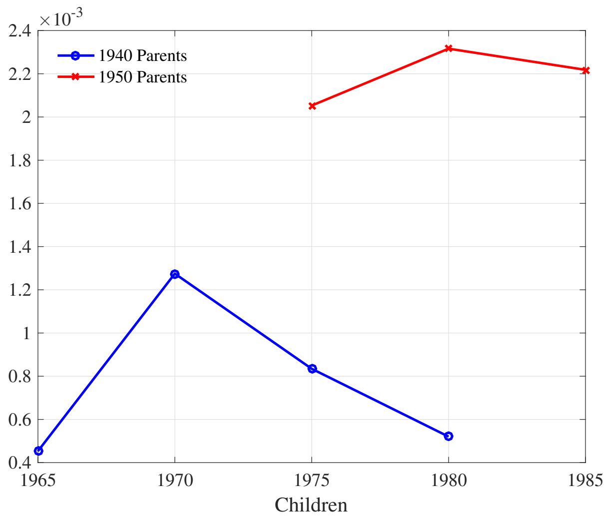
[45] The minimized criterion function for our estimates (see Section 4) is SSR = 0.049.



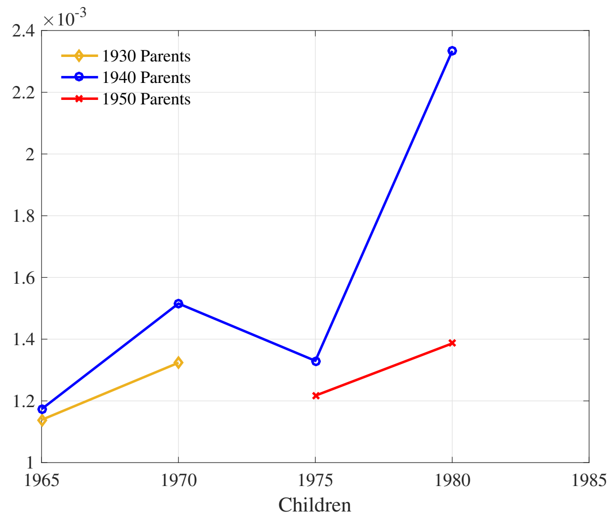
(a) Var(ψ_p)



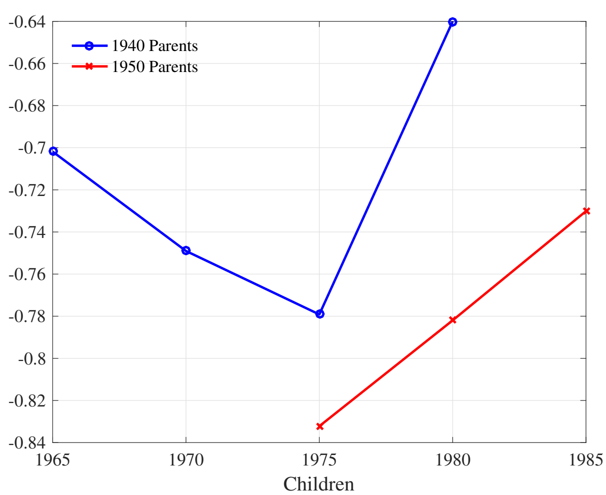
(b) Var(ψ_k)



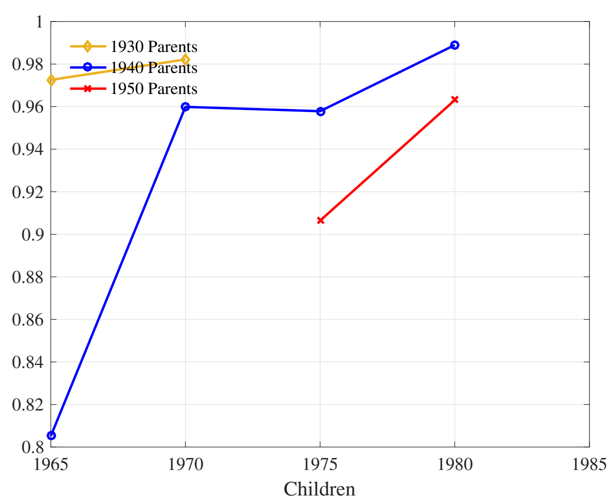
(c) Var(δ_p)



(d) Var(δ_k)

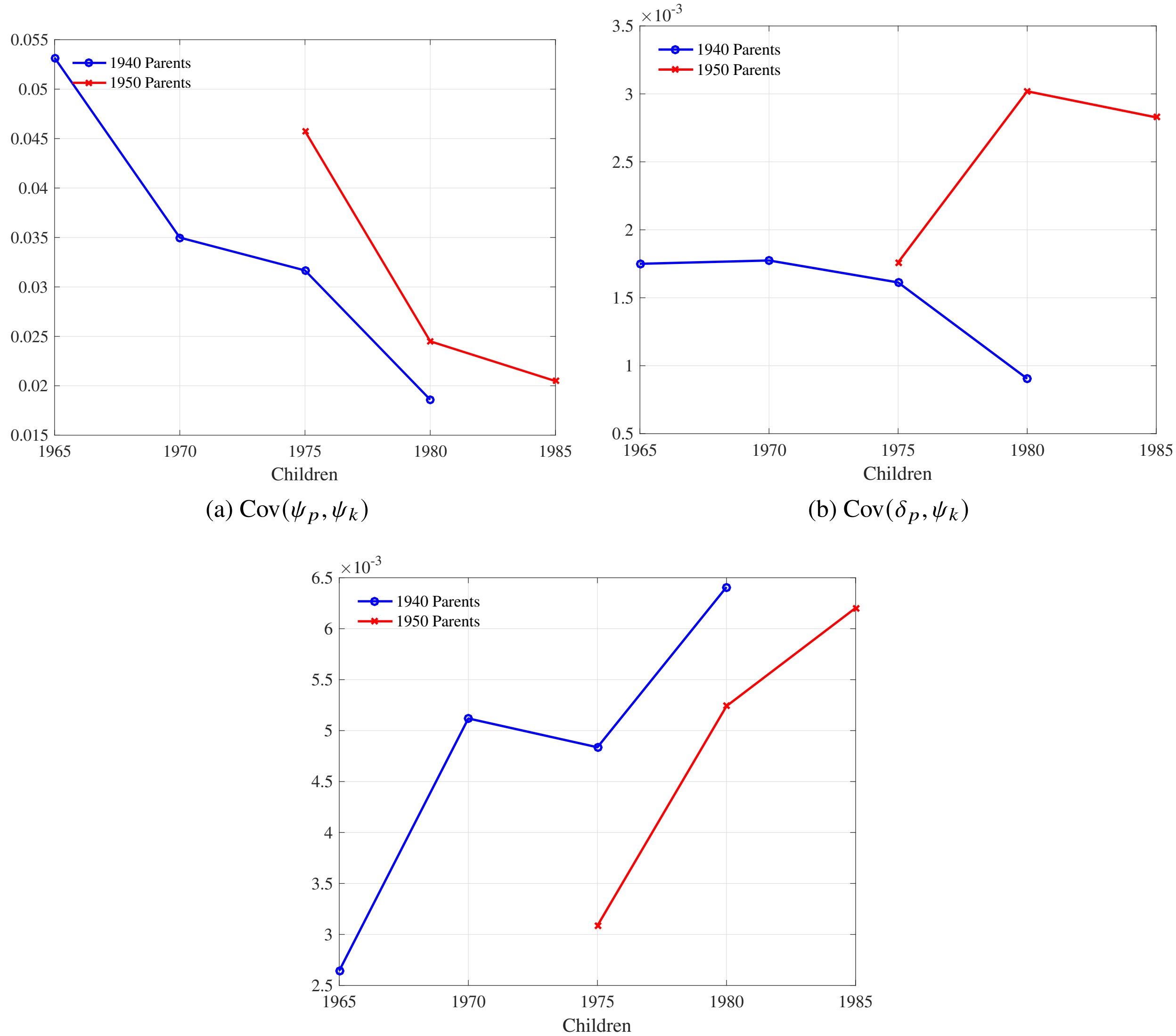


(e) Corr(ψ_p, δ_p)



(f) Corr(ψ_k, δ_k)

Figure G5: Covariances of Skill Factors ($\mathbf{\Omega}_p$ and $\mathbf{\Omega}_k$) by Cohort

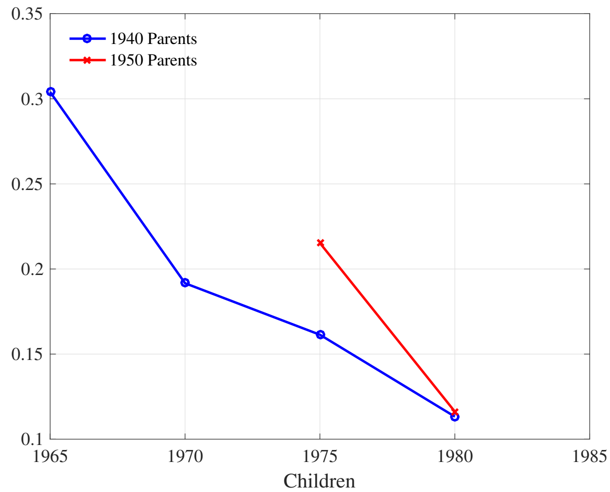


(a) Cov(ψ_p, ψ_k)

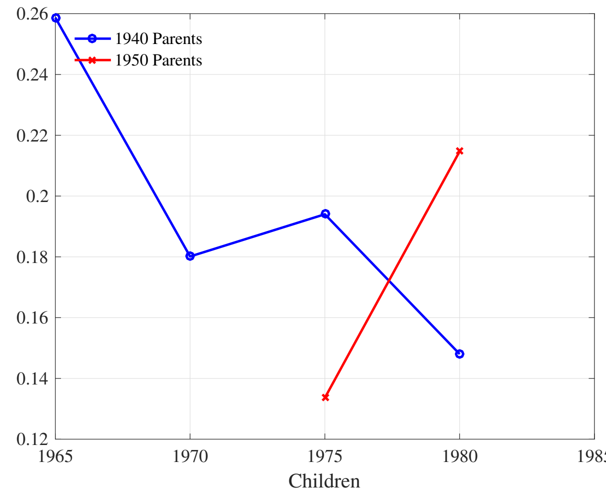
(b) Cov(δ_p, ψ_k)

(c) Cov(ψ_p, δ_k)

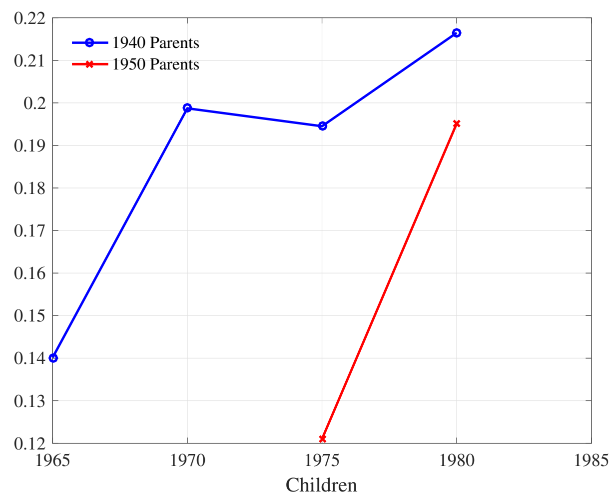
Figure G6: Intergenerational Covariances of Skill Factors ($\mathbf{\Omega}_{p,k}$) by Cohort



(a) Corr(ψ_p, ψ_k)



(b) Corr(δ_p, ψ_k)



(c) Corr(ψ_p, δ_k)

Figure G7: Intergenerational Correlations of Skill Factors by Cohort

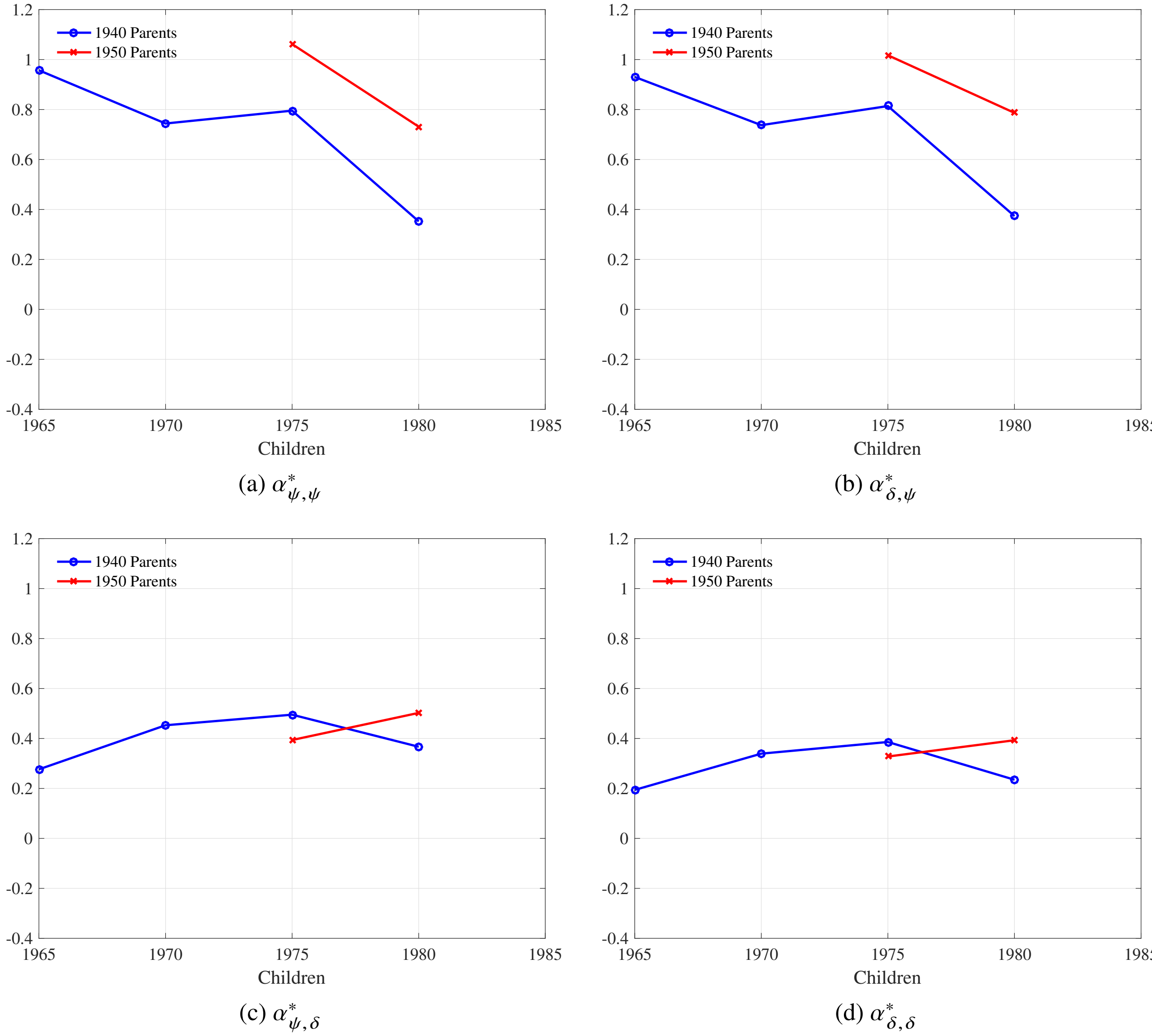
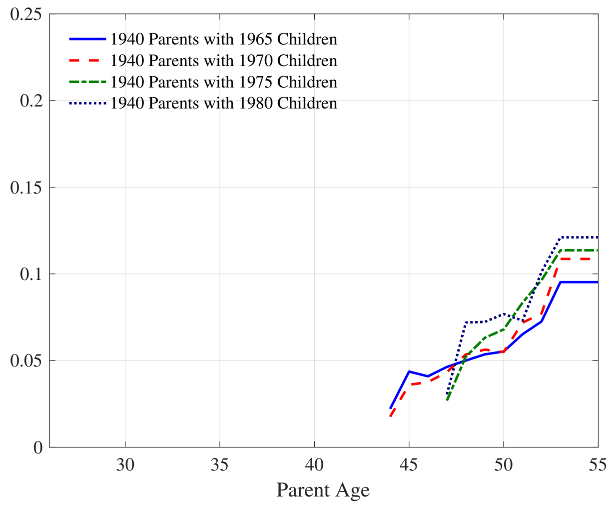
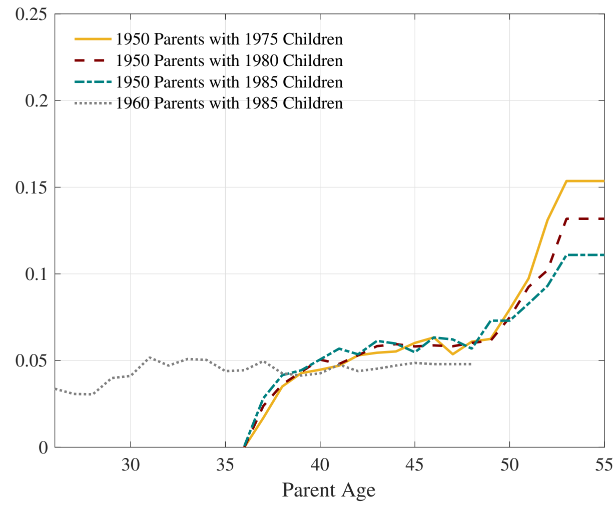


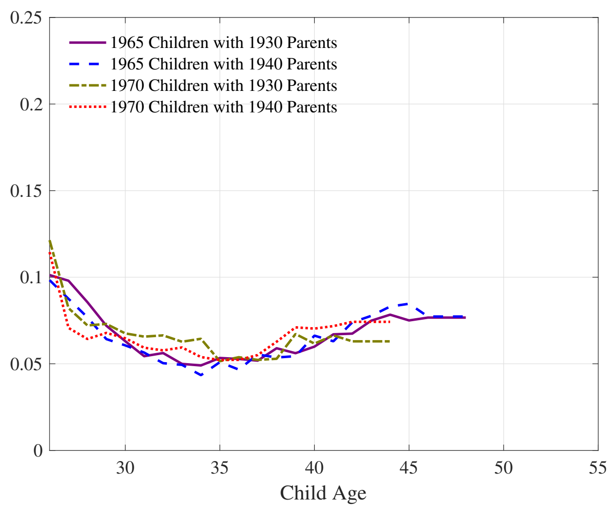
Figure G8: Standardized Linear-Projection Coefficients ($\mathbf{A}^*$) by Cohort



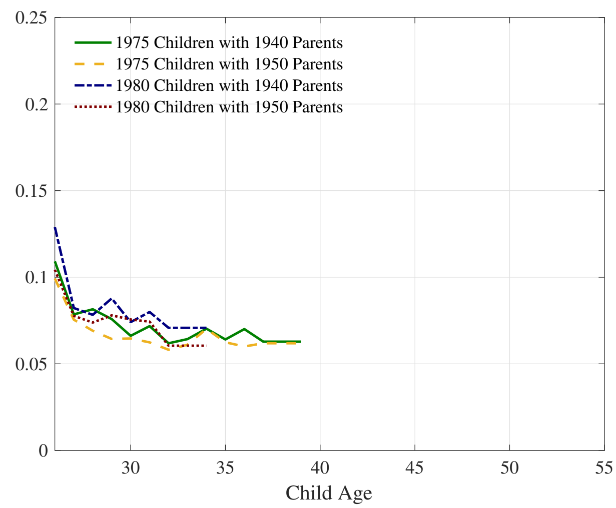
(a) 1940 Parents



(b) 1950 and 1960 Parents



(c) 1965 and 1970 Children



(d) 1975 and 1980 Children

Figure G9: Persistent Shock Variances (Var($\nu_{j,t}$)) by Age and Cohort

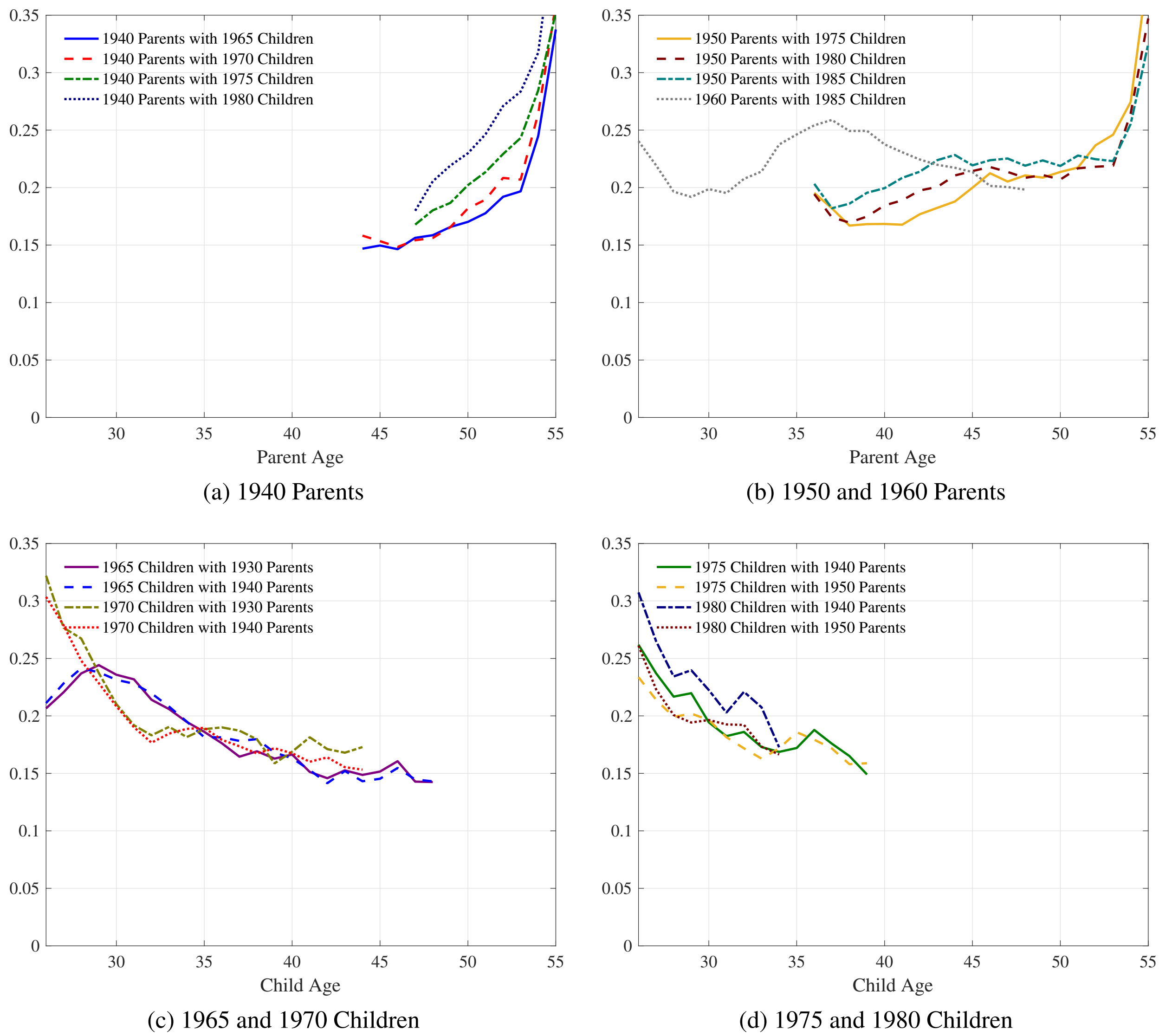


Figure G10: Transitory Shock Variances (Var($\xi_{j,t}$)) by Age and Cohort

H Additional Results for National Analysis

	Cohort Group							
Parents	1940				1950			1960
Children	1965	1970	1975	1980	1975	1980	1985	1985
Annual	0.312	0.342	0.353	0.326	0.261	0.284	0.300	
	(0.006)	(0.006)	(0.015)	(0.024)	(0.003)	(0.004)	(0.003)	
5-Year Average	0.441	0.484	0.504	0.473	0.372	0.397	0.418	0.321
	(0.009)	(0.009)	(0.021)	(0.035)	(0.004)	(0.004)	(0.004)	(0.005)
9-Year Average	0.462	0.501	0.517	0.497	0.379	0.415	0.446	0.375
	(0.008)	(0.008)	(0.019)	(0.033)	(0.006)	(0.006)	(0.005)	(0.006)
13-Year Average					0.712	0.705	0.703	0.502
					(0.004)	(0.003)	(0.007)	(0.007)
23-Year Average								0.556
								(0.007)

Notes: Annual measures are based on ages 30 for children and 50 for fathers, 5-year averages based on ages 28–32 for children and 48–52 for fathers, 9-year averages based on ages 26–34 for children and 47–55 for parents, 13-year averages based on ages 36–48 for fathers, and 23-year averages based on ages 26–48 for children and fathers. Standard errors are based on 100 bootstrap samples.

Table H6: Father's Share of Earnings Variance from Skills ($\mathrm{Var}(\bar{\theta}_p)/\mathrm{Var}(\bar{y}_p)$) by Cohort

	Cohort Group					
Parents	1930		1940			1950
Children	1965	1970	1965	1970	1975	1975
			A. 13-Year Average			
$\text{Var}(\overline{\theta}_k)/\text{Var}(\overline{y}_k)$	0.543	0.519	0.550	0.522	0.516	0.528
	(0.004)	(0.006)	(0.003)	(0.004)	(0.006)	(0.006)
$\text{Var}(\overline{\hat{\theta}}_k)/\text{Var}(\overline{y}_k)$			0.198	0.105	0.111	0.130
			(0.029)	(0.018)	(0.024)	(0.010)
$\text{Var}(\overline{\hat{\theta}}_k)/\text{Var}(\overline{\theta}_k)$			0.360	0.201	0.215	0.245
			(0.052)	(0.033)	(0.047)	(0.019)
			B. 23-Year Average			
$\text{Var}(\overline{\theta}_k)/\text{Var}(\overline{y}_k)$	0.702		0.702			
	(0.003)		(0.003)			
$\text{Var}(\overline{\hat{\theta}}_k)/\text{Var}(\overline{y}_k)$			0.225			
			(0.032)			
$\text{Var}(\overline{\hat{\theta}}_k)/\text{Var}(\overline{\theta}_k)$			0.321			
			(0.046)			

Notes: 13-year averages are based on ages 26–38, and 23-year averages are based on ages 26–48. Standard errors (in parentheses) are based on 100 bootstrap samples.

Table H7: Children's Skill and Earnings Variances by Cohort: 13- and 23-Year Averages

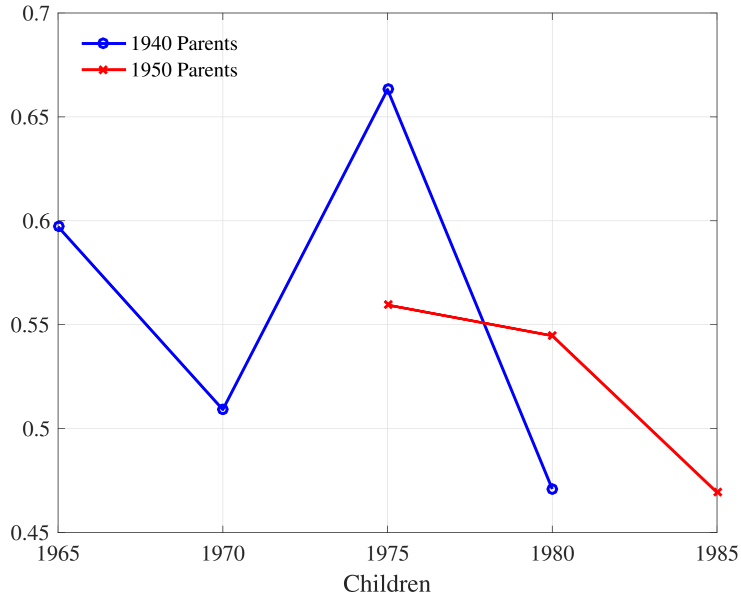
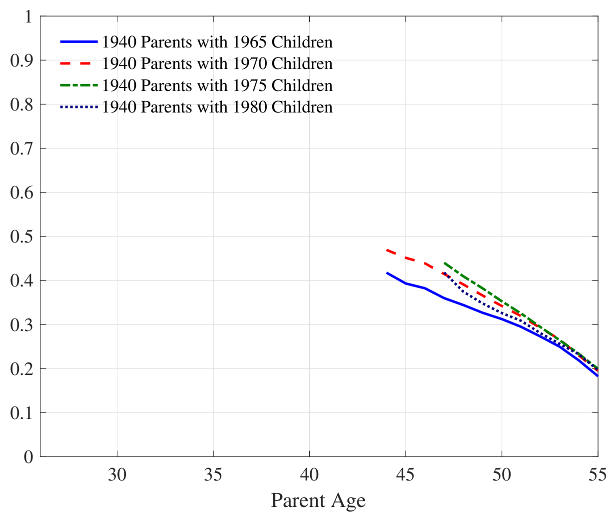
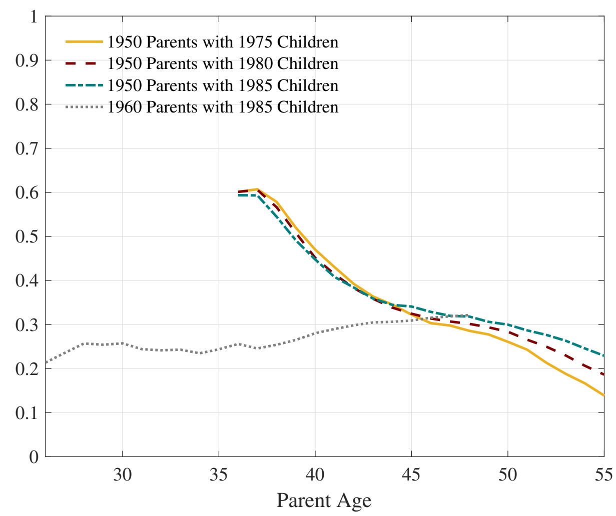


Figure H11: $\mathrm{E}\left[\mathrm{Var}(\hat{\overline{\theta}}_k|\overline{\theta}_p)\right]/\mathrm{Var}(\hat{\overline{\theta}}_k)$ Based on Lifetime Average Skills, by Cohort

Notes: $\hat{\overline{\theta}}_k$ reflects average child projected skills over ages 26–48, while $\overline{\theta}_p$ reflects average parental skills over ages 26–55.

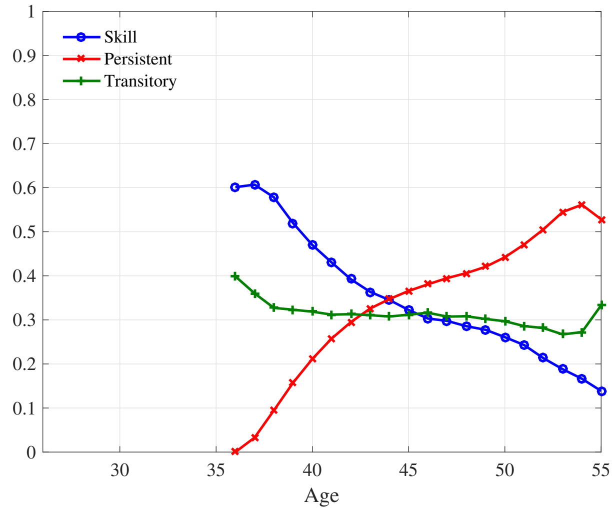


(a) 1940 Parents

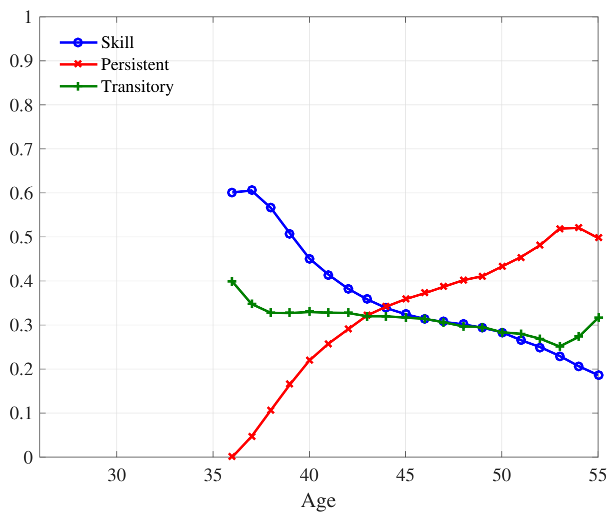


(b) 1950 and 1960 Parents

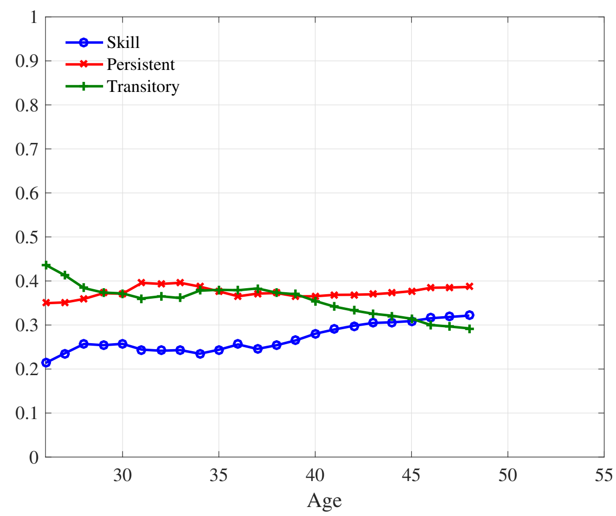
Figure H12: $\mathrm{Var}(\theta_{p,t})/\mathrm{Var}(y_{p,t})$ by Age and Cohort



(a) 1975 Children, 1950 Parents

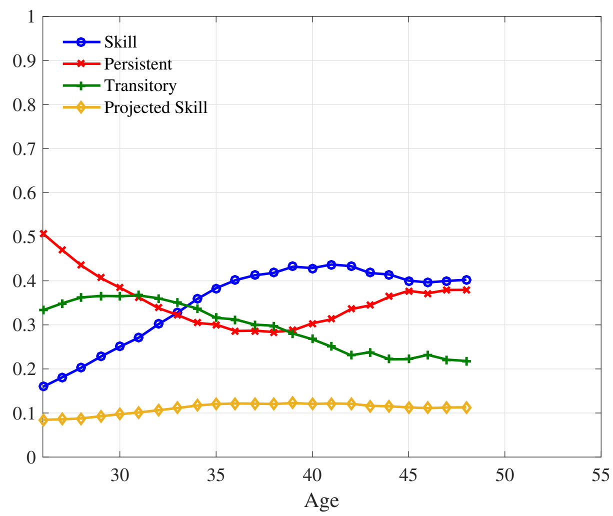


(b) 1980 Children, 1950 Parents

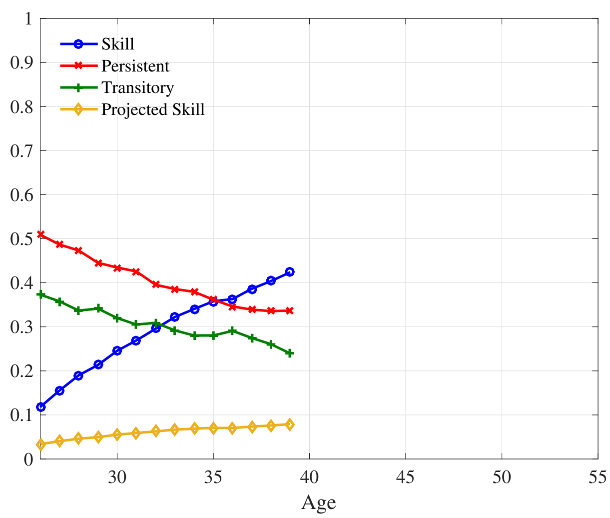


(c) 1985 Children, 1960 Parents

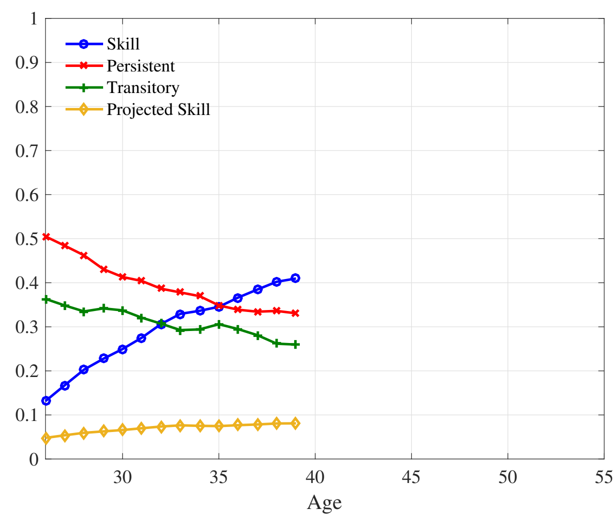
Figure H13: Share of Earnings Variances for Parents by Age



(a) 1965 Children, 1940 Parents



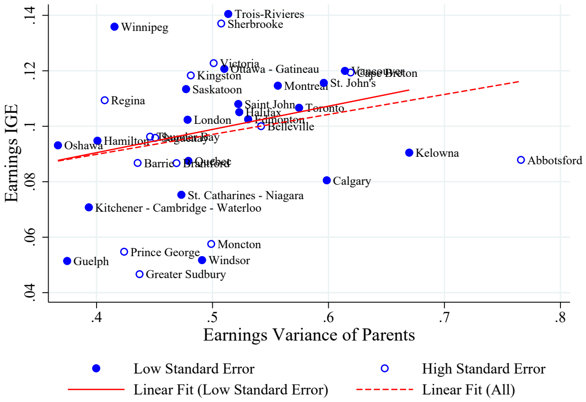
(b) 1975 Children, 1940 Parents



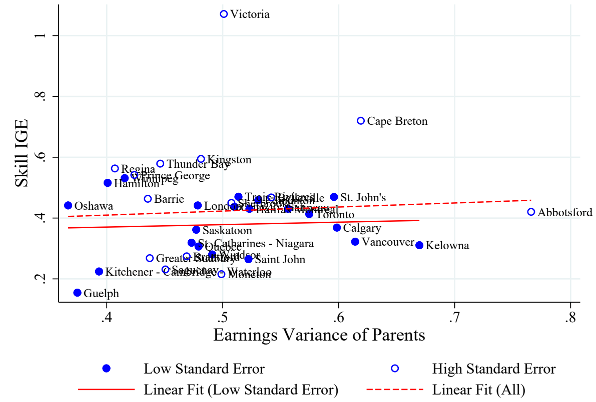
(c) 1975 Children, 1950 Parents

Figure H14: Share of Earnings Variances for Children by Age

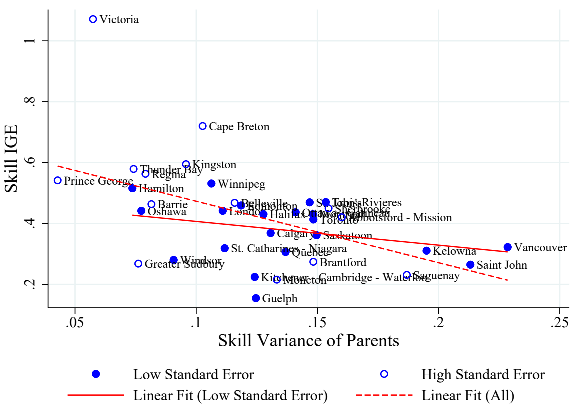
I Additional Results for Regional Analysis



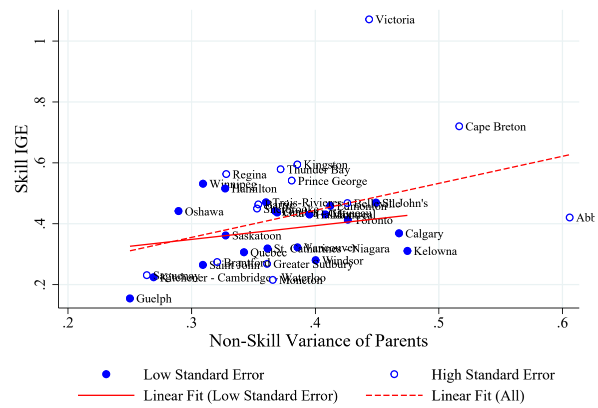
(a) Earnings IGE vs. Earnings Variance



(b) Skill IGE vs. Earnings Variance

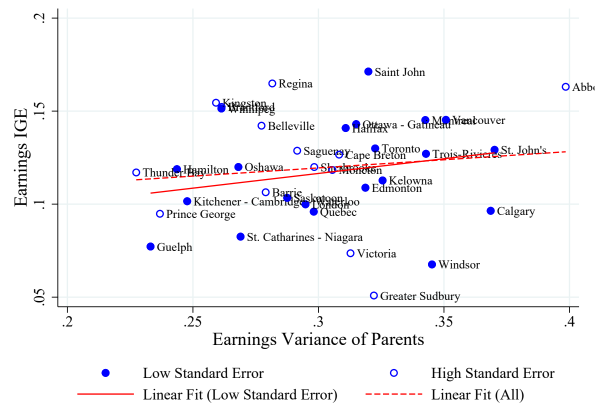


(c) Skill IGE vs. Skill Variance

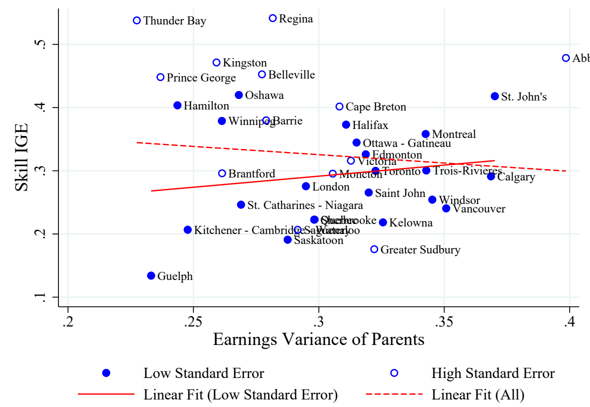


(d) Skill IGE vs. Non-Skill Variance

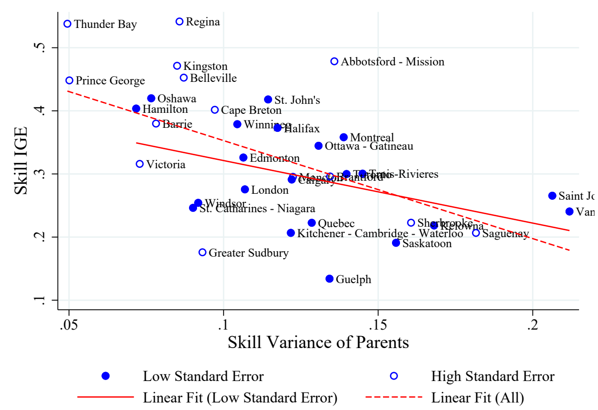
Figure I15: Intergenerational Mobility and Inequality across Canadian Cities: Annual Measures



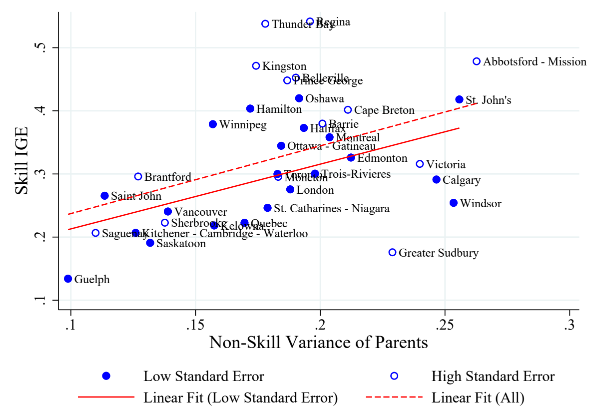
(a) Earnings IGE vs. Earnings Variance



(b) Skill IGE vs. Earnings Variance



(c) Skill IGE vs. Skill Variance



(d) Skill IGE vs. Non-Skill Variance

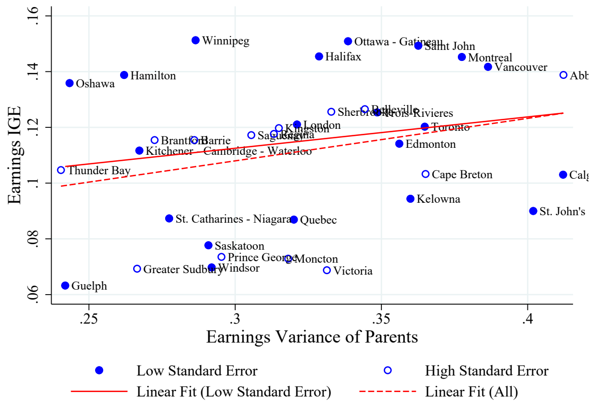
Figure I16: Intergenerational Mobility and Inequality across Canadian Cities: 9-Year Averages

J Controlling for Location in Each Year

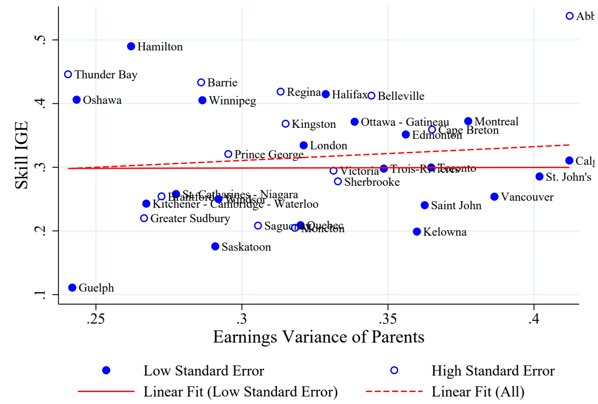
In this appendix, we reproduce the main results of Section 6 while controlling for where individuals live in each year. Separately for each generation and year, we work with log earnings residuals from regressions of log earnings on indicators for (i) the full interaction between birth year and city of residence when the parent–child linkage was formed and (ii) the current location where each individual lives. Our sample is still based on individuals from the 35 largest CMA/CA when the parent–child linkage was formed. We consider 61 mutually exclusive current location categories: (i) 35 largest CMA/CA, (ii) one category for all other CMA/CA in each province/territory (13 total), and (iii) other non-CMA/CA provinces/territories (13 total).

The results presented in Figure J17 and Table J8 are similar to those in Figure 15 and Table 5, suggesting that controlling for the location each year does not change our main results.[46] This is partly because individuals from big cities tend to stay in the same city or move to other big cities. For example, around 90% of fathers and sons in our sample (i.e., those living in the 35 largest CMA/CA in the linkage year) are still living in the 35 largest CMA/CA in the 2010s. Moreover, around 85% of fathers and 75% of sons remain in the same CMA/CA in the 2010s.

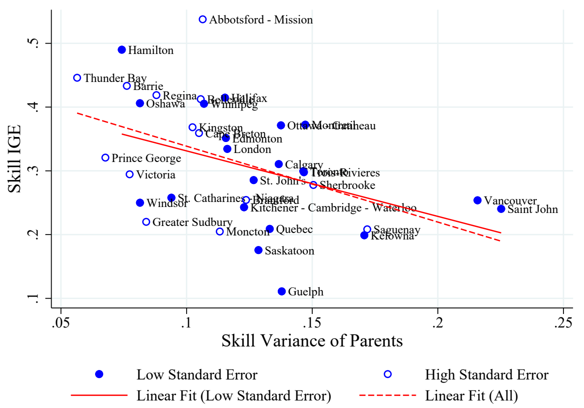
[46]The identity of "Low Standard Error" cities are unchanged.



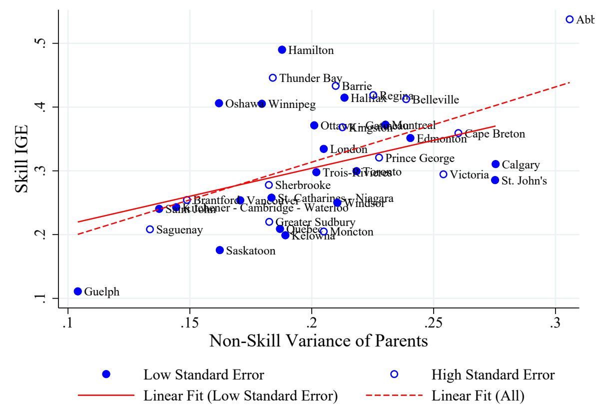
(a) Earnings IGE vs. Earnings Variance



(b) Skill IGE vs. Earnings Variance



(c) Skill IGE vs. Skill Variance



(d) Skill IGE vs. Non-Skill Variance

Figure J17: Intergenerational Mobility and Inequality across Canadian Cities: 5-Year Averages

Notes: These results are based on log earnings residuals, controlling for each individual's city of residence in each year.

Dependent	Independent	Slope Estimate
Earnings IGE	Earnings Variance	0.112
Skill IGE	Earnings Variance	0.009
Skill IGE	Skill Variance	-1.027
Skill IGE	Non-Skill Variance	0.878
Skill Share of Earnings Variance	Earnings Variance	0.140

Table J8: Slope Coefficients Among "Low Standard Error" Cities: 5-Year Average

Notes: These results are based on log earnings residuals, controlling for each individual's city of residence in each year.