EARNINGS DYNAMICS AND INTERGENERATIONAL TRANSMISSION OF SKILL

Lance Lochner
Youngmin Park

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 (less than 0.2, even when based on 5- and 9-year average earnings); 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.

We estimate significant heterogeneity in both initial skills and skill growth rates, showing that intergenerational transmission of these factors 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 with age. Although "inherited" initial skills (compared to skill growth) are a more important determinant of children's skills throughout their lives, 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, determining the extent to which considerable differences in earnings and skill IGEs vary with the extent of local heterogeneity in parental skills vs. earnings instability.

Lance Lochner
Department of Economics
Faculty of Social Science
Western University
1151 Richmond Street, North
London, ON N6A 5C2
CANADA
and NBER
llochner@uwo.ca

Youngmin Park
Bank of Canada
234 Wellington Street
Ottawa, Ontario
K1A 0G9
Canada
plyoungmin@gmail.com

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; Cunha, Heckman, and Schennach, 2010; Gayle, Golan, and Soytas, 2015; Lee and Seshadri, 2019; Caucutt and Lochner, 2020).

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 IGE) in estimates based on annual measures of earnings. Jenkins (1987) and, more recently, Haider and Solon (2006) further argued that correlation between earnings at any age and lifetime earnings varies systematically over the lifecyle, generating a different form of "lifecycle bias" (for the lifetime IGE). 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) 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).

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

---

[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.

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 being larger than the latter due to non-skill earnings shocks. We confirm prior studies by Corak and Heisz (1999) and Chen, Ostrovsky, and Piraino (2017) that show that standard earnings IGEs are quite low in Canada (less than 0.2 for our sample), even when using 5- or 9-year averages for fathers' and sons' earnings. 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 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 alterative summary measure of intergenerational mobility that quantifies the extent to which intergenerational transmission explains the 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. Thus, 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 the 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 initial skills (compared to skill growth rates) is a more important source of skill variation throughout life. An alternative variance decomposition reveals that heterogeneity in 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' skill levels at a given age (but not their initial skill and growth rates) explains at most 55% of all projected skill variation among their sons, highlighting the importance of accounting for the transmission of both skill factors.

Given the recent interest in regional variation in intergenerational mobility (e.g., Chetty et al.,

2014a; Connolly, Corak, and Haeck, 2019; Corak, forthcoming), 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 stronger links between skills and earnings (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 would instead focus attention on 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.

Consistent with prior studies for Canada (Connolly, Corak, and Haeck, 2019; Corak, forthcoming) and the Great Gatsby Curve, we estimate higher earnings IGEs for cities with greater parental earnings inequality; however, our estimated relationship is only statistically significant at around the 10% level. 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 (and much noisier) than that observed for 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 more earnings inequality conditional on skills.

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 levels 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 earnings of family member j in family i and at age t, and let $y_{i,j,t}$ be the log-earnings residual, after taking out age/time and cohort effects of log earnings for each j. 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.

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 "skill" and $\varepsilon_{i,j,t}$ reflects the "non-skill" components at age t. We normalize both components so they have mean zero and are orthogonal to each other:[3]

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

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 skill and earnings 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 (i, j) at age $\underline{t}$, the beginning of individual's careers. 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

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

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

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

---

[3]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 $\mathrm{Cov}(x_{j,t}, x_{j',t'})$.

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).[4]

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 $\lambda_{j,t}$:

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

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

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

Finally, 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:[5]

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

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 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.[6] Accordingly, our notion of

---

[4]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. Since we do not observe education in our data, heterogeneity in $(\psi_{i,j}, \delta_{i,j})$ will reflect any systematic effects of education on skill levels and growth rates (see, e.g., Heckman, Lochner, and Todd, 2006, 2008).

[5]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.

[6]Consistent with this assumption, Gallipoli, Low, and Mitra (2020) estimate insignificant effects of permanent shocks to parental earnings on the permanent earnings shocks of their children.

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 $\underline{t}$.[7]

2.2 Intergenerational Transmission of Skill

In our framework, the intergenerational transmission of skill (and earnings) comes entirely from the intergenerational transmission of the 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'):

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

where

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

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

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

represents the analogous skill IGE. (We note that the 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.[8] 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 measured at any given age.[9] It is worth further noting that average realized lifetime earnings also

---

[7]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); Gayle, Golan, and Soytas (2015); Lee and Seshadri (2019); Caucutt and Lochner (2020).

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

[9]The distinction between expected earnings at a given age (like our measure of skill) and "average lifetime earnings" leads

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).[10] 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:

$$\hat{\psi}_k := P[\psi_k | \psi_p, \delta_p] = \alpha_{\psi,\psi}\psi_p + \alpha_{\delta,\psi}\delta_p$$
$$\hat{\delta}_k := P[\delta_k | \psi_p, \delta_p] = \alpha_{\psi,\delta}\psi_p + \alpha_{\delta,\delta}\delta_p,$$

where $P[y|x]$ denotes the linear projection of y onto x and the linear-projection coefficients are given by

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

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 the "projected skill" at age t:

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

---

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.

[10]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 D-12 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.

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

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

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

$$\frac{\text{Var}(\hat{\theta}_{k,t})}{\text{Var}(y_{k,t})} = \frac{\text{Var}(\hat{\theta}_{k,t})}{\text{Var}(\theta_{k,t})} \, s_{k,t}, \tag{5}$$

where $s_{k,t}$ is the share of children's earnings explained by skills (see equation (2)). Intergenerational transmission explains a smaller share of the earnings variation than the skill variation due to the non-skill earnings shocks children experience throughout their lives.[11]

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

$$\hat{\theta}_{k,t} = \underbrace{\left(\alpha_{\psi,\psi} + \alpha_{\psi,\delta}\Lambda_{k,t}\right)}_{:=\beta_{\psi,t}} \psi_p + \underbrace{\left(\alpha_{\delta,\psi} + \alpha_{\delta,\delta}\Lambda_{k,t}\right)}_{:=\beta_{\delta,t}} \delta_p. \tag{6}$$

Thus, equations (4) and (6) can be used for two distinct projected skill variance decompositions:

$$\text{Var}(\hat{\theta}_{k,t}) = \text{Var}(\hat{\psi}_k) + \Lambda_{k,t}^2 \text{Var}(\hat{\delta}_k) + 2\Lambda_{k,t} \text{Cov}(\hat{\psi}_k, \hat{\delta}_k) \tag{7}$$

$$= \beta_{\psi,t}^2 \text{Var}(\psi_p) + \beta_{\delta,t}^2 \text{Var}(\delta_p) + 2\beta_{\psi,t}\beta_{\delta,t} \text{Cov}(\psi_p, \delta_p). \tag{8}$$

The first decomposition is informative about the relative importance of intergenerational transmission through the projected initial skill vs. the skill growth factors, while the second decomposition clarifies the relative importance of each parental skill factor for child skills. Since parental skill factors only influence child 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 earnings of Canadian parents and their children from the Intergenerational Income Database (IID).[12] 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 tax files (T1) within four years of the

---

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

[12]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).

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 and 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 use the earnings of the eldest sons and their fathers when they are ages 26–55. We select fathers ages 16–50 when the eldest son was born. Because we do not observe earnings before 1978, fathers who were older than 55 in 1978 (born before 1923) are excluded from our sample. Birth years are more dispersed among parents due to variation in their 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 parents and children.

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

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.

Our empirical analysis focuses on log-earnings residuals $y_{i,j,t}$ and their various covariances. Within generation, age, 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, and family cohort group) in order to calculate log-earnings residuals. We calculate two sets of covariances between

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

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 contributed to each covariance ranges from 35,650 to 272,290, with the average number 153,876.[13]

While autocovariances are commonly exploited to identify the earnings process for a single generation (e.g., Lillard and Weiss, 1979), intergenerational covariances play a central role in identifying the transmission of skills across generations, as discussed in Section 2.2.[14] Although the intergenerational mobility literature often focuses directly on the earnings IGE, it rarely discusses intergenerational covariances themselves. These covariances are a transparent source of identification in our two-factor model of intergenerational transmission.

Figure 1 reports the covariances between parents' and children's log earnings when the children'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 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-

---

[13]We only consider families that contribute to any intergenerational covariances. 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.

[14]See Appendix Figures B-1 and B-2 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.
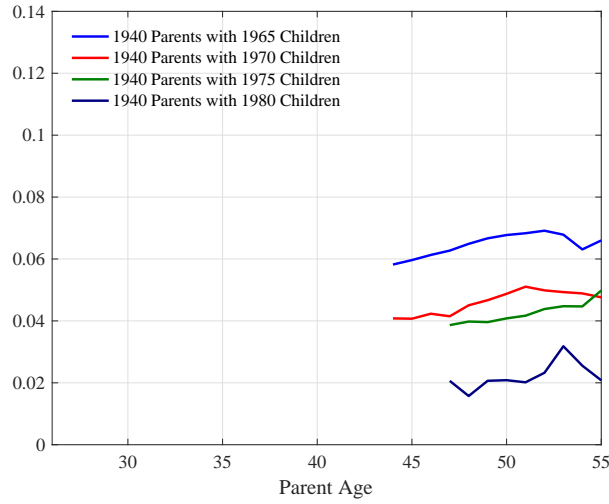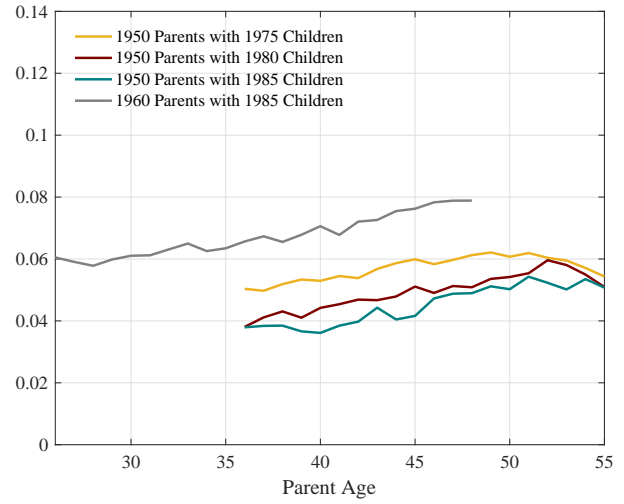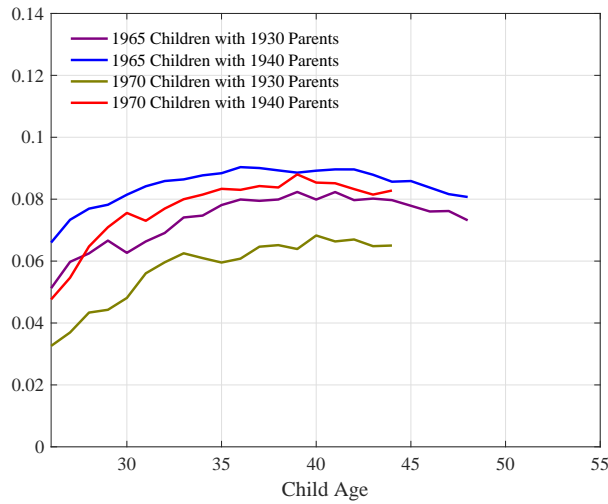
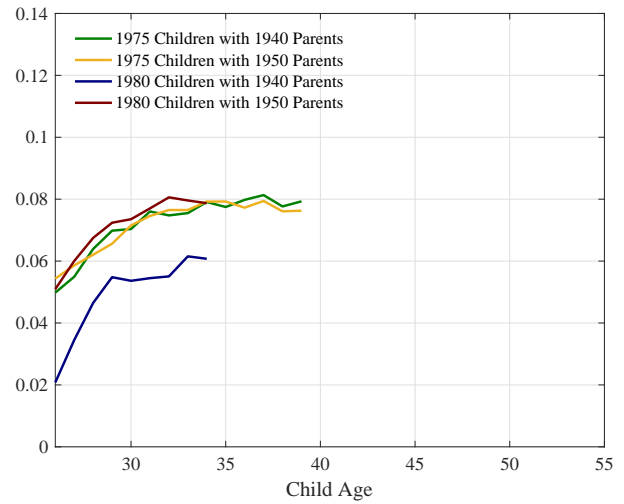(a) 1940 Parents   (b) 1950 and 1960 Parents

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

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

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

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.

The covariances in log earnings also allow us to examine how earnings *growth*, defined as $\Delta y_{i,j,t} := y_{i,j,t} - y_{i,j,t-1}$, is correlated across generations. Figures 3 and 4 show the intergenerational covariances in earnings growth by parents' and children's ages, respectively. In contrast to the covariances between log earnings, the covariances between earnings growth are centered around zero and do not exhibit clear lifecycle or cohort patterns. Indeed, we cannot reject that all of the 1,130 age- and cohort-specific intergenerational covariances of earnings growth equal zero at the .05 significance level.[15]

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

$$\text{Cov}(y_{p,t}, y_{k,t'}) = \text{Cov}(\psi_p, \psi_k) + \Lambda_{p,t}\Lambda_{k,t'}\text{Cov}(\delta_p, \delta_k) + \Lambda_{k,t'}\text{Cov}(\psi_p, \delta_k) + \Lambda_{p,t}\text{Cov}(\delta_p, \psi_k),$$

$$\text{Cov}(y_{j,t}, \Delta y_{j',t'}) = \lambda_{j',t'}\text{Cov}(\psi_j, \delta_{j'}) + \Lambda_{j,t}\lambda_{j',t'}\text{Cov}(\delta_j, \delta_{j'}), \quad \text{for } j \neq j',$$

$$\text{Cov}(\Delta y_{p,t}, \Delta y_{k,t'}) = \lambda_{p,t}\lambda_{k,t'}\text{Cov}(\delta_p, \delta_k).$$

The fact that $\text{Cov}(y_{p,t}, y_{k,t'})$ varies with parents' and children's ages implies that $\lambda_{p,t}$ and $\lambda_{k,t}$ are non-zero (for at least some ages). This coupled with the fact that $\text{Cov}(\Delta y_{p,t}, \Delta y_{k,t'}) = 0$ for all ages of parents and children indicates that the skill growth factor is uncorrelated across generations: $\text{Cov}(\delta_p, \delta_k) = 0$. However, the skill growth factor may still play an important role in earnings transmission—lifecycle variation in $\text{Cov}(y_{p,t}, y_{k,t'})$ implies that $\text{Cov}(y_{j,t}, \Delta y_{j',t'}) \neq 0$, so the skill growth factor of one generation must be correlated with the initial skill of the other.

It is noteworthy that these intergenerational covariance patterns in earnings are inconsistent with a single-factor model of skill transmission. To see this, suppose that $\theta_{i,j,t} = \chi_{j,t}\bar{\theta}_{i,j}$ as assumed by much of the recent literature (Jenkins, 1987; Grawe, 2006; Haider and Solon, 2006; Nybom and Stuhler, 2016).[16] This single-factor model implies that $\text{Cov}(y_{p,t}, y_{k,t'}) = \chi_{p,t}\chi_{k,t'}\text{Cov}(\bar{\theta}_p, \bar{\theta}_k)$, so any variation 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 earnings growth (i.e., $\text{Cov}(\Delta y_{p,t}, \Delta y_{k,t'}) = \Delta\chi_{p,t}\Delta\chi_{k,t'}\text{Cov}(\bar{\theta}_p, \bar{\theta}_k)$) for all ages (see Figure 4).[17]

---

[15]For some perspective, the largest correlation (in absolute value) is less than 0.017. We separately test whether each of the 1,130 covariances of the earnings growth in our sample equals zero using the bootstrap approach in Section 2 of MacKinnon (2009). Based on 100 bootstrap samples, the symmetric bootstrap p-values range from 0.37 to 1, while the equal-tail bootstrap p-values range from 0.72 to 1.

[16]If $\bar{\theta}_{i,j}$ reflects the 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).

[17]Haider and Solon (2006) initially 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 $u_{i,j,t} := y_{i,j,t} - \chi_{j,t}\bar{\theta}_{i,j}$ where $\chi_{j,t}$ is the linear-projection coefficient and $u_{i,j,t}$ is uncorrelated with $\bar{\theta}_{i,j}$. 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 reverse regression) provide an incomplete characterization of the lifecycle bias for IGEs in $\bar{\theta}_{i,j}$ 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
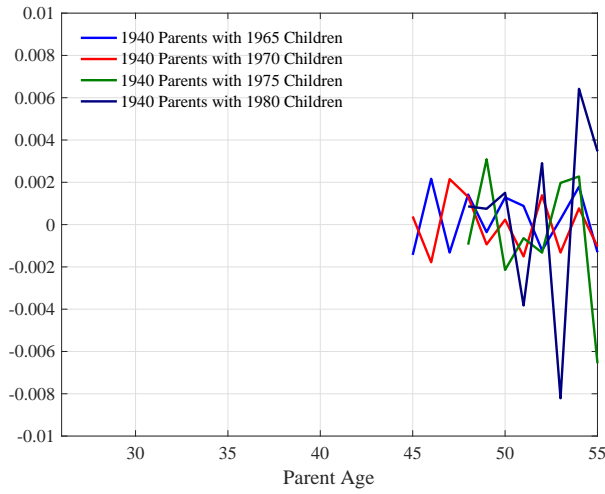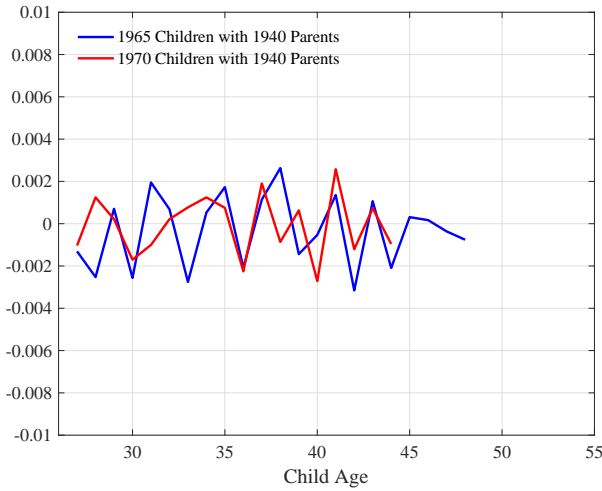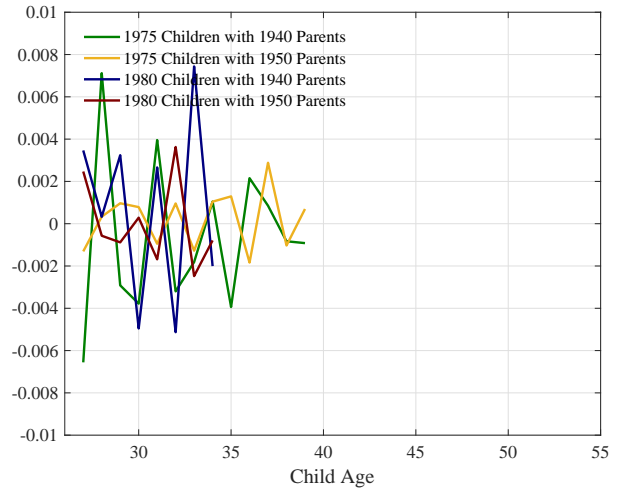
(a) 1940 Parents
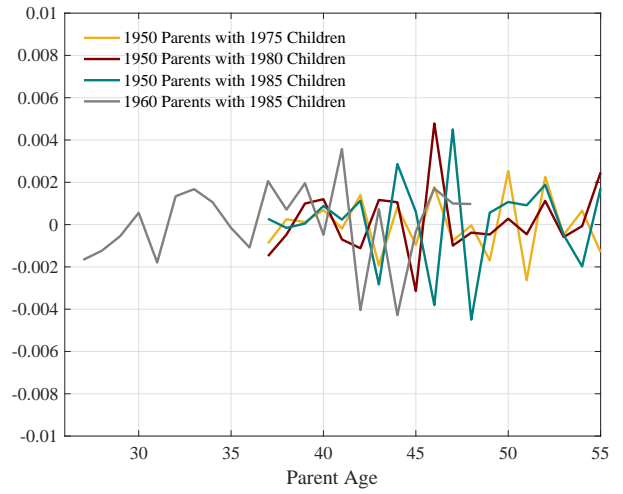
(b) 1950 and 1960 Parents

Figure 3: Intergenerational Covariance of 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 Earnings Growth by Children's Age (at Parents' Age 55)

14

The discussion above suggests that intergenerational covariances in 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 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 the log earnings and the theoretical covariances implied by the model. The parameters related to the dynamic process and the variances of the non-skill shocks, the cross-sectional distributions of the 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 the initial skills and the skill growth factor are identified from the 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 only for 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 the 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 $\boldsymbol{\Omega}_p$, (ii) the intergenerational skill covariance matrix $\boldsymbol{\Omega}_{p,k}$, and (iii) the covariance matrix for children's skills

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

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

Haider and Solon (2006).

$\mathbf{\Omega}_p$, $\mathbf{\Omega}_{p,k}$, and $\mathbf{\Omega}_k$ separately for each family cohort group. Based on the findings in Section 3 that $\text{Cov}(\Delta y_p, \Delta y_k) = 0$, we assume $\text{Cov}(\delta_p, \delta_k) = 0$.

The variances of persistent and transitory shocks are estimated separately for each generation, age, and family cohort group; however, we need to impose some normalizations. 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.[18]

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 C.

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 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

---

[18]We do not observe earnings at age 27 for most of the parent cohort groups. By assuming that $\Lambda_{j,t}$ does not vary across cohort groups, we avoid setting the scale of the skill growth factor based on different ages for different cohort groups. This makes it easier to compare the estimated variances/covariances of the skill factors across cohort groups.

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.
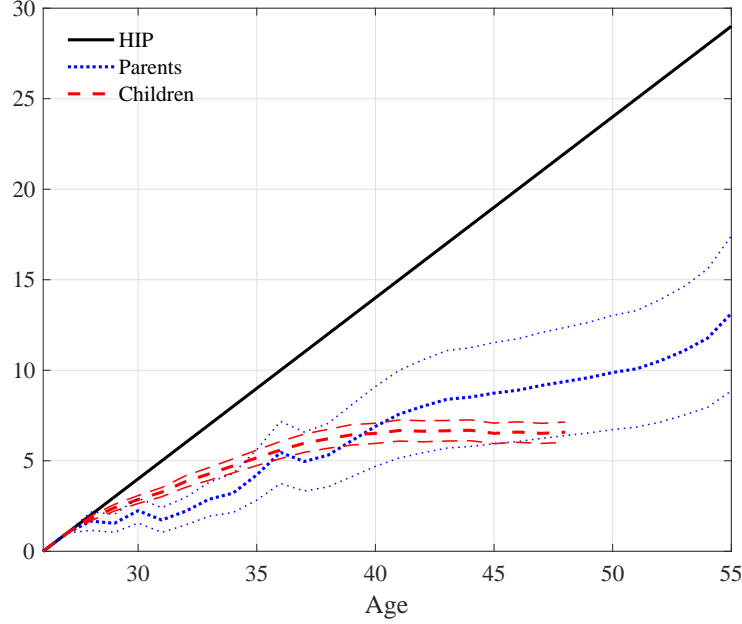

Figure 5: Coefficient on the Skill Growth Factor ($\Lambda_{j,t}$, Thick Lines) and 95% Confidence Interval (Thin Lines)

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 over the lifecycle, with no further growth heterogeneity beyond age 40 for the child's generation. We continue to observe some increase in the variance of skill growth for the generation of fathers, but it is much less than at the beginning of their careers.

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 $\text{Corr}(\delta_p, \psi_k)$ but the opposite for $\text{Corr}(\psi_p, \delta_k)$, with both of these correlations ranging from 0.14 to 0.26.[19] See Appendix C for other cohorts and estimates of (standardized) linear-projection coefficients **A**.

Finally, Appendix Figures C-8 and C-9 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).

---

[19]Recall our assumption $\text{Corr}(\delta_p, \delta_k) = 0$, given the lack of any systematic relationship between son and father earnings growth.

5.2 Intergenerational Earnings and Skill Elasticities

Before we discuss the intergenerational transmission of skills, we document basic earnings IGEs for each 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.[20] 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 Figure 6 (or Table 3). First, as previously noted by Corak and Heisz (1999) and Chen, Ostrovsky, and Piraino (2017), earnings IGEs in Canada are quite low relative to those of the U.S. None of the earnings IGEs exceeds 0.2, with annual earnings IGEs as low as 0.06. Second, earnings IGEs are declining for more recent cohorts of children. For 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 of sons. 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. Fourth, 5-year average earnings IGEs are about 40% larger than annual IGEs, while 9-year averages are another 5–10% larger.

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. Table 3 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 the 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 don't 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 also reported in Table 3 and are graphed in Figure 7.[21] These skill IGEs show similar

---

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

[21]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.

	Cohort Group							
Parents	1940				1950			1960
Children	1965	1970	1975	1980	1975	1980	1985	1985
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		
	(0.01)	(0.007)	(0.012)	(0.017)	(0.007)	(0.006)		
$\mathrm{Var}(\theta_{p,t})/\mathrm{Var}(y_{p,t})$	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)	
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		
	(0.01)	(0.007)	(0.012)	(0.017)	(0.007)	(0.006)		
$\mathrm{Var}(\bar{\theta}_p)/\mathrm{Var}(\bar{y}_p)$	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)
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)		
$\mathrm{Var}(\bar{\theta}_p)/\mathrm{Var}(\bar{y}_p)$	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)
D. 13-Year Average								
$\mathrm{Var}(\bar{\theta}_p)/\mathrm{Var}(\bar{y}_p)$					0.712	0.705	0.703	0.502
					(0.004)	(0.003)	(0.007)	(0.007)
E. 23-Year Average								
$\mathrm{Var}(\bar{\theta}_p)/\mathrm{Var}(\bar{y}_p)$								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. 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 and Father's Share of Earnings Variance from Skills by Cohort
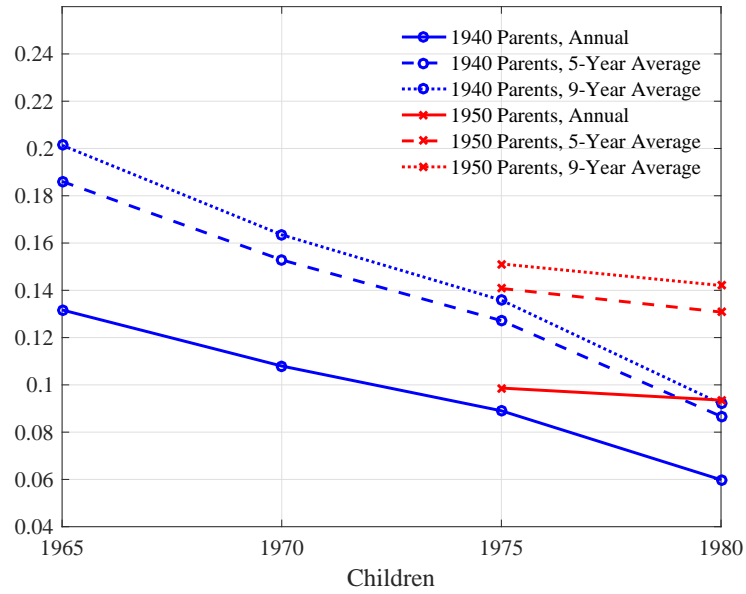
Figure 6: Earnings IGE
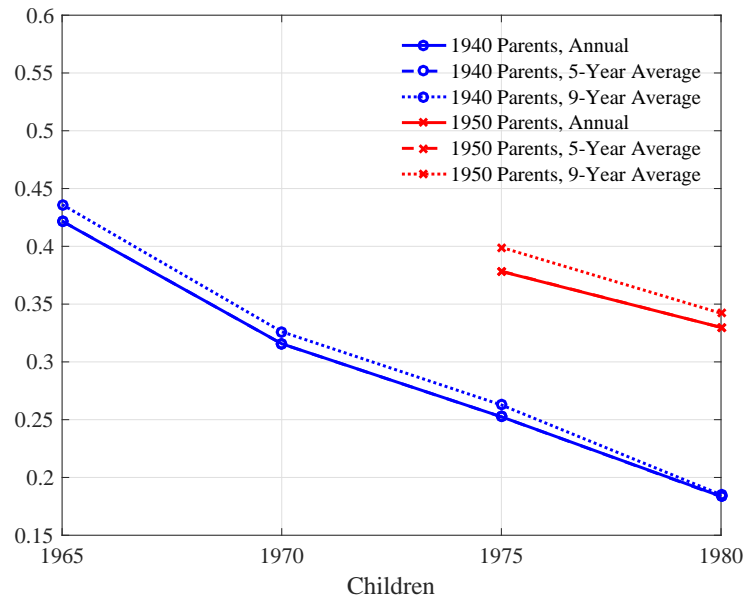
Figure 7: Skill IGE

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 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 variations 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 the earnings variation explained by skills for all cohorts of children with fathers from the 1940 or 1950 cohorts. We report these shares for annual measures (at age 30), 5-year averages (ages 28–32), 9-year averages (ages 26–34), 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 and reach a maximum of around 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 the variation in projected skill levels, $\hat{\theta}_k$, for annual measures as well as averages over several years. Figure 8 displays the estimated shares of the skill variation for all cohorts of children with fathers from the 1940 and 1950 cohorts.[22] 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

---

[22]Recall that the share of skill variation is simply the share of earnings variance explained by the projected skills divided by the share of the children's earnings variance that is explained by their skills.

	Cohort Group							
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)	(0.05)	(0.017)	(0.021)	(0.012)

B. 5-Year Average

$\mathrm{Var}(\bar{\theta}_k)/\mathrm{Var}(\bar{y}_k)$	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)	(0.01)	(0.007)	(0.005)
$\mathrm{Var}(\bar{\hat{\theta}}_k)/\mathrm{Var}(\bar{y}_k)$			0.150	0.077	0.082	0.030	0.100	0.076
			(0.022)	(0.013)	(0.018)	(0.006)	(0.008)	(0.004)
$\mathrm{Var}(\bar{\hat{\theta}}_k)/\mathrm{Var}(\bar{\theta}_k)$			0.389	0.212	0.225	0.085	0.264	0.202
			(0.056)	(0.036)	(0.05)	(0.017)	(0.021)	(0.012)

C. 9-Year Average

$\mathrm{Var}(\bar{\theta}_k)/\mathrm{Var}(\bar{y}_k)$	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)	(0.01)	(0.008)	(0.005)
$\mathrm{Var}(\bar{\hat{\theta}}_k)/\mathrm{Var}(\bar{y}_k)$			0.174	0.087	0.093	0.034	0.114	0.086
			(0.026)	(0.015)	(0.021)	(0.007)	(0.009)	(0.005)
$\mathrm{Var}(\bar{\hat{\theta}}_k)/\mathrm{Var}(\bar{\theta}_k)$			0.395	0.214	0.228	0.086	0.268	0.204
			(0.057)	(0.037)	(0.051)	(0.017)	(0.022)	(0.012)

D. 13-Year Average

$\mathrm{Var}(\bar{\theta}_k)/\mathrm{Var}(\bar{y}_k)$	0.543	0.519	0.550	0.522	0.516		0.528	
	(0.004)	(0.006)	(0.003)	(0.004)	(0.006)		(0.006)	
$\mathrm{Var}(\bar{\hat{\theta}}_k)/\mathrm{Var}(\bar{y}_k)$			0.198	0.105	0.111		0.130	
			(0.029)	(0.018)	(0.024)		(0.01)	
$\mathrm{Var}(\bar{\hat{\theta}}_k)/\mathrm{Var}(\bar{\theta}_k)$			0.360	0.201	0.215		0.245	
			(0.052)	(0.033)	(0.047)		(0.019)	

E. 23-Year Average

$\mathrm{Var}(\bar{\theta}_k)/\mathrm{Var}(\bar{y}_k)$	0.702		0.702					
	(0.003)		(0.003)					
$\mathrm{Var}(\bar{\hat{\theta}}_k)/\mathrm{Var}(\bar{y}_k)$			0.225					
			(0.032)					
$\mathrm{Var}(\bar{\hat{\theta}}_k)/\mathrm{Var}(\bar{\theta}_k)$			0.321					
			(0.046)					

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

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

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 the skill variation explained by 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 (5) makes clear, intergenerational transmission will explain a greater share of the earnings variation when skills are a more important component of children's earnings.[23] Figure 9 shows that the fraction of the earnings variation explained by skills grows steadily over the first half of children's careers.[24] Figure 10 documents a very similar lifecycle pattern for the share of the 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).

While 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 the skill variation explained by projected skills—equivalently, the share of the 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. 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 the skills themselves.

To better understand these patterns, we consider variance decompositions that shed light on the roles of the initial skill and the skill growth heterogeneity. Figure 12 shows the contribution of initial skills and skill growth rates (as well as their covariance) to the variance of skills over the careers of children from the 1965 and 1970 cohorts (both with fathers from the 1940 cohort). 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

---

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

[24]By contrast, Appendix Figures D-11 and D-12 show that the share of the earnings variation explained by skills declines over the later part of the lifecycle, at least for the generation of fathers.

Figure 8: $\mathrm{Var}(\bar{\bar{\theta}}_k)/\mathrm{Var}(\bar{\theta}_k)$

(a) 1965 and 1970 Children
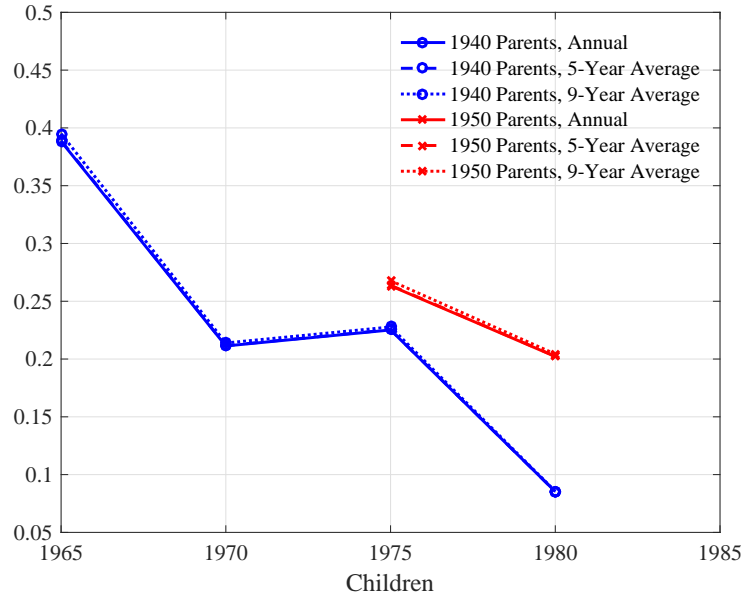
(b) 1975 and 1980 Children
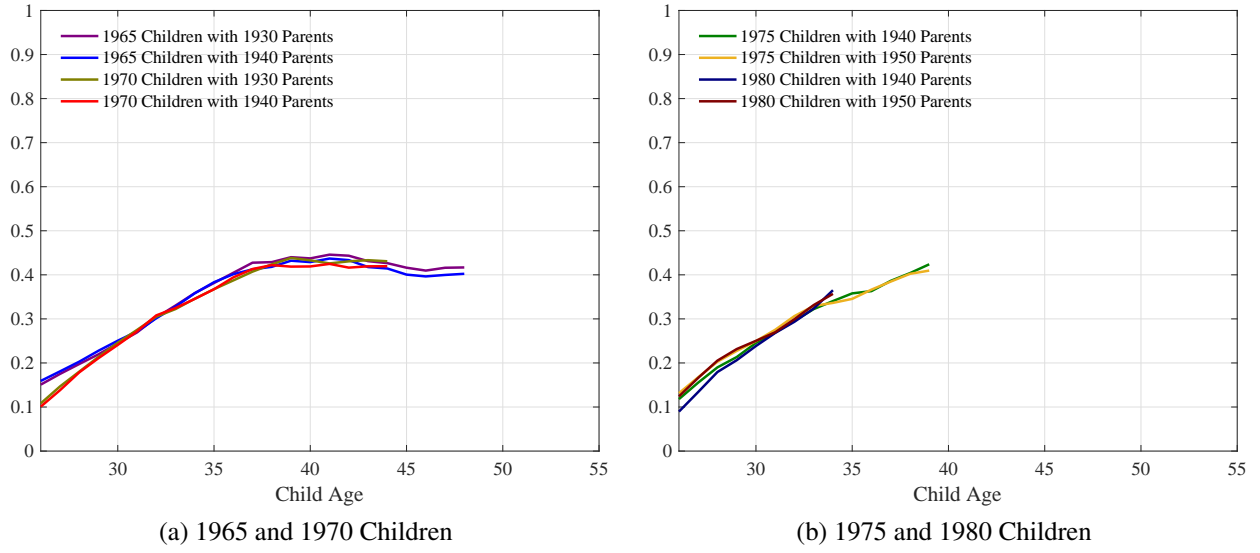
Figure 9: $\mathrm{Var}(\theta_{k,t})/\mathrm{Var}(y_{k,t})$
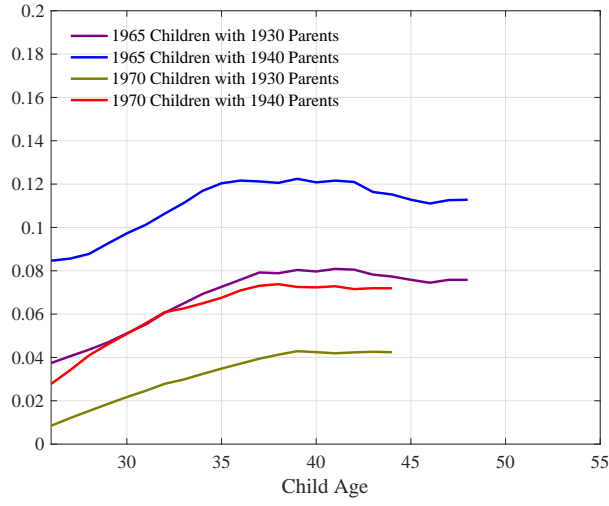
(a) 1965 and 1970 Children
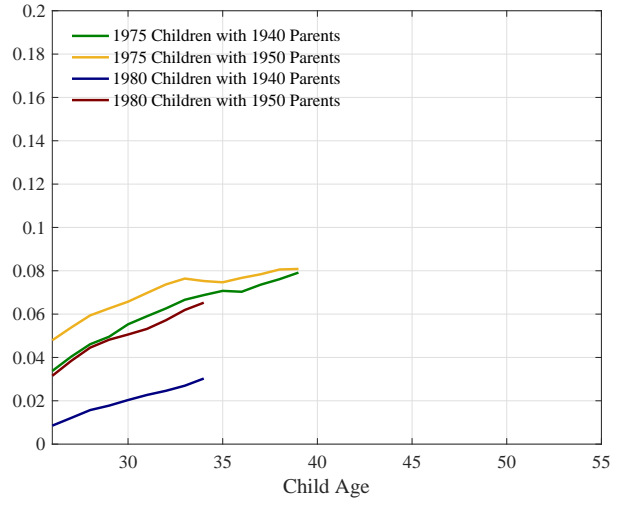
(b) 1975 and 1980 Children

Figure 10: $\mathrm{Var}(\hat{\theta}_{k,t})/\mathrm{Var}(y_{k,t})$
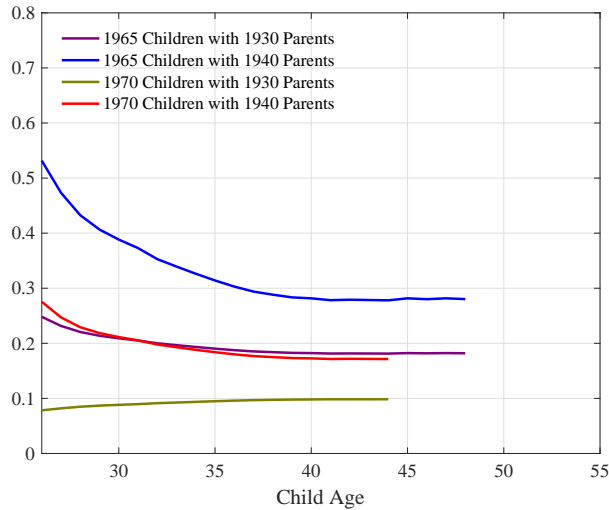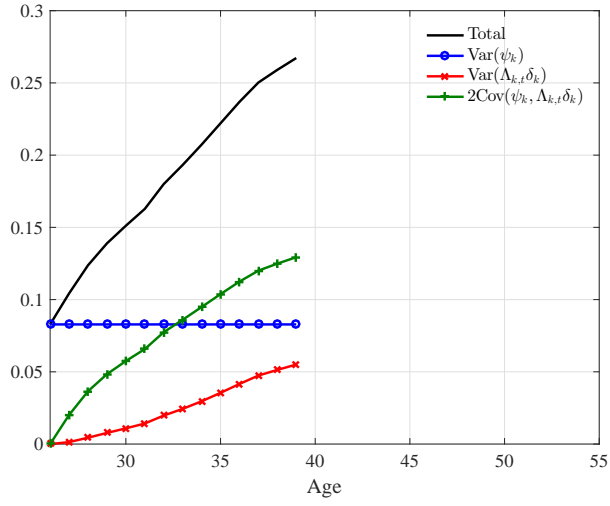

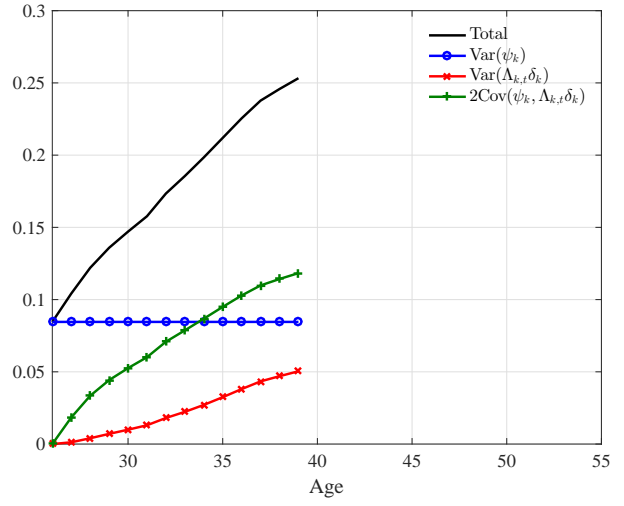(a) 1965 and 1970 Children
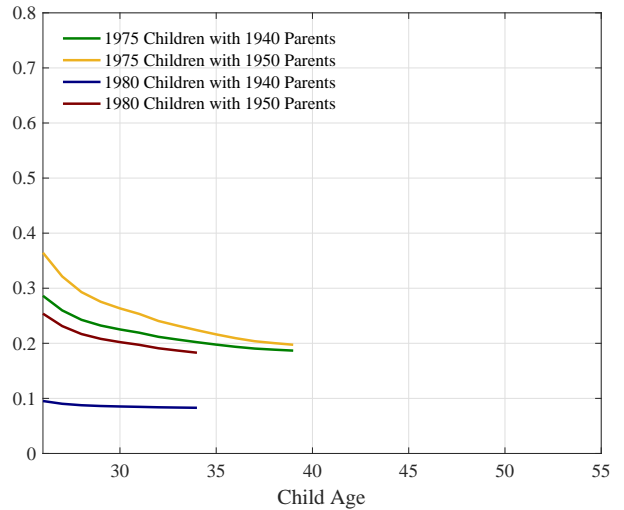
(b) 1975 and 1980 Children

Figure 11: $\mathrm{Var}(\hat{\theta}_{k,t})/\mathrm{Var}(\theta_{k,t})$
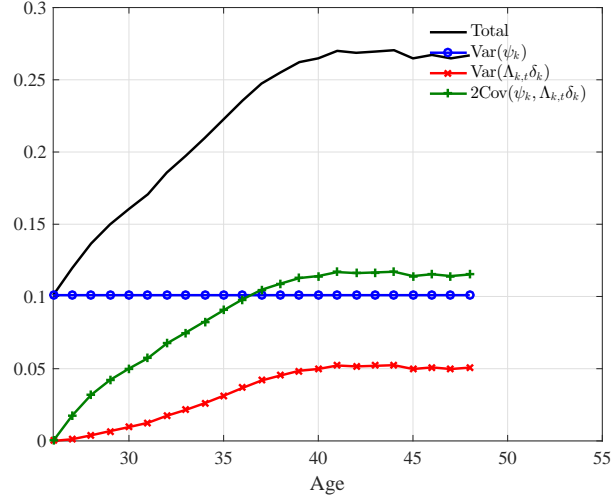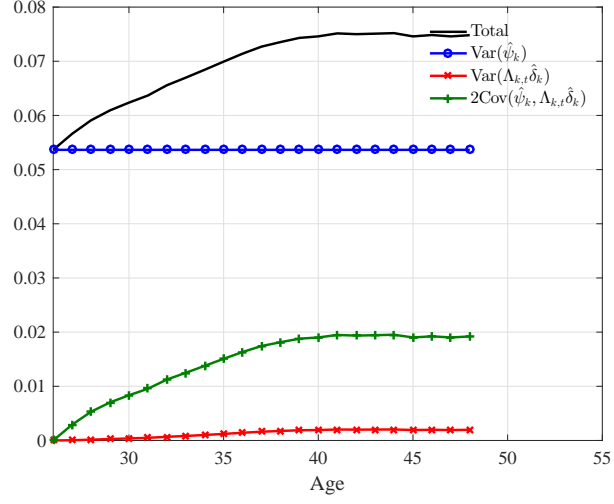
(a) 1965 Children, 1940 Parents

(b) 1975 Children, 1940 Parents

(c) 1975 Children, 1950 Parents

Figure 12: Skill Variance Decomposition for Children

(a) 1965 Children, 1940 Parents
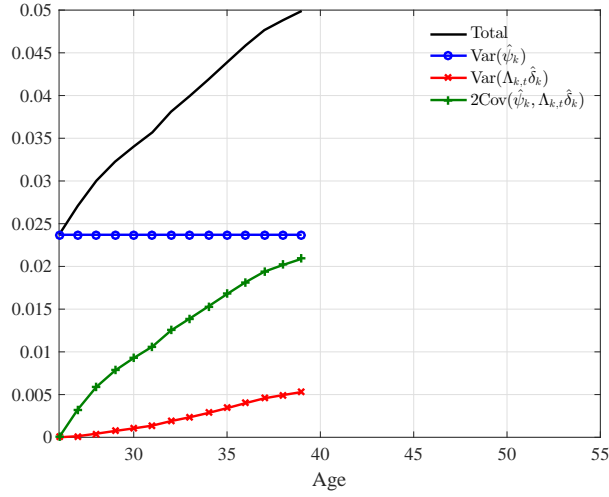
(b) 1975 Children, 1940 Parents
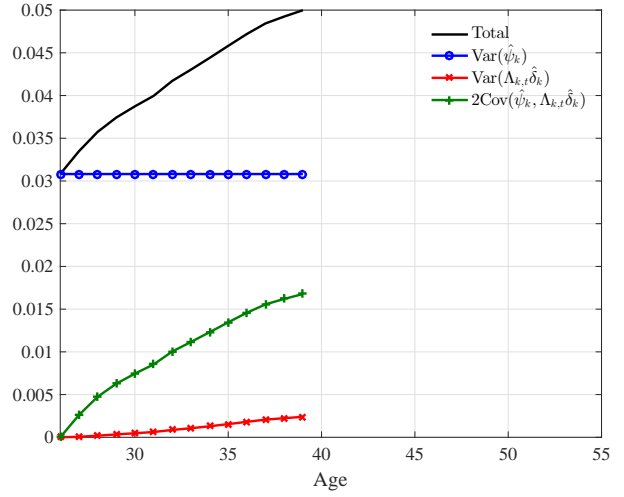
(c) 1975 Children, 1950 Parents

Figure 13: Projected Skill Variance Decomposition for Children
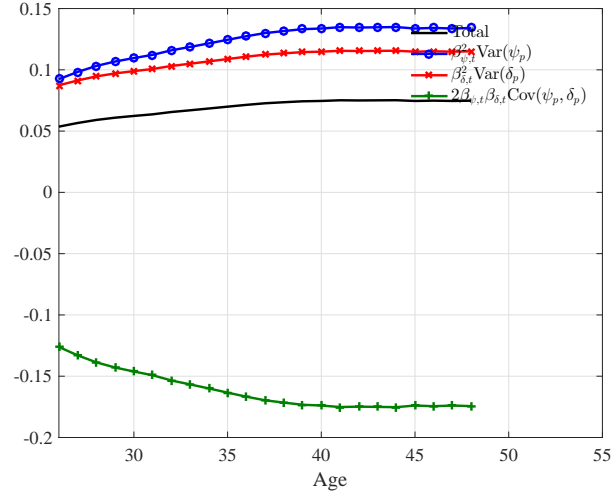
(a) 1965 Children, 1940 Parents
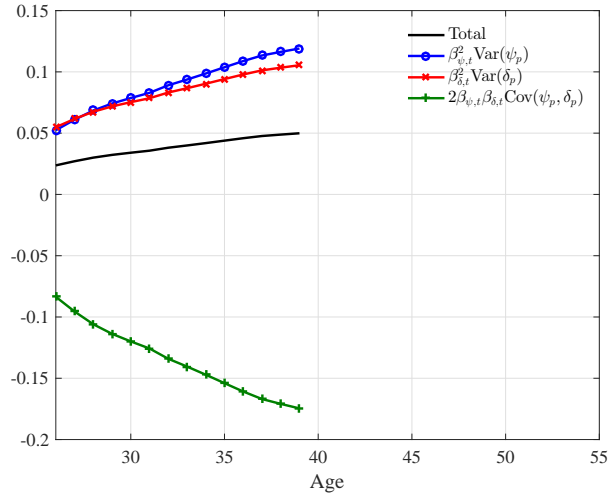
(b) 1975 Children, 1940 Parents
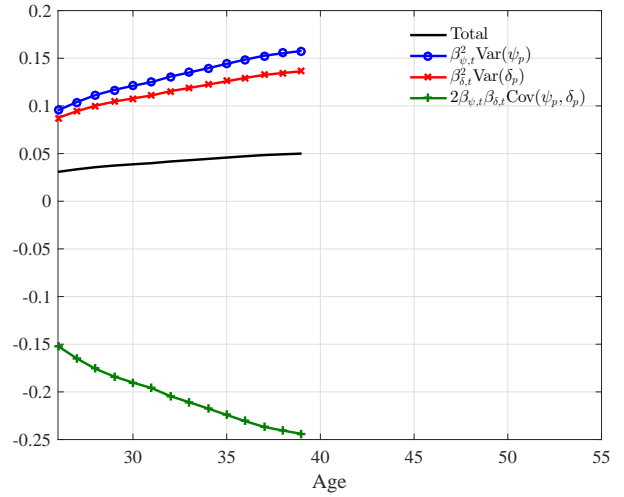
(c) 1975 Children, 1950 Parents

Figure 14: Alternative Projected Skill Variance Decomposition for Children

their covariance) yields very similar lifecycle trajectories (Figure 13); however, predicted initial skill levels explain much more of the variation in projected skills. Thus, the intergenerational transmission of initial skills is a 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 the 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. See equation (8). Figure 14 shows that the variation in parental initial skills and skill growth plays roughly an equal role in explaining the level and growth in children's projected skill variation. Because ψ_p and δ_p are (strongly) negatively correlated, their covariation offsets much of the positive effects of the variation in the two skill factors.

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 the variation in parental skills alone. In a single-factor model of intergenerational transmission, the variation in parental skills fully explains all of the variation in the projected skills (i.e., $\text{Var}(\hat{\theta}_{k,t}|\theta_{p,t'}) = 0$); however, this is not the case in our two-factor model.[25] 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

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

where the second equality follows from the fact that $\text{Var}(\hat{\theta}_{k,t}|\theta_{p,t'})$ 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 as

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

Appendix Figure D-10 shows that this fraction ranges from 0.28 to 0.55 for different cohorts (at child's age 30 and parent's age 50), suggesting that knowledge of parental skill levels alone explains no more than 55% of the variation in the projected child skill levels. While predicting a child's skill level is already imprecise given information about their parent's initial skill and skill growth factor, that prediction becomes substantially less precise when only information about the parent's skill level is available.

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, forthcoming). This literature has further emphasized the positive correlation

---

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

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). 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 address when the parent-child linkage is formed.[26] To ensure sufficient sample sizes, we consider families living 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., 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, $s_{p,t}$.[27] 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.[28]

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).[29] Because sample sizes are small for some of the smaller cities, some of the estimates (especially estimated skill IGEs) are imprecise.[30] 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).[31] 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

---

[26]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 population center.

[27]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).

[28]We use 200 bootstrap samples separately for each city to conduct inference using MD estimates of $s_{p,t}$.

[29]Appendix Figures E-14 and E-15 show analogous results for annual and 9-year averages.

[30]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.
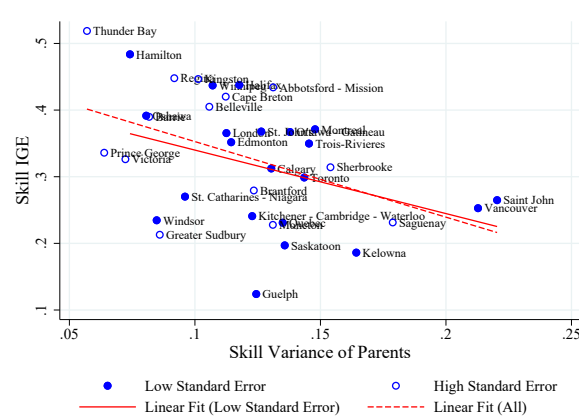
[31]"Low Standard Error" cities include the 15 largest cities, plus 6 other cities spread throughout the next 20 largest cities.
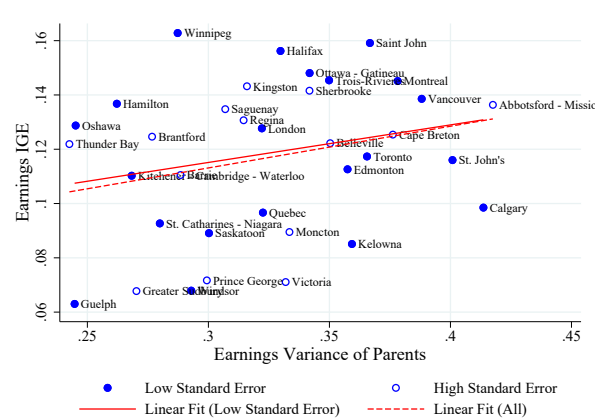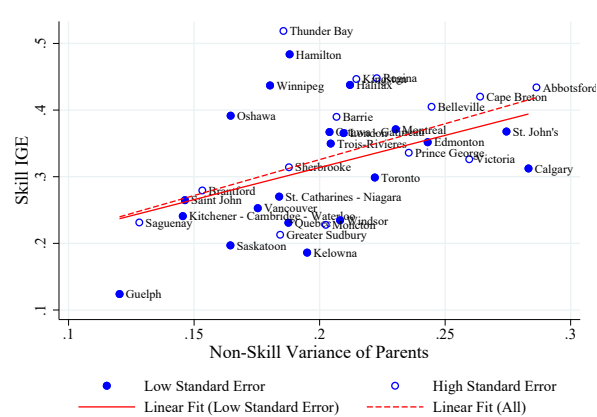
(a) Earnings IGE vs. Earnings Variance
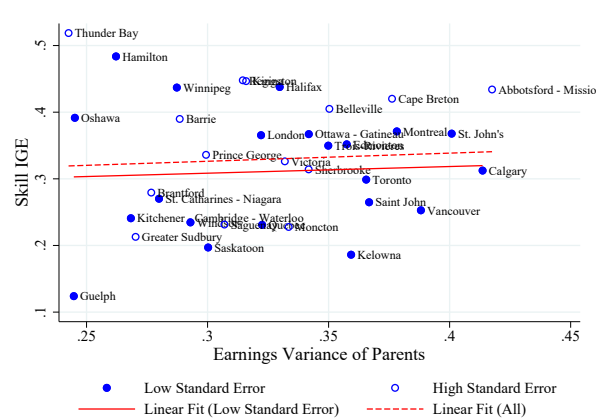
(b) Skill IGE vs. Earnings Variance

(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
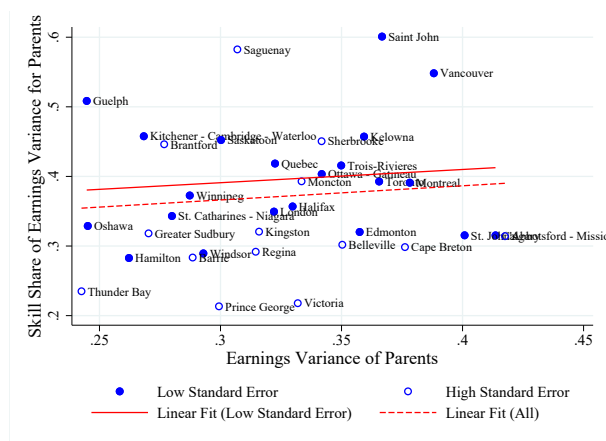

Figure 16: Skill Share of Earnings Variance vs. Earnings Variance: 5-Year Averages

inequality may not be too distorted by sampling variation of the estimators; however, we explore this issue more formally below.

Variables of Regression		(1) (2) Uncorrected			(3) (4) Bias-Corrected		
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 by the Great Gatsby Curve and is largely consistent with the conclusions of Corak (forthcoming), 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. Indeed, the 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.[32]

Following MacKinnon (2002), we use the bootstrap approach to correct the finite-sample bias in

---

[32] All slope coefficients in Figure 15 and column (1) of Table 5 may be biased; this is because (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 the non-linear estimation (MD).

point estimates and confidence intervals for the slope estimates in Figures 15 and 16.[33] 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 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).[34] 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.

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 intergenerational mobility, whereas inequality in earnings conditional on skills appears to be the driving force for the Great Gatsby phenomenon.

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 are an important component of earnings but that other idiosyncratic factors (some 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 rates, where interpersonal differences in skill growth decline over the course of individuals' careers (as human capital theory predicts). As a simple metric for characterizing the importance of both skill factors, we show that knowledge of only parental skill levels at a given age explains no more than 55% of the variation in children's projected skill levels. Consequently, while parents' initial skill and skill growth rates explain 20–40% of the variation in children's skill (10–20% of their earnings variation)

---

[33]Let $\hat{\beta}$ be a slope estimate. For each bootstrap sample $b = 1, 2, \ldots, 200$, we calculate the bootstrap version of the slope estimate, β_b^*. Let $\overline{\beta}^*$ and s_β^* be the mean and standard deviation of β_b^* over 200 bootstrap samples. Then the bias-corrected point estimate is $2\hat{\beta} - \overline{\beta}^*$ and the 90% confidence interval is $[2\hat{\beta} - \overline{\beta}^* - 1.645 \times s_\beta^*, 2\hat{\beta} - \overline{\beta}^* + 1.645 \times s_\beta^*]$.

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

for most cohorts, knowledge of only parental skill levels 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 the skills themselves. Our estimates also reveal that the intergenerational transmission of 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. Yet, we also find that parents' initial skills and skill growth rates are equally important for children's skill levels and lifecycle trajectories, in large part 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 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. We find a positive relationship between the earnings IGEs and the variance of parental earnings across cities, so earnings mobility is decreasing in earnings inequality, consistent with the Great Gatsby Curve. 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. In short, skill diversity fosters mobility, while instability inhibits it.

References

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.

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.

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.

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.

———. forthcoming. "The Canadian Geography of Intergenerational Income Mobility." *The Economic Journal* .

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.

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 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.

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.

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.

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

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.

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

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

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.

———. 2009. "Bootstrap hypothesis testing." In *Handbook of Computational Econometrics*, edited by David A. Belsley and Erricos John Kontoghiorghes, chap. 6. Chichester: Wiley, 183–213.

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.

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

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

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

$$\Delta y_{i,j,t} - \rho_j\Delta y_{i,j,t-1} = (\lambda_{j,t} - \rho_j\lambda_{j,t-1})\delta_{i,j} + \Delta\nu_{i,j,t} + \sum_{l=0}^{\min\{q,t-\underline{t}\}} \kappa_{j,l}(\Delta\xi_{i,j,t-l} - \rho_j\Delta\xi_{i,j,t-1-l}),$$

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

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

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

For any l,

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

Therefore,

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

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

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

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 $t' - t > q + 2$,

$$\text{Cov}\left(\Delta y_{j,t}, \Delta y_{j,t'} - \rho_j \Delta y_{j,t'-1} | c_j\right) = \lambda_{j,t}(\lambda_{j,t'} - \rho_j \lambda_{j,t'-1}) \text{Var}(\delta_j | c_j), \tag{9}$$

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

$$\text{Cov}\left(\Delta y_{j,t}, \Delta y_{j,t'-1} - \rho_j \Delta y_{j,t'-2} | c_j\right) = \lambda_{j,t}(\lambda_{j,t'-1} - \rho_j \lambda_{j,t'-2}) \text{Var}(\delta_j | c_j). \tag{11}$$

From (9) and (10), 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 (9) and (11), we have

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

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 $\text{Var}(\delta_j | c_j)$ Once ρ_j and $\lambda_{j,t}$ are identified, $\text{Var}(\delta_j | c_j)$ for all j and c_j can be identified from (9).

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

$$\text{Cov}(y_{j,t}, y_{j',t'} | c_j, c_{j'}) = \text{Cov}(\psi_j, \psi_{j'} | c_j, c_{j'}) + \Lambda_{j,t} \lambda_{j',t'} \text{Cov}(\delta_j, \delta_{j'} | c_j, c_{j'})$$
$$+ \Lambda_{j,t} \text{Cov}(\delta_j, \psi_{j'} | c_j, c_{j'}) + \lambda_{j',t'} \text{Cov}(\psi_j, \delta_{j'} | c_j, c_{j'}).$$

B   Additional Covariances (National Sample)
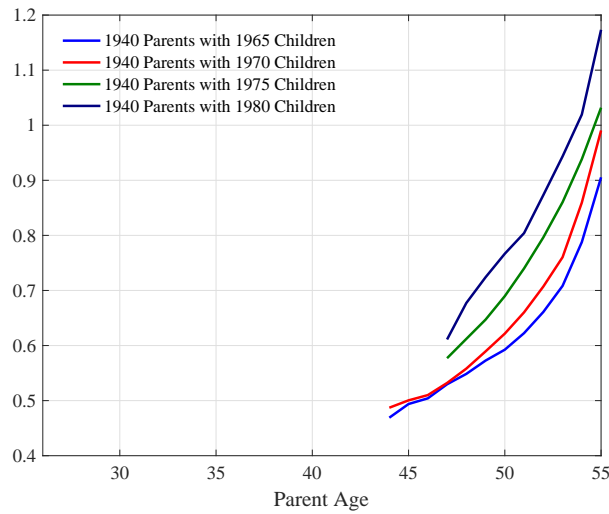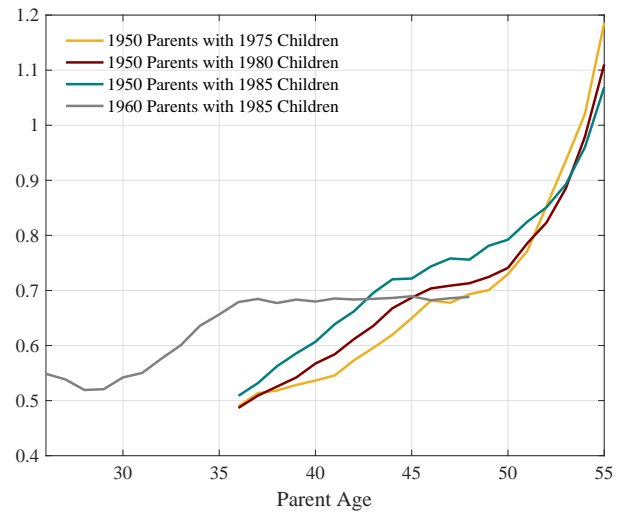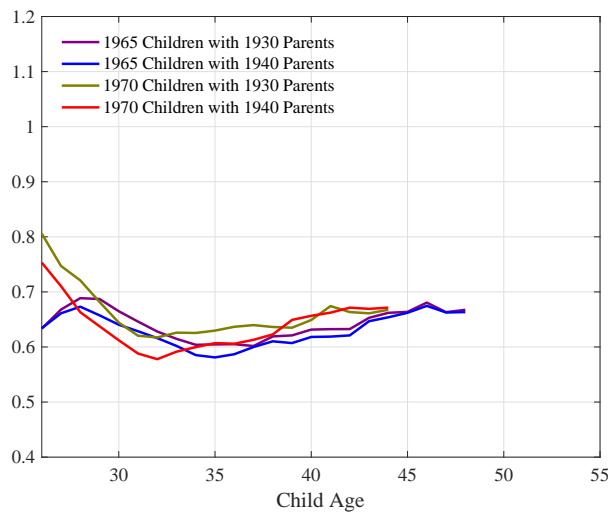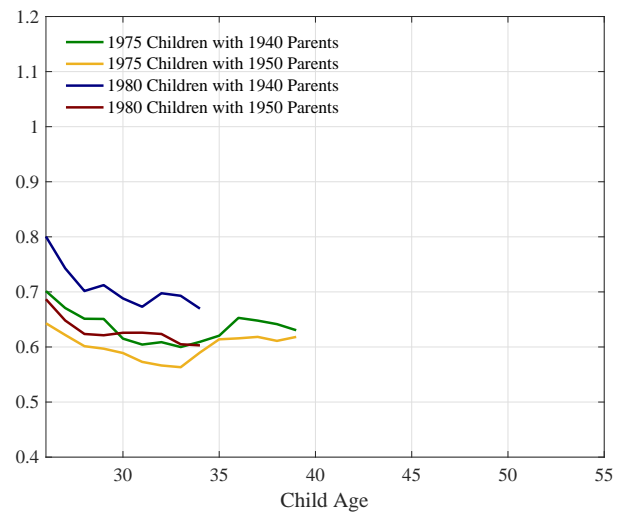
Figure B-1: Variance of Earnings by Cohort
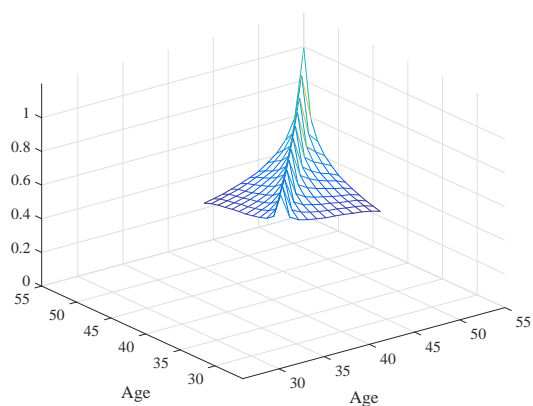
(a) Parents
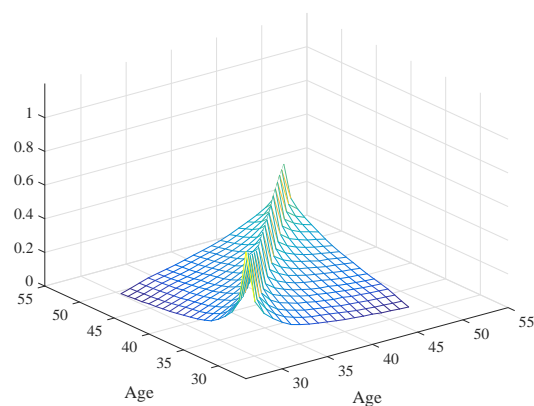
(b) Children

Figure B-2: Autocovariances for 1970 Children and 1940 Parents
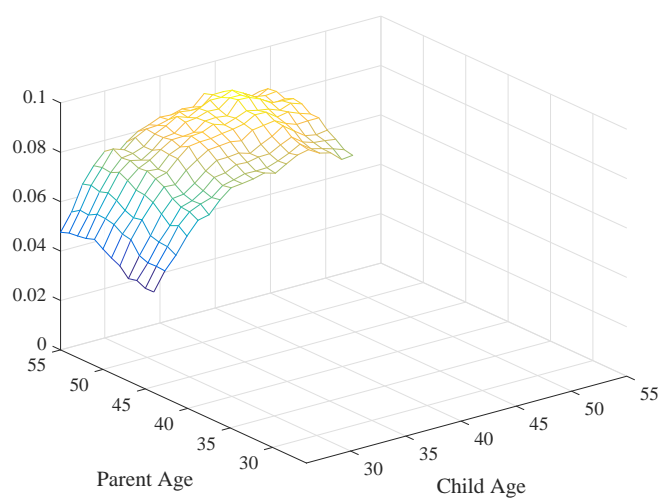
Figure B-3: Intergenerational Covariances for 1970 Children and 1940 Parents

44

C MD Parameter Estimates (National Analysis)

This appendix reports key parameter estimates from our MD approach using the Canadian national sample.[35] Table C-1 shows estimated AR(1) and MA(1) coefficients and their standard errors for the persistent and transitory components. Figure C-4 displays estimated covariances of skill factors (within generation) for fathers and sons. Figure C-5 displays intergenerational covariance estimates for the skill factors, while Figure C-6 reports the corresponding intergenerational correlations.

Figure C-7 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 C-8 and C-9 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 C-1: AR(1) and MA(1) Coefficients

---

[35]The minimized criterion function for our estimates (see Section 4) is SSR = 0.049.
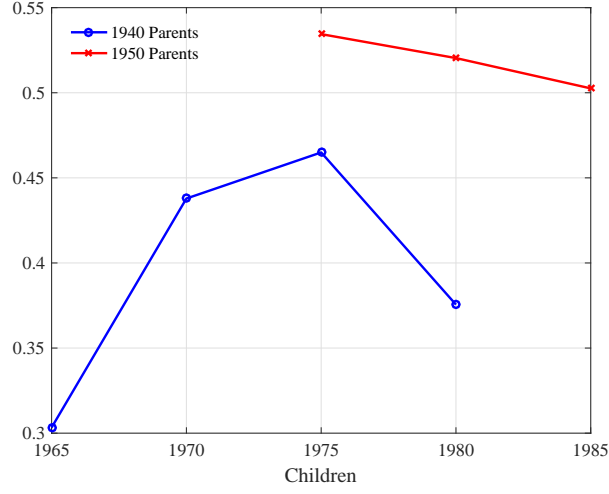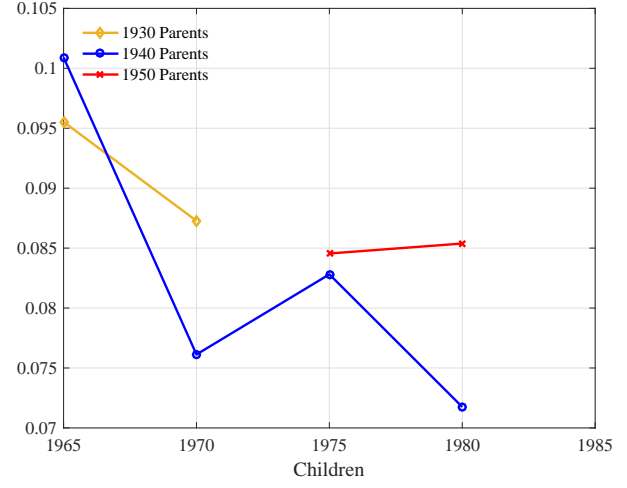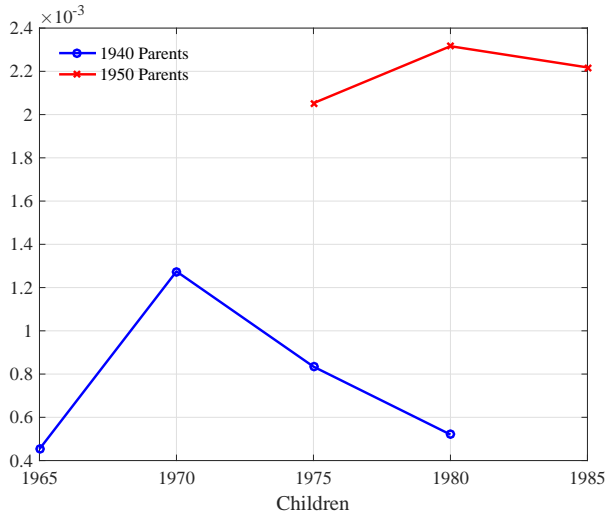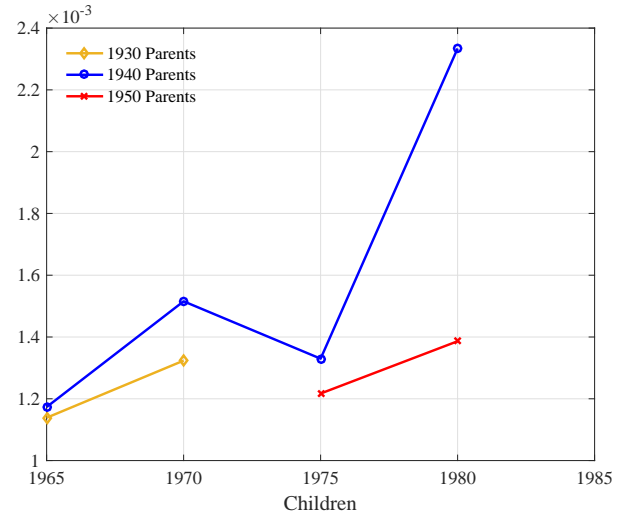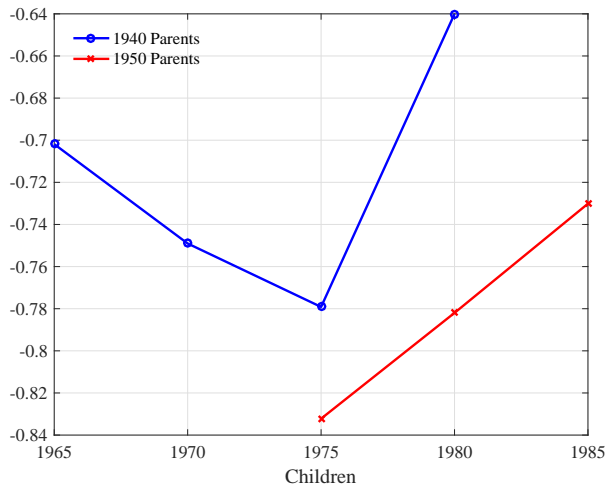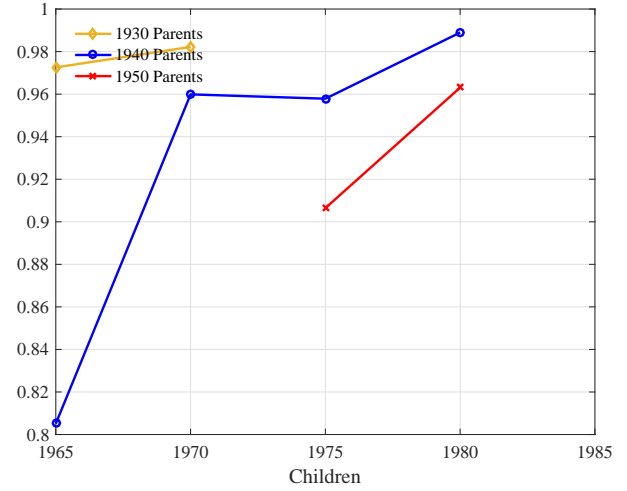
Figure C-4: Covariances of Skill Factors ($\mathbf{\Omega}_p$ and $\mathbf{\Omega}_k$)

(a) $\text{Cov}(\psi_p, \psi_k)$

(b) $\text{Cov}(\delta_p, \psi_k)$

(c) $\text{Cov}(\psi_p, \delta_k)$
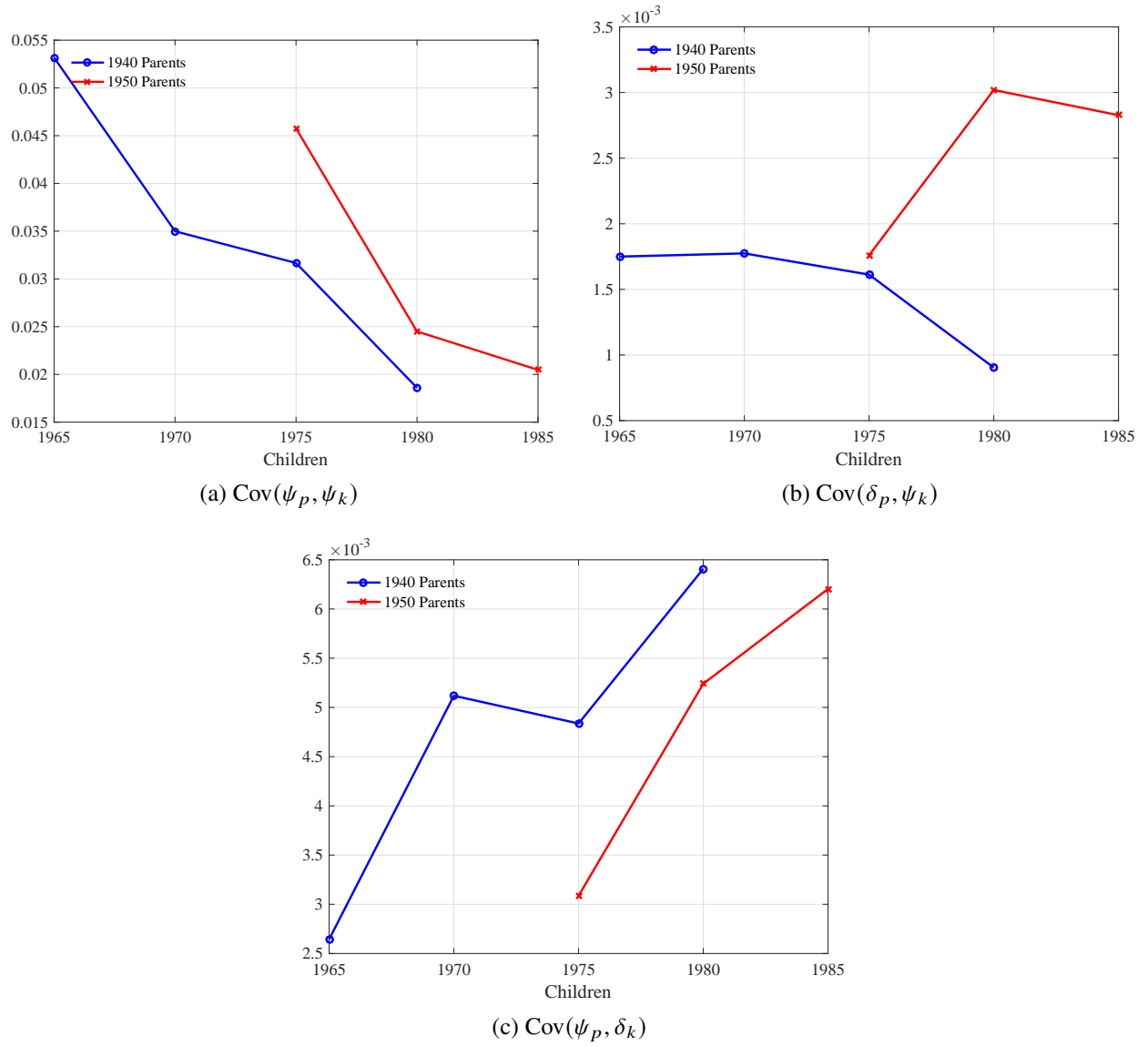
Figure C-5: Intergenerational Covariances of Skill Factors ($\mathbf{\Omega}_{p,k}$)

(a) Corr(ψ_p, ψ_k)

(b) Corr(δ_p, ψ_k)

(c) Corr(ψ_p, δ_k)
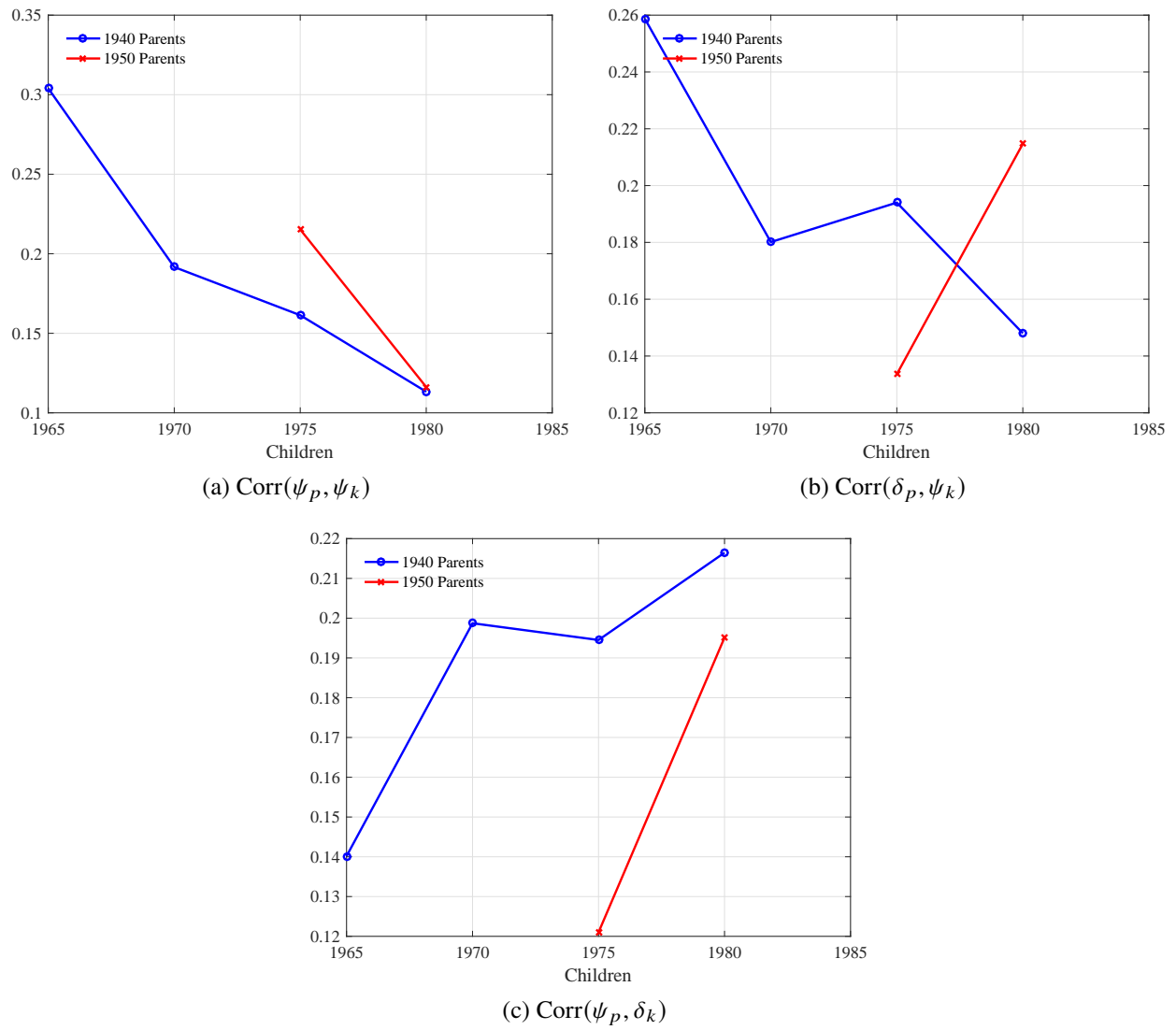
Figure C-6: Intergenerational Correlations of Skill Factors
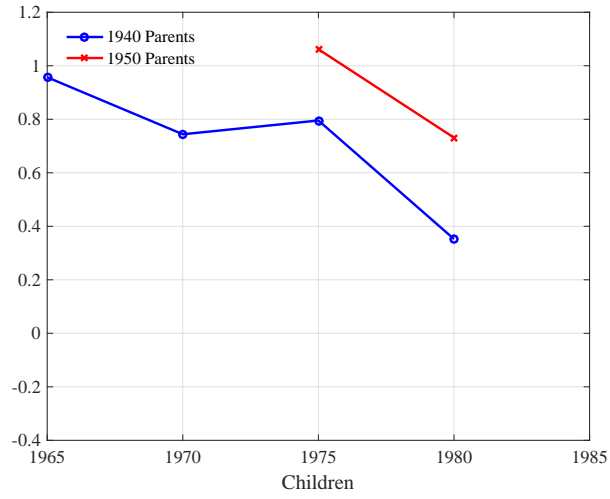
(a) $\alpha^*_{\psi,\psi}$
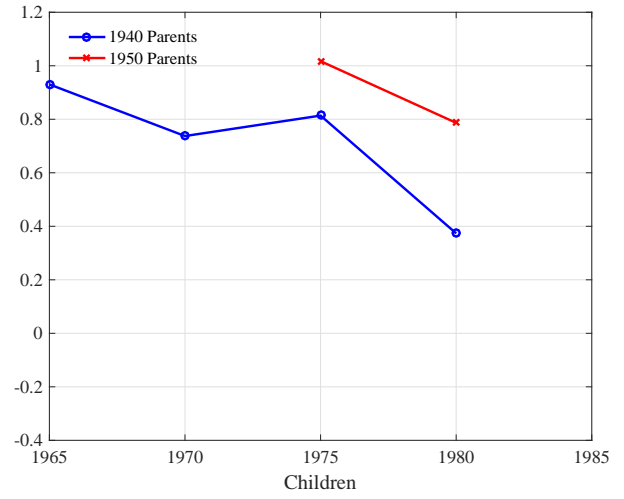
(b) $\alpha^*_{\delta,\psi}$
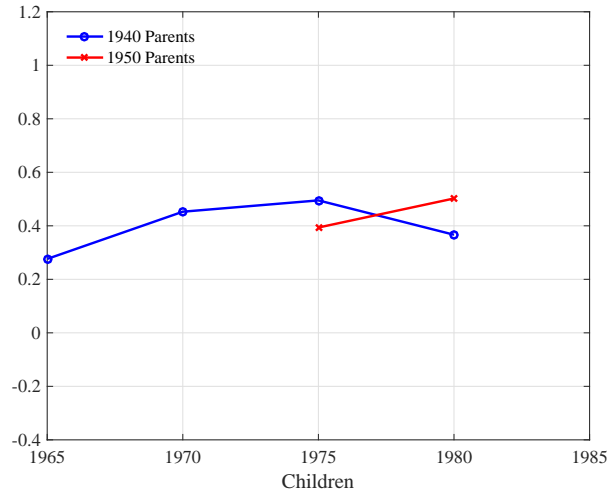
(c) $\alpha^*_{\psi,\delta}$
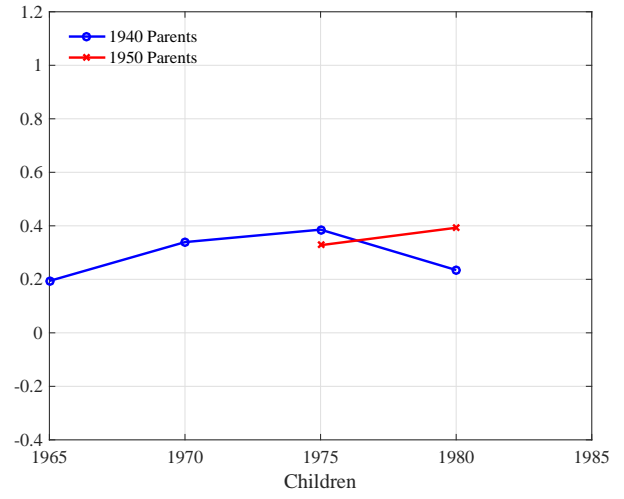
(d) $\alpha^*_{\delta,\delta}$

Figure C-7: Standardized Linear-Projection Coefficients ($\mathbf{A}^*$)

Figure C-8: Persistent Shock Variances (Var($\nu_{j,t}$))
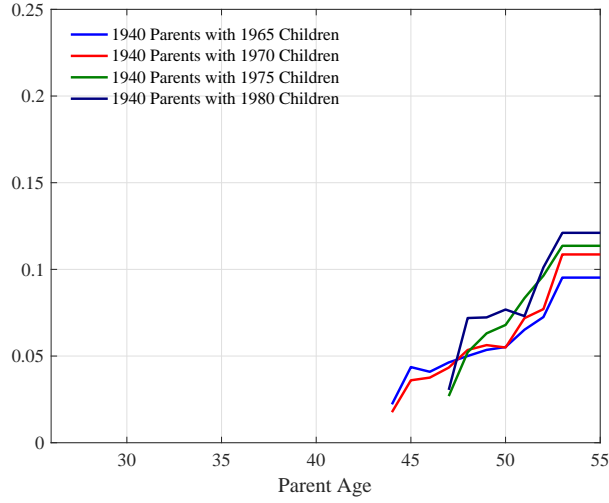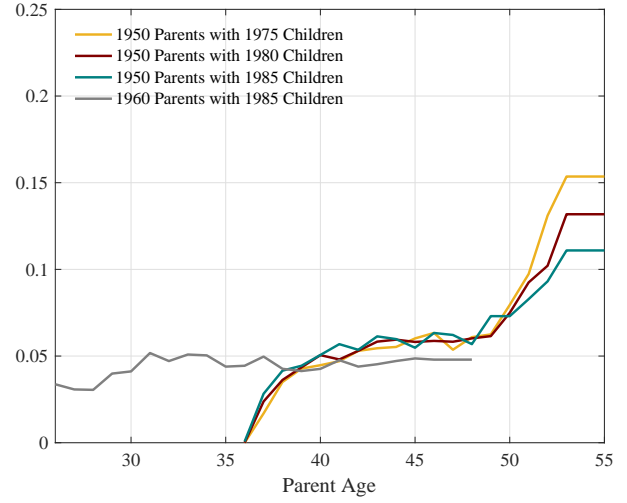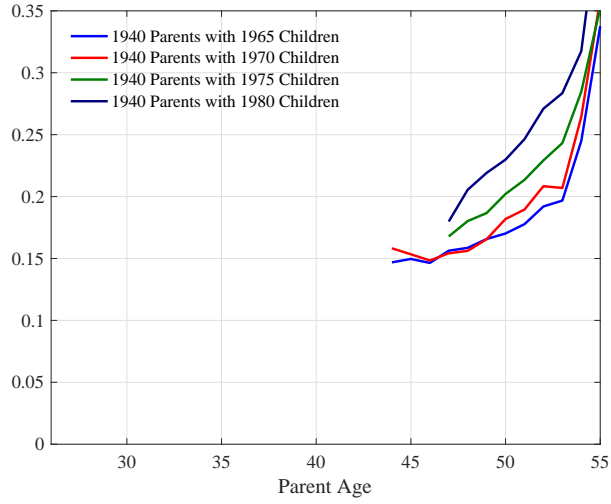
(a) 1940 Parents
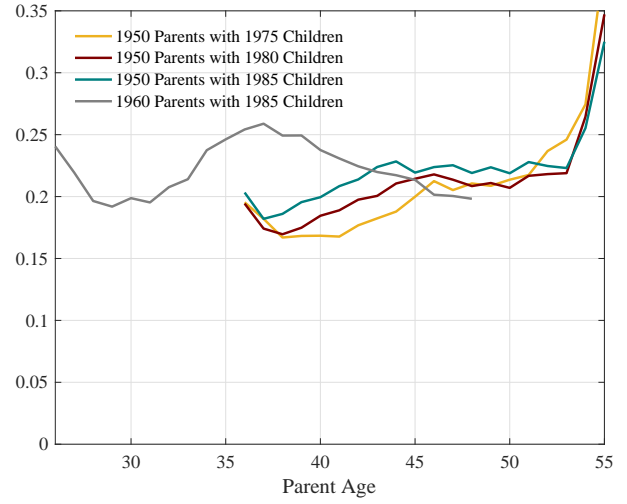
(b) 1950 and 1960 Parents
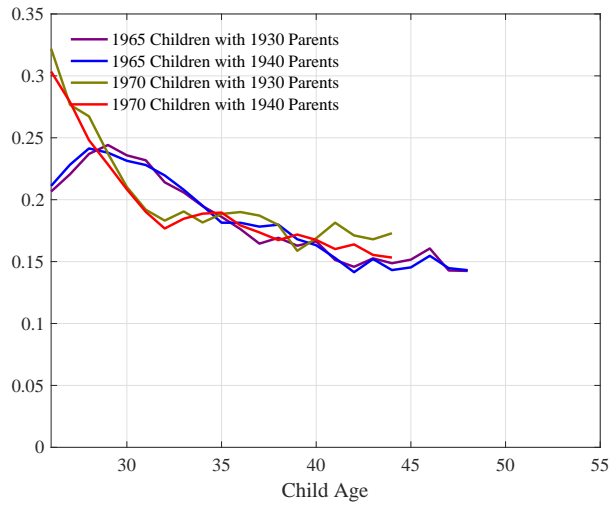
(c) 1965 and 1970 Children
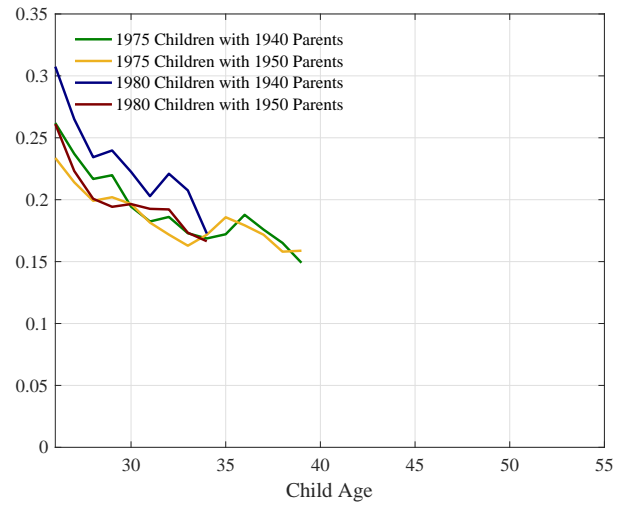
(d) 1975 and 1980 Children

Figure C-9: Transitory Shock Variances ($\text{Var}(\xi_{j,t})$)

D   Additional Results for National Analysis
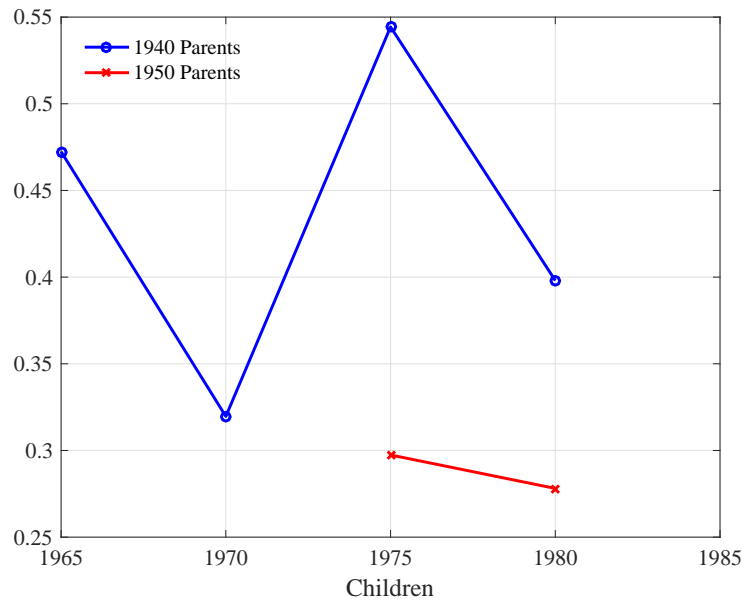
Figure D-10: $E\left[\mathrm{Var}(\hat{\theta}_{k,t}|\theta_{p,t'})\right]/\mathrm{Var}(\hat{\theta}_{k,t})$ (at Parent Age 50 and Child Age 30)

(a) 1940 Parents

(b) 1950 and 1960 Parents
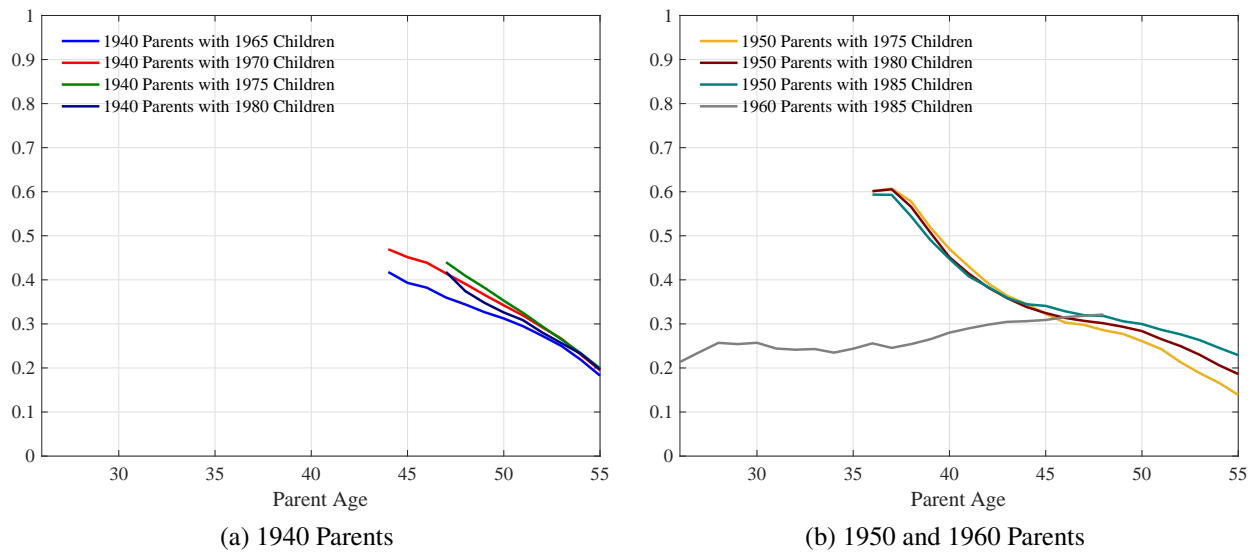
Figure D-11: $\mathrm{Var}(\theta_{p,t})/\mathrm{Var}(y_{p,t})$
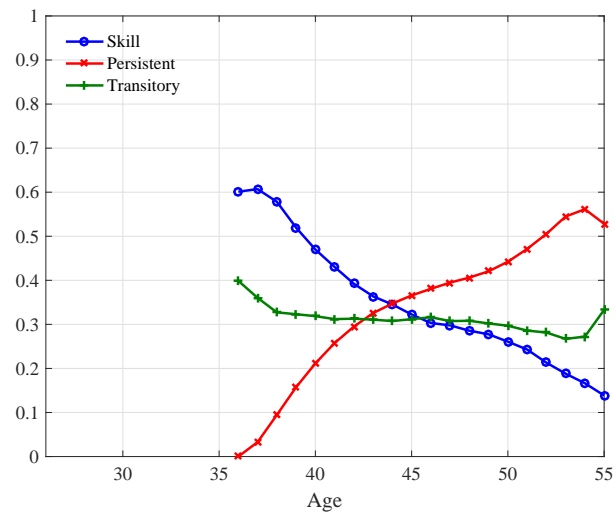
(a) 1975 Children, 1950 Parents
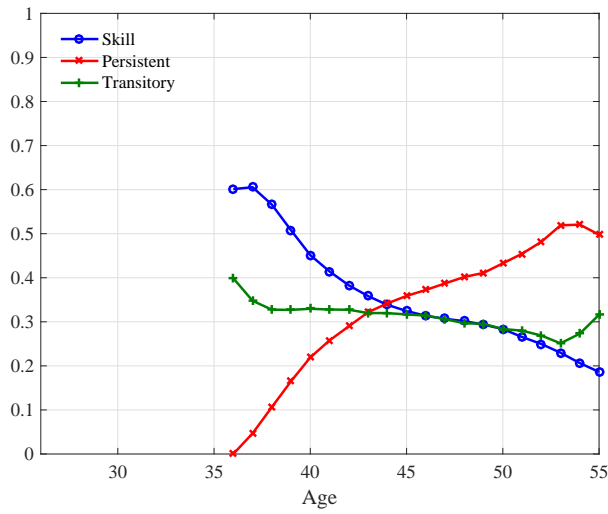
(b) 1980 Children, 1950 Parents
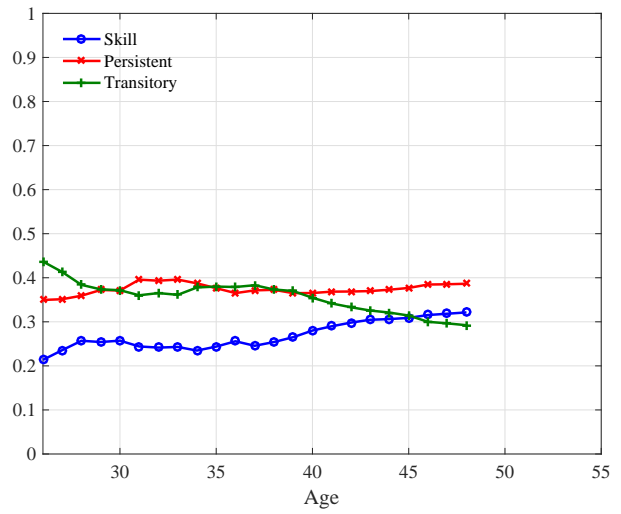
(c) 1985 Children, 1960 Parents

Figure D-12: Share of Earnings Variances for Parents
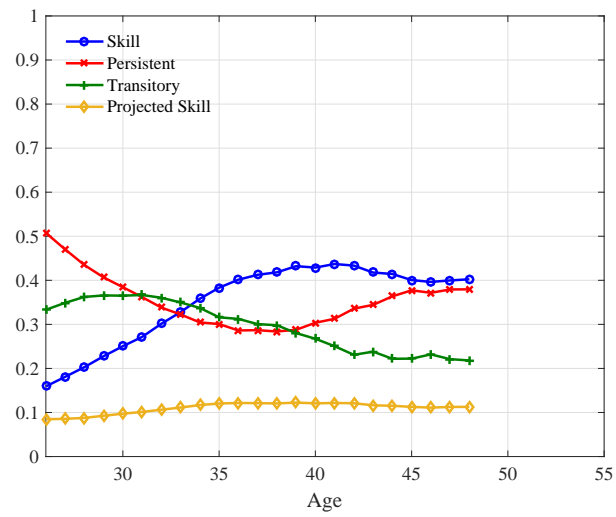
(a) 1965 Children, 1940 Parents
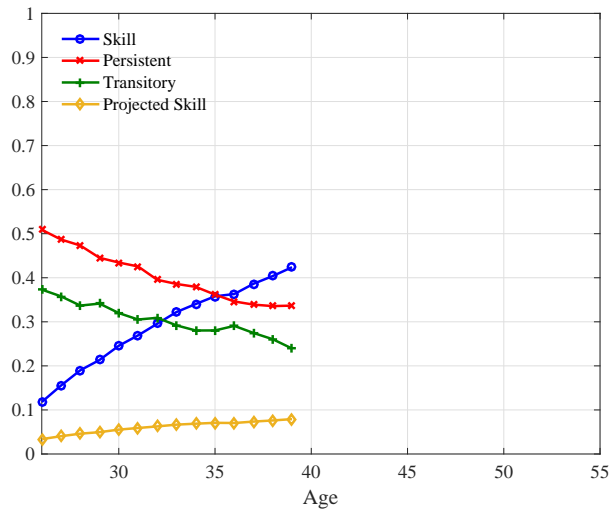
(b) 1975 Children, 1940 Parents
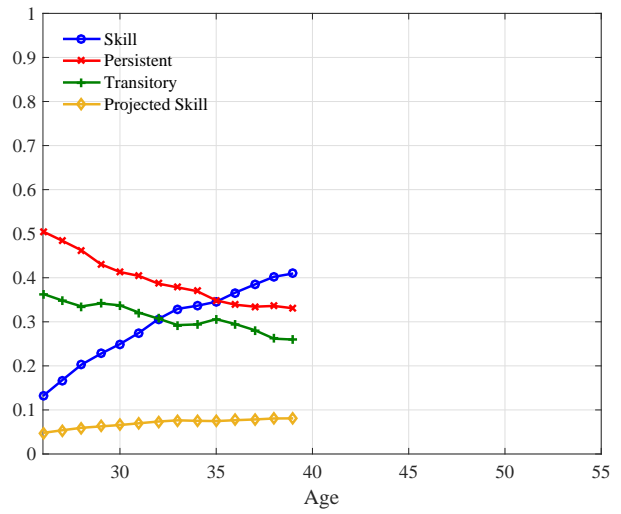
(c) 1975 Children, 1950 Parents

Figure D-13: Share of Earnings Variances for Children

E   Additional Results for Regional Analysis
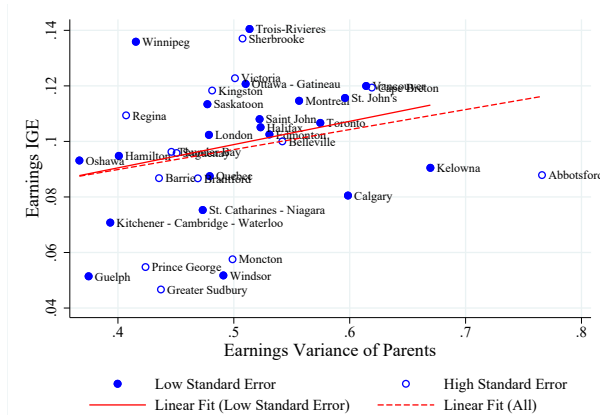
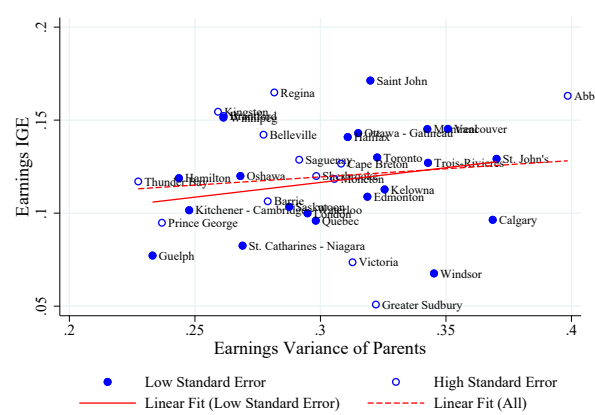(a) Earnings IGE vs. Earnings Variance
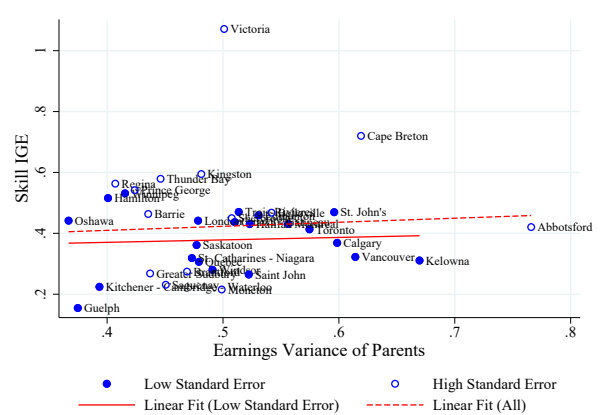
(b) Skill IGE vs. Earnings Variance
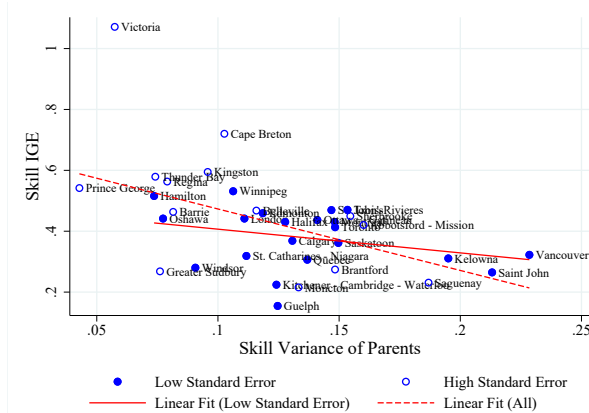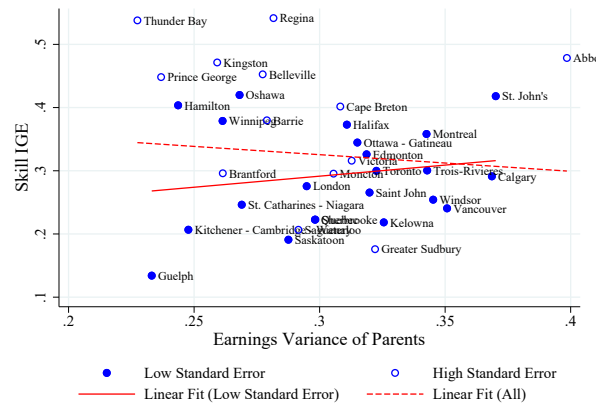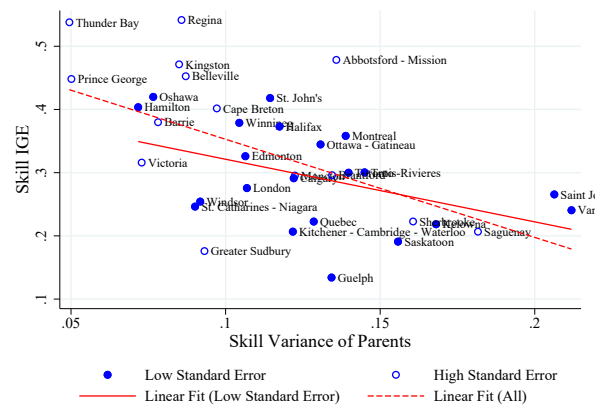
(c) Skill IGE vs. Skill Variance
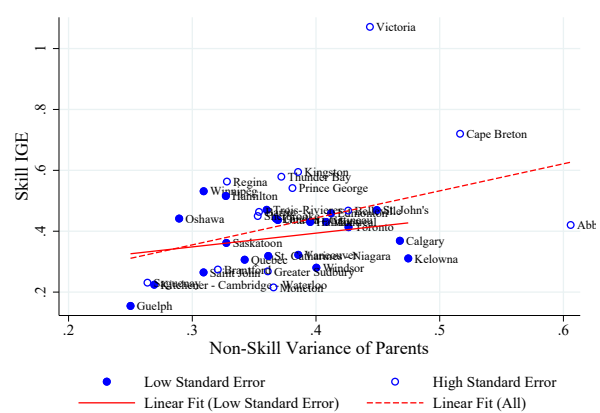
(d) Skill IGE vs. Non-Skill Variance

Figure E-14: Intergenerational Mobility and Inequality across Canadian Cities: Annual Measures
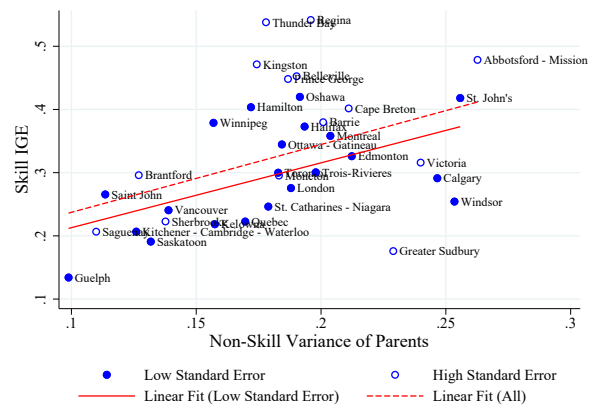
(a) Earnings IGE vs. Earnings Variance

(b) Skill IGE vs. Earnings Variance

(c) Skill IGE vs. Skill Variance

(d) Skill IGE vs. Non-Skill Variance

Figure E-15: Intergenerational Mobility and Inequality across Canadian Cities: 9-Year Averages