



HAL
open science

Intergenerational earnings mobility in Japan among sons and daughters: levels and trends *

Arnaud Lefranc, Fumiaki Ojima, Takashi Yoshida

► To cite this version:

Arnaud Lefranc, Fumiaki Ojima, Takashi Yoshida. Intergenerational earnings mobility in Japan among sons and daughters: levels and trends *. *Journal of Population Economics*, 2014, 27 (1), pp.91 - 134. 10.1007/s00148-012-0464-2 . hal-01648188

HAL Id: hal-01648188

<https://hal.science/hal-01648188>

Submitted on 24 Nov 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Intergenerational earnings mobility in Japan among sons and daughters : levels and trends *

Arnaud LEFRANC[†] Fumiaki OJIMA[‡] Takashi YOSHIDA[§]

December 6, 2012

Abstract

This paper estimates the extent of intergenerational income mobility in Japan among sons and daughters born between 1935 and 1975. Our estimates rely on a two-sample instrumental variables approach using representative data from the Japanese Social Stratification and Mobility (SSM) surveys, collected between 1965 and 2005. Father's income is predicted on the basis of a rich set of variables and we discuss changes in the Japanese earnings structure for cohorts born between the early 1900s and the 1960s. Our main results indicate that the intergenerational income elasticity (IGE) for both sons and daughters, in Japan lies around .35, which is an intermediate value, by international standards. We discuss the sensitivity of the IGE to using either personal or family income as the income variable for both fathers and children. We also examine changes across cohorts in the IGE. Results indicate that intergenerational mobility has been roughly stable over the last decades.

JEL Codes: D1, D3, J3

Keywords: Intergenerational mobility, income inequality, Japan, assortative mating, education.

*We thank Jaap Dronkers, Nobuo Kanomata, Tadashi Yagi and seminar participants at Keio university and Doshisha university for useful comments. Access to the SSM data was granted through the SSM 2005 Research Committee, whose support is gratefully acknowledged. This research received financial support from the French National Research Agency, under the grant TRANSINEQ (ANR-08-JCJC-0098-01) and from the Japanese Society for the Promotion of Science (Grant-in-Aid for Scientific Research #22330161).

[†]Corresponding author. Université de Cergy-Pontoise, THEMA, and IZA. Corresponding address : Université de Cergy-Pontoise, 33 boulevard du Port, THEMA, F-95011 Cergy-Pontoise. Email: arnaud.lefranc@u-cergy.fr. Tel : +33 1 34 25 60 62. Fax : +33 1 34 25 62 33.

[‡]Department of Sociology, Doshisha University. Shinmachi-dori Imadegawa Agaru, Kamigyō-ku, Kyoto, 602-8580 Japan.

[§]Faculty of Humanities and Social Sciences, Shizuoka University, 836 Ohya, Suruga-ku, Shizuoka, 422-8529 Japan.

1 Introduction

Over the last fifteen years, an abundant literature has analyzed the extent of intergenerational economic mobility and has revealed that, in many developed countries, a large fraction of economic inequality is transmitted from one generation to the next, within families (Solon 1999, Black & Devereux 2010). While available evidence has contributed to a rich description of the extent of intergenerational transmission, the mechanisms responsible for this transmission remain largely unstudied. In particular, although evidence seems to suggest that countries where income inequality is lower also exhibit more intergenerational mobility, as discussed for instance in Björklund & Jäntti (2009) and Lefranc, Pistoiesi & Trannoy (2008), the contribution of key ingredients of the mobility process, such as such labor market institutions, wage inequality and educational institutions is still, from an empirical perspective, unclear. To this end, the gathering of international evidence on the extent of intergenerational mobility in countries with different social and economic structure appears as an important step. Another important limitation of existing evidence, however, in this perspective, is that it has, until recently, mostly focused on western developed countries and much less is known about the extent of intergenerational mobility in other parts of the world, including Asian countries. The objective of this paper is to fill in this gap and to measure the extent of intergenerational income mobility in Japan.

Several characteristics of Japan make it an interesting case for the study of intergenerational earnings mobility, in particular from a cross-country comparative perspective. First, Japan is often seen as a fairly equal society characterized by compressed income differentials and limited poverty. Indeed, this image seems largely sustained empirically, at least compared to other countries and until the increase in inequality that occurred in the 1990s (Gottschalk & Smeeding 2000, Tachibanaki 2009).¹ The extent to which this high degree of income and, more generally, social equality translates into a high level of economic mobility is of course an important question, both for the comprehension of contemporary Japan and for the understanding of the intergenerational income mobility process. In this respect, opposing views can be found. On the one hand, occupational and educational success are often thought, at least in Japanese popular views, to depend little on family origin but rather on individual effort. This is summarized by the popular meritocratic belief : "Doryoku wa mukuwareru" (effort pays).²

¹Whether Japan still is, in the most recent period, an "equal" society seems highly debatable, as discussed in Tachibanaki (2009).

²The relevance of this view was challenged with great force in general-audience debates by the book of Toshiki Sato (Sato 2000) who made the (controversial) claim that the Japanese society had been closing down after the bubble-growth era of the 1990s.

On the other hand, Japan is characterized by a strongly differentiated and very competitive educational system, where success at entering the most prestigious universities, largely conditions future labor market prospects and henceforth motivates considerable financial investment by the parents. As a consequence, one would expect family background to strongly influence individual outcomes. Which view is the most relevant is the key question addressed in this paper.

Evidence on the extent of intergenerational economic mobility in Japan is currently rather limited, in comparison to the vast literature that has focused on social and educational mobility (e.g. Ishida 1993, Ojima 1998, Imada 2000, Kondo 2000). To our knowledge, only two papers have examined the extent of intergenerational income mobility in Japan.³ Lefranc, Ojima & Yoshida (2008) provided a first assessment. As the present paper, it relies on the SSM dataset, i.e. a large representative data set covering the second half of the twentieth century. But the analysis is limited to father-son pairs and the prediction of fathers income rests on a narrow set of individual characteristics. Ueda (2009) offers estimates of intergenerational income elasticities for male and female based on the Japanese Panel Survey of Consumers. However, this data set has several limitations that may influence the estimation of the intergenerational elasticity, as discussed below.

The main contribution of our paper is to offer a robust and in-depth analysis of intergenerational income mobility in Japan among sons and daughters. We measure mobility by the now standard intergenerational earnings elasticity (IGE), which can be obtained by regressing the log of individual annual earnings on the log of their father's earnings. Our analysis relies on the Social Stratification and Mobility (SSM) survey, a rich survey conducted between 1955 and 2005 that gathers information on individual income as well as family and social background. In the lack of direct observation of father's income, we use a two-sample instrumental variables approach as in e.g. Björklund & Jäntti (1997). Father's income is predicted on the basis of a rich set of variables including education, occupation and job characteristics. The SSM data allow to measure intergenerational mobility for a representative sample of children born between 1935 and 1975 and we also examine changes across cohorts in the IGE.

We also perform several robustness analysis. First, our data includes two distinct measures of income : the first one refers to the individual's own income; the second to his or her family income. Having both measures allows in particular to isolate the contribution of marital outcomes to the intergenerational mobility process and we examine the sensitivity of the IGE to

³A third paper is that of Yoshida (2008) but it is only available in Japanese.

the income measure used in the analysis. We also evaluate the robustness of our estimates of the IGE to various changes in the econometric specification, including changes in the first-step equation used to predict father's income and changes in the sample of children used to estimate the IGE.

Incidentally, the prediction of father's income requires us to examine changes in the Japanese earnings structure for a large set of cohorts born between the early 1900s and the 1960s, using income data for the years 1965 to 1995. As a result, an important by-product of this study is to document long-term trends in the Japanese earnings structure. Compared to existing studies, our analysis of trends in inequality in Japan offers several advantages. First, it covers a longer time span than what is usually studied in other papers. Second, it relies on microeconomic data unlike most studies that have, up until recently, largely used aggregated data. Third, our analysis covers all employment sectors, firm size and employment status unlike many studies that have focused on salaried workers in large firms from the manufacturing sector.

Lastly, we also provide two extensions of our basic estimates of the IGE. The first one examine possible changes across cohorts in the IGE. The second extension estimates the intergenerational correlation of earnings (IGC). One of the limitations of the intergenerational elasticity is indeed its sensitivity to changes in the extent of cross-sectional inequality over time, between the generation of the parents and that of the children. The intergenerational correlation (IGC) is, on contrary, not influenced by such changes and offers a useful complementary measure of the degree of intergenerational mobility.

The main result of our analysis is that intergenerational income elasticity in Japan is around .35, which is an intermediate figure in comparative perspective. This elasticity is rather similar for sons and daughters. It seems constant over time. Lastly, in the case of sons, their overall family income seems less related to parental income than their own individual incoming, suggesting that marital sorting dampens slightly the intergenerational persistence of inequality. The opposite holds true for daughters. The rest of the paper is organized as follows. Section 2 discusses the estimation procedure and section 3 presents the data used in the analysis. Section 4 presents the results of the first-step estimation and analyzes the main characteristics of the income structure in Japan and its evolution over time. Section 5 presents the main results for the estimation of the IGE for sons and daughters and section 6 discusses the sensitivity of these results to changes in the specification and sample selection rules. Section 7 examine time-trends and estimates the intergenerational correlation.

2 Estimation method

Most of the economic analysis of intergenerational mobility focuses on estimating the IGE in permanent (or long-term) earnings.⁴ This elasticity is given by the coefficient β in the following intergenerational regression model :

$$Y_i = \beta_0 + \beta X_i + \epsilon_i \quad (1)$$

where Y_i denotes the log of individual i 's long-term earnings X_i denotes his father's long-term earnings. As abundantly discussed in the literature, β should not be seen as a structural parameter measuring the causal effect of parental resources on child's earnings, but rather as a descriptive measure of the intergenerational association in earnings, capturing all possible channels of transmission.⁵

Direct estimation of equation 1 requires a considerable wealth of information. Not only does it call for a linked data set in which both father and child's earnings are observed but it furthermore requires one to observe a time-series of individual earnings in order to measure long-term earnings. Very few data sets satisfy this requisite.

Without such data, β can be estimated using a two-sample instrumental variables (TSIV) estimation, an approach originally derived in Angrist & Krueger (1995) and Arellano & Meghir (1992). This method was first applied to the estimation of the IGE by Björklund & Jäntti (1997). The basic principle behind TSIV estimation is to replace X_i in equation 1 by a prediction \hat{X}_i formed on the basis of some observable father's characteristics, Z_i .

The data requirements for TSIV estimation are significantly less stringent. The prediction is derived from a first-step equation which is estimated on a sample that is representative of the fathers' population, and in which we observe both earnings and the characteristics Z_i . Given the estimation of the first step, the data requirement for the estimation of β is to observe both child's income and father's characteristics.

TSIV has been extensively used for the estimation of the IGE and its properties are discussed in several papers including Solon (1999) and Nicoletti & Ermisch (2007). These properties depend on the choice of the instrument. If the instrument only affects child's earnings through its effect on father's earnings, the estimation of β is consistent. Indeed, in this case TSIV estimation

⁴An alternative approach is to estimate siblings correlation (e.g. Björklund, Eriksson, Jäntti, Raaum & Österbacka 2002).

⁵For general discussions and surveys of the causal mechanisms underlying the intergenerational transmission of inequality, see in particular Bowles & Gintis (2002) and Black & Devereux (2010). For recent analysis, see for instance Anger & Heineck (2010) and Liu & Zeng (2009).

offers the significant advantage of over-riding the *attenuation bias* that typically arises, because of classical measurement errors, when estimating equation 1 with long-term earnings replaced by current earnings (Solon 1992, Zimmerman 1992, Mazumder 2001). However, if the instrument has a direct effect on the child’s outcome, than the TSIV estimates are biased and the direction of the bias depends on the sign of the direct effect. For most of the instruments used in the present paper, the expectation is that the direct effect will be positive, hence resulting in an overestimation of the IGE. However, in practice, the papers that have estimated the extent of this bias generally conclude that the order of magnitude is rather small (see for instance Björklund & Jäntti 1997). In section 6, we address this issue and examine the sensitivity of our IGE estimates to the characteristics used to predict fathers’ income.

The recent literature has also pointed to another potential source of bias in the estimation of the IGE. This second bias, referred to as the *life-cycle bias* (Jenkins 1987, Grawe 2006, Haider & Solon 2006), arises when using current earnings instead of permanent earnings in the estimation of the IGE. In the presence of individual heterogeneity in earnings growth over the life-cycle, current earnings measure permanent earnings with error. Furthermore, the error is not of the classical type and is correlated with both true permanent earnings and age.⁶ If age-earnings profiles are steeper for high income individuals, current income differentials, measured at an early (respectively late) stage of the life-cycle, will underestimate (resp. overestimate) permanent income differentials. This introduces an asymmetric bias in the estimation of β . Using current earnings early (resp. late) in the life-cycle, as a proxy for *child’s* permanent earnings will lead to underestimate (resp. overestimate) β . Conversely, using current earnings early (resp. late) in the life-cycle, as a proxy for *father’s* permanent earnings will lead to overestimate (resp. underestimate) the IGE. To address this source of bias, we rely on the results of Haider & Solon (2006) who indicate that life-cycle bias is small if not nil around age 40.⁷ As a consequence, our measure of father’s income is predicted at age 40 (see below) and our sample of children is centered around age 40. We also perform several robustness in section 6.

In the end, the specification for the second-step equation comes as follows:

$$Y_{it} = \alpha_t + \beta \hat{X}_i + g(\text{age}_{it}) + e_{it} \quad (2)$$

⁶The classical measurement error case refer to the situation where measurement error is independent of the true value.

⁷Since their study is based on US data, one should of course question the relevance of their results in the Japanese context. In the absence of direct replication of their study using Japanese data, this question cannot be addressed directly.

where i and t are indices for individual and time. The α_t s denote time dummies and g is a fourth order polynomial function in age.⁸ \hat{X}_i is predicted father's earnings at age 40; the variable age is normalized to zero at age 40.

Let us now turn to the econometric model used in the first-step to predict fathers' income. Our objective is to predict father's income at the age of 40 on the basis of his education, occupation and job characteristics. In practice, two main issues have to be taken into consideration in the specification of this equation. First, some cohorts, in particular the oldest ones, are observed far away from their mid-career. Hence, to accurately predict earnings differentials at age 40 we need to account for heterogeneity in age-earnings profiles. This is done by introducing group-specific age-earnings profiles where groups are defined on the basis of their education and employment status.⁹ Second, since the first-step equation is estimated on a large range of cohorts, covering more than 60 years and there is no reason to expect the effect of individual characteristics on earnings to be constant over such a long period. Hence we allow the effect of individual characteristics to vary across cohorts.¹⁰

Given that extreme cohorts, old and young, are not observed over their full working career, it is not possible to be fully flexible when modeling the interaction of age, individual characteristics and birth cohorts in the wage equation. In particular, it is not possible to allow age profiles to vary with both individual characteristics and birth cohort. So we only allow age profiles to vary with individual characteristics. Furthermore, given the size of our sample, letting the effect of individual characteristics vary freely with birth cohorts would lead to imprecise estimates. As a result, we impose some parametric restrictions on the shape of cohort trends. We let the effect of individual characteristics vary by interacting these characteristics with a function of birth cohort. We consider three possible functional form for cohort trends : (i) a quadratic function of birth cohort; (ii) a piece-wise linear trend in birth cohorts that allows trends to differ between cohorts born before and after 1945; (iii) a step-wise function that allows the effect of individual characteristics to differ across three broad groups of cohorts : before 1930, 1931-1950, after 1950. We defer to the next section the discussion of the relevance of these functional forms.

⁸In principle, the use of polynomial function for age would allow to simultaneously include time and cohort dummies. Cohort dummies however turn out to be insignificant when added to this specification and their inclusion does not affect the results.

⁹A larger set of interaction terms could be introduced, such as occupation-age interactions. This is done in section 6. In practice, they turned out to be non-significant.

¹⁰In most specifications, the characteristics whose effects are allowed to vary across cohorts are age, education and self-employment status. We consider other variables in section 6.

In the end, the first-step model we estimate can be summarized by the following equation :

$$X_{ict} = \alpha_t + \phi(Z_{ic}, c) + \psi(\text{age}_{ict}, Z_{ic}) + e_{ict} \quad (3)$$

where i, c, t are indices for individual, cohort and time; Z_{ic} are the father's individual characteristics, ϕ captures the effect of these characteristics, which is allowed to vary with c ; ψ denotes the heterogenous age profile which is allowed to vary with individual characteristics and is normalized to be equal to zero when $\text{age}_{ict} = 40$.

Lastly, predicted father's earnings is given by :

$$\hat{X}_{ict} = \hat{\phi}(Z_{ic}, c) \quad (4)$$

where $\hat{\phi}$ is a estimate of the function ϕ .

3 Data

Our data come from the Social Stratification and Social Mobility (SSM) surveys. The SSM survey has been the primary data source for studies of social and educational mobility in Japan (Ishida 1993, Ojima 1998, Imada 2000, Ishida & Miwa 2008). The first wave of the survey was conducted in 1955 by the Japanese Sociological Society. Since then, similar surveys were conducted at intervals of ten years. The earliest waves (1955, 1965 and 1975) focused only on males. A female sample was collected since the 1985 survey. The questionnaire of the last wave of the survey (2005) has also been used for similar surveys in Korea and Taiwan.

The SSM samples are designed to provide a national representative sample of the population between 20 and 70 years old. Across the different waves, the size of the male sample varies between two and three thousands individuals. The questionnaire focuses on the description of social status, educational attainment, social origin, class identification and the perception of inequality. The most important variables in our analysis are income, which is the main variable of interest, and educational and occupational attainment, which serve to predict father's income in the first-step equation. Respondents to the SSM survey are asked to report their income, education, occupation and job characteristics as well as the education, occupation and job characteristics of their father. As often the case, father's information is reported *ex post* by the survey respondent and refer to father's main occupation.

All SSM waves record two distinct measures of income. The first one is individual own

income. The second one is family total pre-fisc income. In both cases, the variable measures annual primary income, in the year preceding the survey, before any tax or transfer and includes both labor and asset income. Furthermore, income information is available for both salaried and self-employed workers. For most individuals of working age and who actually work, the primary component of pre-fisc income is labor earnings. Family income is the total of each family member's individual income in the respondent's household. Income is available in all waves of the survey in bracketed form, except for 1965 where income is coded continuously. The bounds and number of brackets vary across waves, between 17 and 30 main brackets. Higher income are not top-coded though and a response beyond the top category was coded continuously. In the regressions, we assign the mid-value of the bracket and use standard linear regression techniques.¹¹

One of the appealing features of the SSM data is to provide individual data on income for the entire Japanese labor force, including all industries and employment status, over a long time period. This represents a significant advantage of SSM, compared to alternative data sources. In studies of the Japanese wage structure and trends in inequality, the most commonly used data set has been the Basic Survey of Wage Structure (BSWS) collected by the Japanese Ministry of Health, Labor and Welfare.¹² However this data suffers several limitations. In particular, it is restricted to salaried workers employed in firms with more than 10 employees. Hence, it excludes in particular self-employed workers, as well as workers in irregular employment forms. Second, unlike the SSM surveys, the sampling design of the BSWS data has changed over time which makes its use problematic for studying long term trends.

The education classification used in the different waves of the survey varies across waves and cohort, reflecting the changes in the Japanese educational system that occurred over the last century. For older cohorts, the classification distinguishes between five educational levels: elementary school (6 years of formal schooling), upper elementary (8 years), middle school (11 years), college (14 years) and university (17 years). For more recent cohorts, the five educational levels are: junior high school (9 years), high School (12 years), junior college (14 years), university (16 years) and graduate school (18 years). Given sample size and to assure cross-year consistency of the education classification, we used a reduced classification that distinguishes between three educational levels: lower secondary education (or lower), upper secondary education and tertiary education. This corresponds, for instance, to the classification used in Kondo

¹¹Lefranc, Ojima & Yoshida (2008) uses interval regression to deal with the bracketed form of income and show that the impact on the estimated IGE is negligible.

¹²Among others, Katz & Revenga (1989) and Kambayashi, Kawaguchi & Yokoyama (2008) analyze the Japanese earnings structure and its evolution using BSWS data.

(2000).

Social status is coded using the Erikson-Goldthorpe-Portocarero (EGP) classification. We use a variant of the classification in 9 groups, since a change of the questionnaire about the number of employees between 1985 and 1995 makes it impossible to distinguish between IVa (Small employer) and IVb (Independent) (Kanomata, Tanabe & Takenoshita 2008). In the end, we use the following eight-categories¹³ : I- Higher-grade professionals, administrators, and officials, managers in large industrial establishments, large proprietors; II - Lower-grade professionals, administrators, and officials; higher-grade technicians, supervisors of non-manual employees; III- Routine non-manual employees, higher grade (administration and commerce) and lower grade (sales and services); IV- small proprietors and self-employed workers ; V- Lower-grade technicians supervisors of manual workers; VI- Skilled manual workers; VII- Semi- and unskilled manual workers; VIII - Farmers and farm workers. In addition to social status, SSM also includes additional characteristics of occupation and job position. We use firm size and an indicator of self-employment status.

In the main samples used in this paper to estimate both the first- and the second-step equations, we exclude those without positive earnings in the year preceding the survey. The sample used in the estimation of the first-step equation draws on five available survey waves, 1965, 1975, 1985 and 2005. The reason for excluding 1955 is that income information is missing in that wave for farmers who accounts for a large share of the employment in older cohorts. The sample is restricted to individuals aged 30 to 59 years old. The main reason for excluding individuals older than age 60, is that many people in this sample start retiring and living on pension from this age in Japan, a problem also encountered in the study of Ueda (2009).

Second-step estimations are based on the three most recent waves (1985, 1995 and 2005).¹⁴ For reasons already discussed, the children sample is restricted to individuals aged between 30 and 50 years old, i.e. close to the middle of their working career. For each individual in the second-step sample, we form a prediction of his father's income using estimates of the first-step equation. The prediction is based on reported father's education, EGP classifications, residential area and other occupational information. In most cases, individual in the second-step sample report their father's birth year. In this case, we use the relevant age-specific returns to education to predict father's income. When information on father's birth year is not available, we predict fathers income on the basis of the observed distribution of father's birth cohort, conditional on

¹³Our analysis relies on the coding developed by Shirakawa in the SSM research group.

¹⁴The reason for not using earlier waves of the SSM for the second-step is that women are only interviewed starting in 1985. Furthermore, some information on family background is missing or not recoded homogenously before 1985.

child's birth cohort. Mean values in the samples are presented in table 1.

!! Table 1 about here !!

Lastly, when predicting father's income for early cohorts, we are limited by the scope of our first-step sample. The first-step equation is only estimated on individuals born after 1905. However, about half of the sample born in 1935 have a father born before 1905. To predict the income of fathers born before 1905, we make the assumption that the earnings structure among cohorts born before that date is similar to the earnings structure of the 1905 cohort. As discussed in the appendix, our results are insensitive to this assumption.

4 First-step estimates and long-term trends in income inequality in Japan

Main effects First-step estimates are presented in tables 2 and 3 for personal and family income respectively. These estimates are based on the sample representative of the fathers cohorts and include only male. The first salient result that emerge from this table is the relatively limited extent of earnings differentials across the various groups. For instance, the gap between the two extremes of the social ladder, higher-grade professionals and farmers, is less than .7 log points. Similarly, the difference between the low- and high-education groups is less than .2 log points. Of course, the *ceteris paribus* analysis of the coefficients is largely artificial given the collinearity of the different dimensions but even if we cumulate the above mentioned earnings gaps, they lead to modest earnings differentials, in comparison to what can be found in other countries. An interesting comparison can be established with the results of Lefranc & Trannoy (2005) who use similar approach and classifications. In the case of France, which is often thought to occupy an intermediate position among developed countries in terms of earnings inequality, they find a gap between the top and bottom social groups of .9 to 1 log points and of .3 to .4 between top and bottom education groups.

!! Tables 2 and 3 about here !!

The finding of a relatively small earnings differentials by occupation and education confirms the previous results obtained in studies based on a narrower period and using the BSWs data which only covers salaried workers outside small firms (Koike 1988, Katz & Revenga 1989). In

particular Koike (1988) also documents that wage differentials by occupation and educational groups are narrower in Japan than in EC countries and the US. Ishida (1990, 1993) using data for 1975, also finds that the college premium in Japan was much smaller than in the US, even before the large rise in US inequality.

In contrast to occupation and education, results in tables 2 and 3 indicate that employment characteristics lead to non-negligible earnings differentials in Japan. First, firm-size has a strong impact on individual earnings : the benefit of being employed in a large firm is larger than the income gap between the high- and low-education groups. This confirms previous findings of large firm-size effects in Japan (Oi 1990, Tachibanaki 2009).¹⁵ On the contrary, the effect of self-employment status on earnings has been seldom analyzed, largely because self-employed workers fall outside the scope of commonly used data sets such as the BSWs. Our results indicate that self-employment also positively affects income and has a significant effect.

Tables 2 and 3 also report earnings differentials by residential area. The reference category are the 6 largest Japanese cities¹⁶. Earnings differentials by residential are sizable, especially given the fact that we already control for education, occupation and employment characteristics.

The main effect of the socio-demographic and employment characteristics tends to be higher in the case of personal income than in the case of family income. This result is indeed rather intuitive since we should expect other family members' income to be less correlated to the individual's characteristics than his own income. The only exception is self-employment that has a stronger impact on family than on personal income. One possible explanation is that the wives of self-employed are often also employed as family workers in family-owned businesses, an activity from which they would derive an income higher than the average female income.

The second part of tables 2 and 3 report the estimated age-earnings profiles by level of education and self-employment status. In the specification of the first-step equation, we only included a quadratic function of age since higher-order terms were not significant. We also present the age-earnings profile in figure 2, panel A. As expected, the slope of age-earnings profiles increases markedly with education. At the beginning of individual careers, the income effect of 10 years of individual experience amounts to about .3 log points for individuals with higher education against about .2 for individuals with secondary education or less. This is consistent with the findings of Shimada (1981) and Koike (1988). On the contrary, the income-

¹⁵Since our estimation does not control for job seniority, one may argue that firm-size effects partly reflect differences in seniority by firm-size. However, the controversy on the extent of returns to job seniority has also suggested that seniority effects might themselves partly arise from firm fixed-effects (see in particular Topel 1991, Altonji & Williams 2005). (Hashimoto & Raisian 1985) discuss seniority distribution and its effect on earnings in Japan.

¹⁶Tokyo, Yokohama, Osaka, Nagoya, Kyoto, Kobe.

age profile of self-employed individuals appears, other things equal, flatter than that of other groups. This confirms the findings of Genda & Kambayashi (2002) using narrower data, from the 1989 and 1994 National Survey on Family Income and Expenditure.

!! Figure 2 about here !!

Cohort trends Tables 2 and 3 report the estimated changes across cohort in the effect of education, self-employment status and type of area of residence. Figure 2 reports these trends in the case of respondent's own income. As already discussed, we consider three possible specifications for cohort trends. In column 1 of the tables and panel B of the figure, trends are captured by a quadratic function of birth cohort. In column 2 and panel C, we assume piece-wise linear cohort trends, where we allow trends to change among cohorts born after 1945. Lastly, in column 3 and panel D, the trend is captured by a step-wise function that allows the effect of individual characteristics to differ across three broad groups of cohorts : before 1930, 1931-1950, after 1950.

For all specifications, trends across cohorts indicate a compression of earnings differentials in the long-run. Let us first focus on income differentials by level of education. The income gap between rural and urban area falls markedly as does the advantage of the larger cities over other rural areas. The earnings gap associated with being self-employed, which is positive among older cohorts, stays approximately constant among cohorts born before 1940. It falls with birth cohort afterwards, to reach a value close to zero. The changes in the youngest cohorts is consistent with the finding in Genda & Kambayashi (2002) who report a fall in the returns to self-employment in the 1990s, controlling for worker's age.

According to the quadratic trend specification, the income gap between the bottom education group and the top two decreases from a high value of .5 log points for cohorts born at the beginning of the twentieth century to a low value of about .2 for cohorts born after World War II. According to the step-wise specification, the gap falls from around .27 log points to about .22 log points between the extreme cohorts. All specifications indicate that this reduction of the education-income gap is not monotonic over the period. It occurs for the most part between cohorts born in the early century and cohorts born around WWII. In contrast, earning gaps are rather constant across cohorts born in the 1945-1965 interval. This time pattern roughly coincides with the pace of educational expansion that occurred in Japan throughout the twentieth century. As shown in figure 1, educational attainment rose markedly over this period, with the bulk of the educational expansion occurring between the early century cohorts

and the cohorts born in the late 1940s. At the end of our period, the downward trend in earnings differentials seems inverted and we observe a slight increase in earnings inequality. This small rise in inequality is however much smaller than the one documented in Tachibanaki (2009). It may be due in particular to the fact that our first-step equation is estimated on a sample of males aged between 30 and 59 and born before 1975, who were less affected by the rise in earnings inequality.

!! Figure 1 about here !!

Evidence on long term changes in the Japanese earnings structure is relatively sparse and not always available in English but is largely consistent with the evidence reported here. In its 1996 White Paper, the Economic Planning Agency estimated the returns to college for three birth cohorts : 1935, 1950 and 1965. The returns to college fall by 25% between the 1935 and 1950 cohorts and later rise by 10% between the 1950 baby-boomers and the 1965 cohort. Other studies have examined changes over time in the returns to education. Most studies conclude to a fall in the returns to education, at least up until the 1980s. For instance Yano (1982), using BSWS, finds evidence of a fall in the return to college between 1967 and 1980. There is less consensus regarding trends at work after 1980. Kambayashi et al. (2008) provide evidence of a fall in the returns to education between 1989 and 2003. However, this fall mostly concerns the gap between junior high-school graduates and more educated workers. On the other hand the gap between high-school graduates and college graduates seems to be relatively flat. Genda (1994) examines trends in the returns to education between 1965 and 1990 using BSWS data for salaried workers outside small firms. For all cohorts grouped together, his results indicate an uninterrupted fall in the returns to education, over the period. However, different patterns emerge by age group. For older workers aged 40 to 50, the returns to education are also continuously falling. On the contrary, in the age group 20-30, there seems to be a slow rise in the returns to education after 1980. In other terms, the returns to education seem to be on the rise, at least in the early stage of the career, for workers born in the 1950s and after. This is precisely the trend inversion indicated by our estimates. It is also interesting to note that all three specifications are consistent on that point.¹⁷

While all three specifications indicate the same overall pattern, they differ in terms of the extent of the changes occurring. In particular, the quadratic and piece-wise linear trends specification suggest a larger fall in the returns to education than what is showed by the step-wise

¹⁷In the case of the piece-wise linear and step-wise functions, the inflexion is of course constrained a priori by the parametric hypothesis but this is not the case for the quadratic function.

function. In order to accurately predict father’s income over a broad range of cohorts, it is important to know which specification is the most accurate. One first indication is provided by the various goodness of fit statistics (R-squared, adjusted R-squared, AIC, BIC). Although the differences are small, all fit measures indicate that the step-wise linear function better adjusts the data. Beyond these internal fit criteria, external evidence tend to support the quadratic or piece-wise linear trends assumptions. One of the assumptions of the step-wise function is that the returns to education and other characteristics are constant within each broad group of cohorts. However, existing studies suggest that this does not hold, especially in the case of older cohorts. The previously cited evidence from Genda (1994) imply that the college wage premium, measured at age 40 to 50, has fallen by about 25% between the cohorts born around 1920 and those born around 1930. Similarly, Yano (1982) indicates a continuous fall in the return to education between the 1910 cohort and the 1930 cohort. Our quadratic and piece-wise linear trends are consistent with these pieces of evidence. On the contrary, the step-wise trends specification is missing part of the long-term changes in the earnings structure.

5 Main results

Table 4 provides estimates of the IGE, based on equation 1 for sons and daughters. Coefficients in columns 1 to 3 of table 4 give the estimated value of the IGEs for different measures of the child and father’s income. In column 1 the child’s own income is regressed on the (predicted) father’s income; in column 2, it is regressed on the father’s family income; in column 3 the child’s family income is regressed on the father’s family income. For each combination of income measures, we present three different IGEs corresponding to the three specifications of the first-step equation discussed in the previous section.

5.1 Results for sons

Our main estimates of the IGE for sons are given in panel A of table 4. Estimates range between .28 and .42, with a mean value of .34. The estimated IGE varies with both the measure of income and the specification of the first-step equation.

!! Table 4 about here !!

Let us first discuss the sensitivity of IGE estimates to the variable used to measure income in both the fathers’ and children’s generations. Most values reported in international assessments

of the intergenerational elasticity are estimated using personal labor income. Hence international estimations are most closely comparable to the coefficients reported in column 1, that are based on reports of both father's and son's personal income. This coefficient lie between .32 and .36. The major discrepancy in the measurement of income, between the data used in column 1 and most international estimates is the inclusion of asset income in our measure of personal income. Strictly speaking, estimates in column 1 represent an average of the intergenerational elasticity for labor earnings and for asset income, weighted by the contribution of both sources to personal income. Previous estimates indicate that the intergenerational elasticity for wealth is significantly larger than for labor earnings (e.g. Mulligan 1997). So for the purpose of international comparisons of the extent of the mobility in the personal labor earnings, the figure of .33, in column 1, should be seen as an upper bound estimate of the elasticity in Japan. Of course, the order of over-estimation is probably rather small, given the share of asset income in overall income.

Compared to column 1, column 2 regresses sons' personal income on the overall family income of their parents. Measuring parental economic status by means of the total family income, instead of the sole father's income leads to a slightly higher IGE in the range .36-.42, i.e .04 to .05 higher than the coefficients in column 1. This suggests that total family income may better capture the broad nexus of factors that shape individual success. This confirms results already obtained in several other papers (e.g. Lucas & Kerr forthcoming). In fact, in Japan, the gap between the two estimates appears smaller than what has been found in other countries, for instance the United States. For this country, Solon (1992) indicates a difference in the estimated IGE using both measures of .09, with the IGEs being respectively .39 and .48 when using father's earnings or family income.¹⁸

Lastly column 3 provides estimates of the IGE based on the regression of sons' family income on their parents' family income. Compared with estimated in column 2, since family income aggregates the income of both spouses, estimates based on this variable will reflect both the influence of assortative mating on mobility, and the direct intergenerational transmission between parents and children. More precisely, as discussed in Chadwick & Solon (2002), the estimated IGE in column 3 is the weighted average of (i) the intergenerational elasticity of son's own income to his parents' family income and (ii) the elasticity of his spouse's personal income to the son's parents' income, where the weights are given by the share of each of the

¹⁸Other evidence on the incidence of the measure of family economic status in the US can be found in Altonji & Dunn (1991) and Mulligan (1997). Our findings of a smaller gap between the IGEs find using both variables may be partly attributable to the fact that our measure of father's income, contrary to earnings measures used in the US, already incorporates the asset earnings of the father.

two spouses' income in the total family income. Using family income (column 3) rather than personal income (column 2) as a dependant variable leads to a fall in the IGE by about .09, with estimates in the range .29-.32. This suggests a relatively low elasticity between the wife's own income and her parents-in-law family resources. Ermisch, Francesconi & Siedler (2006) report similar results for the UK.

Finally, it is worth discussing the sensitivity of the estimated IGEs with respect to the specification of cohort trends in the first-step equation. Estimates tend to be lower under the assumption of a continuous trends (quadratic trends (QT) and piece-wise linear (PW) specifications) than under the assumption of a stepwise (SW) change in the earnings structure. The gap between the SW and the other two specifications ranges between .035 and .055, ie. 10 to 15% of the estimated IGE. This mostly arises from differences in the prediction of father's income in older cohorts. Observed inequality among children is, of course, unaffected by the choice of the first-step specification. However, predicted inequality among older fathers is lower under the piece-wise specification than under the QT and PW specifications. Since the IGE roughly measures what share of father's inequality is transmitted to their children, a lower predicted inequality among fathers mechanically translates, for a given degree of inequality among children, into a higher estimated IGE. As already discussed we believe that the step-wise specification does not accurately capture trends among fathers, and we believe that specifications QT and SW better capture trends in the Japanese earnings structure and henceforth provide a better estimate of the IGE.

5.2 Results for daughters

Estimates of the IGE for daughters are given in panels B and C of table 4. OLS estimates are presented in panel B for the subsample of individuals reporting positive personal income. This the case of only 60% of our daughters sample, because of non-participation to the labor force. Since non-participation is likely to be endogenous, we report selectivity-corrected estimates in panel C.

The OLS estimates of the IGE for daughter's personal income, reported in colums 1 and 2 of panel B, appear particularly low. With one exception (SW specification, father's family income) all estimates lie in in the interval .21-.26, which is low by any standard. On the contrary, the OLS estimates obtained when regressing daughters' family income on their parents' predicted family income is significantly higher, in the interval .37-.39.

Of course, the large gap between the two sets of estimates might partially arise from the

existence of sample selection in the equation for personal income, since non-response is much more frequent for this variable. This is confirmed by the estimates in panel C, obtained using Heckman's sample selection model. In the selection equation, an indicator for the report of income is regressed on marital status, indicators of the number of children, spouse's income and education when married, as well as age profile and year indicators. Overall, IGE estimates obtained in the case of personal income are markedly higher than those obtained using OLS. These results indicate negative sample selection : women with the greatest income potential are less likely to participate to the labor market and thus report their own income, which leads OLS to underestimate the extent of the IGE. Once sample selection is taken into consideration, the IGE rises by about .15.

In the end, this suggests a value of about .35 when regressing daughter's own income on fathers' own income, predicted under the assumption of quadratic or piece-wise linear trends. As in the case of sons, the IGE appears slightly higher when regressing child's income on total parental income rather than father's personal income : about .4, again, under the QT or PW specification of the first-step equation. However, it should be noted that once sample selection is accounted for, the IGE for daughters appears slightly higher (although not significantly so) than the IGE for sons : this indicates a relatively high degree of correlation between daughter's potential income and the income of their parents.

Lastly, it is worth noting that the IGE for daughters is less sensitive than for sons to the use of family income as a measure on child's outcomes. In the light of the model developed by Chadwick & Solon (2002), this first suggests a rather high correlation between daughter's parents' income and her husband's own income. Compared to what we found for sons, this points to a gender asymmetry in the elasticity of own income to parents-in-law's income. This asymmetry could largely be explained by a more compressed earnings distribution for daughter's in Japan, together with gender neutral correlation between the characteristics of spouses and that of their in-laws.

5.3 Comparison with previous estimates

Overall, our results point to a value for the IGE in Japan slightly higher than .3 for sons and slightly lower than .4 for daughters. More precisely, for sons, the average IGE across all estimated specifications is .34 (.33 when ignoring coefficients based the step-wise trends specification of the first-step equation) and the average IGE for daughters is .39 (.37 without the step-wise trends specification). This represents an intermediate degree of intergenerational

mobility for sons, compared to other developed countries. Based on available estimates, Japan turns out to be more mobile than countries such as the United-States, France and Italy. It is also less mobile than most Scandinavian countries.

The results found here also differ slightly from previous results obtained for Japan, in Lefranc, Ojima & Yoshida (2008) and Ueda (2009). First, the IGE reported in table 4 are slightly higher than those reported for sons in our previous study (Lefranc, Ojima & Yoshida 2008). This study suggested a value of the IGE in the interval 0.22-0.31. However, this estimate was based on a less detailed first-step equation. Furthermore, self-employed workers were excluded from the main sample used in the analysis.

Our results are also differ from those of Ueda (2009). Based on her estimates, she concludes that : “*The estimated intergenerational elasticity using predicted parental income is 0.41-0.46 for married sons, 0.30-0.38 for married daughters, and marginally less than 0.30 for single daughters*”. Her estimates are based on the *Japanese Panel Survey of Consumers* for years 1994 to 2004. One major advantage of this data set over the SSM surveys used is that it directly asks children to report their father’s earnings and income. However, for most individuals in the sample, this information is collected in bracketed form using only 7 intervals.¹⁹ Furthermore, only one observation of parental income is available, which leads Ueda to use predicted father’s earnings in her analysis, in order to avoid attenuation bias, which is rather similar to our approach. Yet, several differences in the sample selection and prediction procedures are likely to explain the gap between her favorite estimates and those of the present paper. First, our estimates are based on a much broader range of cohorts. Second, income is observed at the same time for children and parents. Hence children are observed fairly early in their life-cycle and parents are observed fairly late. Ueda accounts for possible life-cycle biases by controlling for age effects but this strategy is only valid if the earnings structure has remained constant over time. On the contrary, section 4 indicates that earnings differentials have diminished over time. Third, in the case of sons, the results of Ueda only concern married individuals, since single men, who represent about 30% of the relevant population, are excluded from her data.²⁰ If there is selection into the marriage market this will bias her estimates, presumably downward though. At the end of the day, we believe that our estimates are based on a broader and more representative sample and that our estimation procedure better takes into account the changes in the earnings structure in the fathers’ generations. Lastly, although Ueda’s conclusions differ from those of the present paper, it is worth noting that our estimates are still consistent with

¹⁹Earnings are only measured in continuous form for single daughters still living with their parents.

²⁰According to the 2005 census, 30% of men between 30 year old and 59 year old and 25% of women are single.

her broader set of estimates.

6 Robustness analysis

Not being able to observe the long-term earnings of both fathers and their children, leads us to impose several restrictions on our empirical model. The first ones pertain to the empirical model used to predict fathers earnings. The second pertain to the way we deal with life-cycle effects in order to derive estimates of the long-term earnings elasticity from point in time measures of income. In this section, we evaluate the sensitivity of our estimates to the first-step specification and to the life-cycle effects.

6.1 Sensitivity to first-step specification

In the previous section, we have already discussed the sensitivity of our IGE estimates with respect to the specification of cohort trends in the earnings structure. We now examine the sensitivity of our estimates to the set of variables used to predict father's income. When presenting our econometric model, we have already mentioned that TSIV estimates by be inconsistent if the variables used as instruments for predicting father's income also have a direct effect on children's earnings. Most studies tend to conclude that this bias is relatively small. Yet, since some instruments might have a more pronounced direct effect on child's outcomes than others, it is worth investigating how IGE estimates vary when using a broader or narrower set of instruments.

Results are presented in table 5 and 6. Each table provides results for five possible sets of explanatory variables in the first-step equation and consider the three possible parametric specifications of cohort trends (QT, PW, SW). Specification A corresponds to the reference specification of the first-step equation whose results were presented in section 4 and used in the IGE estimates of the previous section. It includes the following father's variables : education, social class, self-employment status, firm size, heterogenous age-effects, and interactions between birth cohort and education, self-employment status and residential area.

!! Tables 5 and 6 about here !!

Specification B allows for more flexibility in cohort trends and age profiles by allowing for cohort trends in the earnings differentials by EGP social class, as well as EGP-class-specific age profiles. For both sons and daughters, this specification leads to slightly lower IGEs. The

difference with the base specification is rather small though and is around .01 to .02. One should also note that these differences, as well as all the differences discussed in this section are not statistically significant. This is an important caveat that should be kept in mind when interpreting the changes in coefficients.

Specification C builds on specification B but removes education main effects as well as interactions between education and cohort or age. Comparing specifications B and C indicates that the inclusion of education and interaction terms as instruments tends to dampen the IGE slightly. Again, the effect stays relatively small. This result differs from the result usually obtained in other studies according to which estimates obtained with using education as an instrument tend to be larger than without, because of a stronger direct effect of education. This does not seem to be the case in our data. Conversely, specification D removes the EGP-social class from the prediction of father's earnings. Comparing specifications B and D indicates that the inclusion of social class as an instrument tends to increase the estimates of the IGE. The effect is relatively large, around .05, for the elasticity of sons' family to the predicted family income of their parents. On the contrary, in the case of daughters, the effect of excluding father's social class from the set of instruments is very small.

Excluding self-employment status and region of residence from the base specification in specification E, leads to slightly lower estimates (.01 to .03) for both daughters and sons and in most cases, although the effect is virtually zero for daughters in column 3 and for daughters in column 1.

All in all, this sensibility analysis suggests that the IGE estimates previously discussed are fairly robust to variation in the set of first-step instruments. If anything, alternative sets of instruments lead to slightly lower estimates.

On the opposite, IGE estimates are strongly sensitive to the adequate modeling of cohort trends in the earnings structure. This is demonstrated by specification F, which builds upon specification A but removes all cohort trends in the effect of education, employment status and geographical area. This results in higher estimates for both sons and daughters. Of course, in the presence of significant trends, these estimates are incorrect.

6.2 Sensitivity to age selection

The recent literature has emphasized the sensitivity of IGE estimates to the age at which both children and their fathers are observed (Jenkins 1987, Grawe 2006, Haider & Solon 2006). This has been referred to as the *life-cycle bias* in IGE estimates. This bias arises from the presence

of individual heterogeneity in age-earnings growth rates, when using current earnings, instead of permanent earnings, in the estimation of the IGE.²¹ As shown in several papers, individual-specific earnings growth rates tend to be positively correlated with initial earnings level. As a consequence, inequality of current earnings underestimates permanent earnings inequality when measured early in the life-cycle and overestimates it when measured at the end of the career. This introduces an asymmetric bias in the estimation of the IGE. Using current earnings early (resp. late) in the life-cycle, as a proxy for *child's* permanent earnings will lead to underestimate (resp. overestimate) β . Conversely, using current earnings early (resp. late) in the life-cycle, as a proxy for *father's* permanent earnings will lead to overestimate (resp. underestimate) the IGE.

Several solutions have been offered to account for this bias. Based on the analysis of individual age-earnings profiles in the United States, Haider & Solon (2006) conclude that earnings observed at ages ranging from the early thirties to the mid forties are fairly representative of lifetime earnings inequality. Furthermore, if idiosyncratic departures from the mean age-permanent-earnings profile are uncorrelated across generations of the same family, then regressing child's current earnings measured in the early thirties to mid forties will provide a consistent estimate. In line with this result, we chose an age interval for sons centered around age 40 and excluding early and late ages (below 30 and above 50). For similar reasons, we predicted father's earnings at the age of 40. Provided that the results of Haider and Solon generalize to the Japanese case, this should purge our estimates from possible life-cycle bias. However, several issues should be discussed and taken into account in our estimation.

One first difficulty is that there might be some residual bias arising from the fact that current earnings are not necessarily representative of lifetime inequality over the entire age interval 30 to 50 years. If this is the case, our results might fail to provide consistent estimates of the IGE in long-term economic status. To address this concern, Lee & Solon (2009) propose to explicitly model the dependency of the IGE with respect to the age of children by including an interaction term between age and father's income. This leads to an equation of the form :

$$Y_{it} = \alpha_t + \beta \hat{X}_i + f(\text{age}_{it}) \times \hat{X}_i + g(\text{age}_{it}) + e_{it} \quad (5)$$

The age profile f will parametrically capture the life-cycle bias, which varies with age. Lee and Solon suggest to normalize the age profile f to be equal to zero at age 40, i.e. at an age where

²¹As already discussed, classical measurement error also introduce an attenuation bias. TSIV estimation corrects for this attenuation bias but is still subject to the life-cycle bias.

current earnings differentials are representative of lifetime earnings inequality. In this case, the main effect of father's income, β is consistent for the IGE under the assumptions of Haider & Solon (2006).²²

A second difficulty is raised by the recent paper of Nybom & Stuhler (2011). These authors challenge the assumption of Haider and Solon that individual departures from the mean age-permanent-earnings profile are uncorrelated across generations of the same family. If this assumption is invalid, the proper age for consistently estimating the IGE might differ from the age at which current earnings are representative of lifetime earnings inequality within a given generation. Using long-term earnings profiles for Sweden, they show that this is indeed the case. The pessimistic conclusion they draw is that the IGE could never be satisfactorily estimated without full earnings chronicles, which is obviously impossible to implement in most countries, including Japan, given data limitations. One important practical, and less radical, conclusion that can be derived from their analysis, however, is that one should carefully examine how the estimated IGE varies with the age selection rule for children.

In the light of these difficulties, it seems important to gauge the robustness of our results to alternative modelings of age effects. We consider two approaches. The first is to implement the parametric control for life-cycle bias introduced in Lee and Solon. The second is to vary the age criterion for selection in our sample of children. Results are given in table 7 for sons and table 8 for daughters. All estimates are presented for the piece-wise linear trends specification of the first-step equation. Column 2 reproduces the coefficients discussed in the previous section. Including an interaction between father's predicted income and a quadratic function of child's age, in column 1, turns out to be statistically insignificant and has no impact on the estimated main effect. This holds true for both sons and daughters and regardless of the combination of income variables used.

!! Tables 7 and 8 about here !!

In columns 3 to 6, we change the age interval for children and consider three partitions of the reference age group (30-50) : 30-40, 40-50, 35-45. We also consider a broader age interval

²²Lee & Solon (2009) suggest including the interaction between child's age and father's income, in an OLS context. A further issue arises in our case from the fact that we are implementing a TSIV estimator, i.e. using predicted father's permanent income instead of the true value (or a good proxy of it, as provided by multi-year averages). As discussed in Nybom & Stuhler (2011), the results of Haider & Solon (2006) (henceforth HS) carry over to the TSIV context. Under the assumptions of HS, predicted father's income at age 40 will consistently estimate lifetime permanent income differentials among fathers. Substituting father's predicted income for father's lifetime income will still be distorted by the presence of non-classical measurement error in the right-hand side variable compared to the TSIV estimate one would obtain in the absence of measurement error in the RHS. However, correcting for life-cycle bias can also be undertaken in this context using the rule of thumb of HS, or using the parametric interaction term of Lee and Solon, provided that the age-profile of the life-cycle bias can be proxied by a parametric functional taking value zero at age 40.

: 25-55. All in all, results are very stable across these different age intervals. For sons results vary by less than .02 in absolute value and the difference in point estimates is never statistically significant. For daughters, the results seem more sensitive in the case of the IGE of personal income to father's personal income. For the younger age group, the IGE is .5 and .27 for the older age group, with a reference value of .35. Yet, estimates are fairly imprecise and the difference in point estimates is not statistically significant. When regressing family income on family income of the parents, the results for daughters are again insensitive to the age selection. These results partially contradict previous evidence that have emphasized the sensitivity of IGE estimates to the age of children. Two effects are likely to account for this result. The first-one is that the age groups we consider are all centered around age 40 and are never restricted to the sole groups of young or old children. Second, when changing the age interval, the set of cohorts used in the estimation will mechanically change and we might mix up cohort heterogeneity in the IGE with age effects. For this reason, we replicate our analysis on the sample of children born between 1945 and 1965 : since data for children are collected in 1985, 1995 and 2005, all cohorts between 1945 and 1965 will be used in the estimation, regardless of the age selection rule that we consider. Results are given in columns 7 to 12. They confirm that our estimates of the IGE are very robust to changes in the age selection rule for children.

7 Complementary results

We now supplement our analysis of intergenerational mobility in Japan with two complementary discussions. The first one focuses on changes across cohorts in the IGE. The second one pertains to the estimation of the intergenerational correlation coefficient.

7.1 Changes across cohorts in the IGE

To examine the possibility of changes over time in the value of the IGE, we interact our measure of predicted father's income with a dummy variable equal to one for cohorts born after 1952. The choice of the cut-off date roughly splits our sample in two equal-sized groups and allows to isolate the most recent cohorts who entered the labor market after the period of high economic expansion that followed World War II.

As discussed in Hertz (2007) and Lee & Solon (2009), one of the difficulties of estimating trends in the IGE is to separate cohort trends from the potential confounding influence of life-cycle effects. As discussed in the previous section, cohorts in our sample are observed at different

points of their life-cycle. In particular, the youngest cohort are only observed in the early stage of their career while the oldest ones are observed at the end of it. This is not true of the middle cohorts : given the age restrictions we impose and the survey waves used, individuals born between 1945 and 1965 are typically observed twice in their mid-career, i.e. in their thirties and in their forties. Hence we test for trends on two distinct sample : first, the entire sample of children; second, the subsample of individuals born between 1945 and 1965.

Results are given in table 9, for sons (columns 1 to 3) and daughters (columns 4 to 6). They indicate that the IGE was remarkably stable across cohorts over the period studied here. This converges with the results of Ishida & Miwa (2008) who find that social mobility has remained stable in Japan in the second half of the twentieth century. For sons, the estimated change is not statistically significant. This is mostly due to the small estimated value of the change in the IGE (less than 3% of the mean sample's value), rather than to an imprecise estimate of changes at work. For instance, we computed the 95% confidence interval for the IGE changes, which shows that we can rule out changes larger than .026 in absolute value. As a comparison, (Aaronson & Mazumder 2008) report that the IGE in the United States fell by over .3 between the cohorts born in the 1900s and those born in the 1930s and Lefranc (2011) reports a fall of almost .15 in two decades in France.

!! Table 9 about here !!

The same picture holds true for daughters when looking at the elasticity of daughter's family income to their parents' family income (column 6) : the change is very small and not statistically significant. Results for the changes in the elasticity of daughters' own income to their father's income are slightly different. The change in the IGE in columns 4 and 5 is larger and suggests a rise in the IGE for daughters in recent cohorts. This change is even significant at the 10% level for column 4, on the subsample of daughters born between 1945 and 1965. The rise in the IGE for female could be explained by changes over time in the Japanese female labor market. For women born before 1952, employment concentrates in low-status and low-wage occupations, so called 'pink-collar' jobs, where the status attainment process is very limited compared with men (Imada 1998) with little promotion perspectives (Suzuki 2005). A large share of part-time work is also observed, together with a high share of family employment. This leads to a fairly compressed wage structure that will be reflected directly into a low IGE. One should of course be cautious when interpreting these trends since they are only marginally significant.

7.2 Intergenerational correlation

As a measure of economic mobility, the IGE has sometimes been criticized for being sensitive to changes in cross-sectional inequality between the parents and the children’s generations. For instance, a reduction in earnings inequality among children will “mechanically” lead to a fall in the IGE. One might of course reasonably argue that such a fall in the IGE would indeed capture a ‘true’ rise in economic mobility. At the same time, this sensitivity might be problematic, for instance when comparing countries that have experienced different trends in cross-sectional inequality.

This suggests to supplement the IGE with an alternative measure of mobility that would be independent of cross-sectional changes in earnings inequality. This is the case of the intergenerational correlation coefficient (IGC). The IGC, which we will denote by ρ , measures positional mobility and is by construction unaffected by changes in the variance of earnings in the children’s or father’s generation. The link between the IGC and the IGE is given by:

$$\beta = \rho \frac{\sigma_Y}{\sigma_X} \tag{6}$$

In the steady state, the variance of earnings is constant across generations and the IGE and IGC are identical. This is no longer the case whenever σ_Y and σ_X differ. Furthermore, this formula makes clear that changes in the IGE will reflect both changes in positional mobility and the evolution across generations of earnings inequality.

To infer the IGC, we rescale our estimates of the IGE using an estimate of the ratio $\frac{\sigma_X}{\sigma_Y}$. The main challenge here is that we do not observe permanent (log) earnings for children (Y) and fathers (X), but only current earnings, denoted respectively by y and x . To estimate permanent earnings inequality, we make the assumption that current earnings inequality at age 40 are representative of lifetime inequality. More precisely we posit the following model, for current log earnings y :

$$y = E(y|Z, age) + u$$

where Z denote some variables used to predict earnings, u is assumed to be distributed independently of age, and the conditional expectation is modeled as the first-step equation discussed in section 2. Given Z and u , and assuming that the error term u is an individual fixed-effect, log

earnings at age 40, \tilde{y} , can be predicted as :

$$\begin{aligned}\tilde{y} &= E(y|Z, age = 40) + u \\ &= E(\tilde{y}|Z) + u\end{aligned}$$

Under the assumption that current earnings inequality at age 40 are representative of lifetime inequality, permanent-earnings inequality can be proxied by $\sigma(\tilde{y}) = \sigma(\tilde{y}|Z) + \sigma(u)$. The assumption that the error term u is an individual fixed-effect is innocuous. The only substantial assumption made on the residual is that it homoscedastic with age and that current-earnings residual inequality is a good proxy for permanent-earnings residual inequality. Alternatively, one might assume that the residual heterogeneity is purely transitory and vanishes in the long-term. This would lead to proxy permanent earnings inequality by $\sigma(\tilde{y}|Z)$. Both approaches can be used to predict the ratio $\frac{\sigma_X}{\sigma_Y}$ in order to derive an estimate of the IGC.

Results are given in table 10. IGC1 is computed under the first approach discussed above to infer permanent earnings inequality. IGC2 corresponds to the second approach. In the case of sons, earnings inequality at age 40 is very similar in the children and fathers' generations and, in most cases, only slightly lower for children. As a result, the adjustment factor $\frac{\sigma_X}{\sigma_Y}$ is close to one and usually slightly higher than one. This produces estimates of the IGC that are only slightly higher than the values of the IGE previously discussed. In fact, in the case of the elasticity of sons' family income to their fathers' family income, the IGC is almost identical to the IGE, at .3, and independent of the method used to estimate permanent earnings inequality. For the elasticity of children's own income to their father's own income, the IGC is slightly higher and lies in the interval .34-.36.

!! Table 10 about here !!

For daughters, the analysis is complicated by the large prevalence of part-time employment on the Japanese female labor market. This prevalence results in a much larger degree of inequality of own income among women than among men : in fact, the standard deviation of own income for women is almost twice as high as the one found for men. As a result, the adjustment factor $\frac{\sigma_X}{\sigma_Y}$ estimated for daughters lies close to one half. And the estimated IGC for women in the case of own income lies between .17 and .19. One can correct this estimates for the incidence of part-time work. This is undertaken in column 4, where we predict potential full-time earnings on the basis of a sample selection model whose wage equation is estimated

on the sole subsample of women working full-time. This model predicts a much lower degree of earnings inequality among women and leads to estimates of the IGC that are much higher, and around .27. In the end, the estimates in columns 1, 2 and 4 indicate that parental income has a sizable influence on female potential full-time earnings but is less strongly associated with the occurrence of part-time work. Since part-time work is an important driver of earnings inequality in Japan, the overall association between women's actual earnings and their father's income turns out to be modest. However, this is no longer the case when looking at family income. While own income is somewhat less correlated to parental income for daughters than for sons, the correlation in family income turns out to be very close in the case of family income, and slightly higher for female (.33) than for male (.30).

8 Conclusion

In this paper, we have examined the extent of intergenerational transmission of income inequality for sons and daughters in Japan. Our estimates suggests that a value of the IGE around .35 adequately summarizes the persistence of inequality between fathers and children for both sons and daughters. Overall, this value puts Japan in an intermediate position, compared to other developed countries, on the scale of intergenerational mobility. For instance, among the countries surveyed in Björklund & Jäntti (2009), Japan appears more mobile than the United States, the United Kingdom, Italy and France, all of which exhibit an IGE greater than .4, comparable to Germany but less mobile than Scandinavian countries, Australia and Canada.

At first sight, this rather high degree of mobility may seem hard to reconcile with the alleged importance of parental investment in child's education at work in Japan. As already abundantly described in several studies, access to higher education in Japan is often expensive and selective, forcing families to elaborate complex educational strategies and to undertake significant financial investments to support them. For instance families cover between 71 and 86 % of the annual expenditures of university students (Kondo 2000). Furthermore, besides tuition fees, parents often invest significant amounts in "shadow education" such as cram schools and private tutoring.

Two factors, however are likely to limit the incidence of these investments for the transmission of income across generations. Both factors, emphasize the role of compressed earnings differentials, but at different stages of the intergenerational mobility process. The first argument emphasizes the small degree of inequality among Japanese parents. In such a context, even if parents invest a large share of their income in their children's education, in the end, the resulting distribution of human capital in the next generation will also be relatively equal. This

is confirmed by the analysis of “shadow education” undertaken in Stevenson & Baker (1992), who emphasize the following three aspects of private investment in education in Japan. First, family financial investment is high on average. Second, such investment is efficient at improving educational attainment. But third, financial investment and the use of shadow education seem to vary little with characteristics of the family background such as parental education or family income. The second argument applies to the end of the intergenerational transmission process. As discussed for instance in Solon (2004), lower returns to human capital, as seems to be the case in Japan compared to most developed countries, will translate into a lower IGE and limit the income consequences of inequalities of parental investment.

Part of the high degree of intergenerational mobility observed in our sample may also be explained by the specificities of the high economic expansion period that develop after World War II. Of course, the high aggregate growth that Japan experienced in this period came along with a wave of rapid industrial development and occupational change that may have fostered intergenerational mobility. However, it is important to stress that our results do not suggest any slowdown in the mobility process for the cohorts of children born after 1952 and who entered the labor market after the high-growth era. Whether this relatively high degree of mobility will be maintained in the face of the recent rise in earnings inequality described in for instance in Tachibanaki (2009) is of course an open question.

References

- Aaronson, D. & Mazumder, B. (2008). Intergenerational Economic Mobility in the United States, 1940 to 2000, *Journal Human Resources* **43**(1): 139–172.
- Altonji, J. G. & Dunn, T. A. (1991). Relationships among the family incomes and labor market outcomes of relatives, *NBER Working Papers 3724*, National Bureau of Economic Research, Inc.
- Altonji, J. G. & Williams, N. (2005). Do wages rise with job seniority? a reassessment, *Industrial and Labor Relations Review* **58**(3): 370–397.
- Anger, S. & Heineck, G. (2010). Do smart parents raise smart children? the intergenerational transmission of cognitive abilities, *Journal of Population Economics* **23**: 1105–1132. 10.1007/s00148-009-0298-8.
- Angrist, J. D. & Krueger, A. B. (1995). Split-sample instrumental variables estimates of the return to schooling, *Journal of Business and Economic Statistics* **13**(150): 225–235.
- Arellano, M. & Meghir, C. (1992). Female labour supply and on-the-job search : empirical model estimated using complementary data sets, *The Review of Economic Studies* **59**: 537–559.
- Björklund, A. & Jäntti, M. (1997). Intergenerational income mobility in Sweden compared to the United States, *American Economic Review* **87**(5): 1009–1018.
- Björklund, A., Eriksson, T., Jäntti, M., Raaum, O. & Österbacka, E. (2002). Brother correlations in earnings in denmark, finland, norway and sweden compared to the united states,

Journal of Population Economics **15**: 757–772. 10.1007/s001480100095.
URL: <http://dx.doi.org/10.1007/s001480100095>

- Björklund, A. & Jäntti, M. (2009). Intergenerational income mobility and the role of family background, in W. Salverda, B. Nolan & T. Smeeding (eds), *Oxford Handbook of Economic Inequality*, Oxford University Press, chapter 20, pp. 491–521.
- Black, S. E. & Devereux, P. J. (2010). Recent developments in intergenerational mobility, in D. Card & O. Ashenfelter (eds), *Handbook of Labor Economics*, Vol. 4B, North-Holland, Amsterdam, chapter 8.
- Bowles, S. & Gintis, H. (2002). The inheritance of inequality, *Journal of Economic Perspectives* **16**(3): 3–30.
- Chadwick, L. & Solon, G. (2002). Intergenerational income mobility among daughters, *American Economic Review* **92**(1): 335–344.
- Ermisch, J., Francesconi, M. & Siedler, T. (2006). Intergenerational mobility and marital sorting, *Economic Journal* **116**(513): 659–679.
- Genda, Y. (1994). Skill premiums and japan’s wage structure in the 1980s, *Gakushuin Economic Papers* **31**(3): 125–158.
- Genda, Y. & Kambayashi, R. (2002). Declining self-employment in japan, *Journal of the Japanese and International Economies* **16**(1): 73–91.
- Gottschalk, P. & Smeeding, T. (2000). Empirical evidence on income inequality in industrial countries, in A. B. Atkinson & F. Bourguignon (eds), *Handbook of Income Distribution*, Vol. 1, North-Holland, pp. 261–308.
- Grawe, N. (2006). Lifecycle bias in estimates of intergenerational earnings persistence., *Labour Economics* **13**(5): 551–570.
- Haider, S. & Solon, G. (2006). Life-cycle variation in the association between current and lifetime earnings., *American Economic Review*, **96**(4): 1308–1320.
- Hashimoto, M. & Raisian, J. (1985). Employment tenure and earnings profiles in japan and the united states, *American Economic Review* **75**(4): 721–35.
- Hertz, T. (2007). Trends in the intergenerational elasticity of family income, *Industrial Relations* pp. 22–50.
- Imada, S. (1998). Status attainment of women: Closed hierarchical space, *International Journal of Sociology* **28**(1): 66–91.
- Imada, T. (2000). Industrialization and the regime of social mobility in postwar Japan, *International Journal of Japanese Sociology* **9**: 35–52.
- Ishida, H. (1990). Educational credentials and the labour market in japan: A cross-national comparison with the united states and britain, *International Journal of Comparative Sociology* **31**(3-4): 139–155.
- Ishida, H. (1993). *Social Mobility in Contemporary Japan: Educational Credentials, Class and the Labour Market in a Cross-National Perspective*, Stanford University Press, Stanford.
- Ishida, H. & Miwa, S. (2008). Trends in intergenerational class mobility and education in japan, in H. Ishida (ed.), *Social Stratification and Social Mobility in Late-Industrializing Countries (The 2005 SSM Research Series/ Volume 14)*, pp. 1–48.
- Jenkins, S. (1987). Snapshots vs movies: ‘lifecycle bias’ and the estimation of intergenerational earnings inheritance, *European Economic Review* **31**(5): 1149–1158.

- Kamayashi, R., Kawaguchi, D. & Yokoyama, I. (2008). Wage distribution in Japan, 1989-2003, *Canadian Journal of Economics* **41**(4): 1329–1350.
- Kanomata, N., Tanabe, S. & Takenoshita, H. (2008). SSM occupational code and international measures of occupational status: Conversion into EGP class schema, SIOPS and ISEI (in Japanese), in T. Maeda (ed.), *Problems in Measurement and Analysis in Social Surveys (The 2005 SSM Research Series/ Volume 12)*, pp. 69–94.
- Katz, L. & Revenga, A. (1989). Changes in the structure of wages: The united states vs japan, *Journal of the Japanese and International Economies* **3**(4): 522–553.
- Koike, K. (1988). *Understanding Industrial Relations in Modern Japan*, Macmillan, London.
- Kondo, H. (2000). Education and social mobility in postwar Japan : Trends and some institutional aspects, *International Journal of Japanese Sociology* **9**(1): 3–19.
- Lee, C.-I. & Solon, G. (2009). Trends in intergenerational income mobility, *Review of Economics and Statistics* **91**(4): 766–772.
- Lefranc, A. (2011). Educational expansion, earnings compression and changes in intergenerational economic mobility : Evidence from french cohorts, 1931-1976, *Technical report*.
- Lefranc, A., Ojima, F. & Yoshida, T. (2008). The intergenerational transmission of income and education: A comparison of Japan and France, *RSCAS Working Papers 2008/25*, European University Institute.
- Lefranc, A., Pistoiesi, N. & Trannoy, A. (2008). Inequality of opportunities vs. inequality of outcomes: Are western societies all alike ?, *Review of Income and Wealth* pp. 513–543.
- Lefranc, A. & Trannoy, A. (2005). Intergenerational earnings mobility in France : is France more mobile than the US ?, *Annales d'Economie et Statistique* (78).
- Liu, H. & Zeng, J. (2009). Genetic ability and intergenerational earnings mobility, *Journal of Population Economics* **22**: 75–95. 10.1007/s00148-007-0171-6.
URL: <http://dx.doi.org/10.1007/s00148-007-0171-6>
- Lucas, R. & Kerr, S. (forthcoming). Intergenerational income immobility in finland: contrasting roles for parental earnings and family income, *Journal of Population Economics* pp. 1–38. 10.1007/s00148-012-0442-8.
URL: <http://dx.doi.org/10.1007/s00148-012-0442-8>
- Mazumder, B. (2001). Earnings mobility in the US: A new look at intergenerational inequality, *University of California, Berkeley, Center for Labor Economics Working Paper* (34).
- Mulligan, C. B. (1997). Galton versus the human capital approach to inheritance, *Journal of Political Economy* **107**(6): S184–224.
- Nicoletti, C. & Ermisch, J. F. (2007). Intergenerational earnings mobility: Changes across cohorts in Britain, *The B.E. Journal of Economic Analysis & Policy* **7**(2).
- Nybom, M. & Stuhler, J. (2011). Heterogeneous income profiles and life-cycle bias in intergenerational mobility estimation, *IZA Discussion Papers 5697*, Institute for the Study of Labor (IZA).
- Oi, W. Y. (1990). Employment relations in dual labor markets (“ it’s nice work if you can get it”), *Journal of Labor Economics* **8**(1): S124–49.
- Ojima, F. (1998). An analysis of trends in educational opportunity, *International Journal of Sociology* **28**(1): 33–65.
- Sato, T. (2000). *Fubyodo Shakai Nihon [Japan as an Unequal Society]*, Chuo Koron Shinsha, Tokyo.

- Shimada, H. (1981). *Earnings structure and human investment : a comparison between the United States and Japan*, Kogakusha, Tokyo.
- Solon, G. (1992). Intergenerational income mobility in the United States, *American Economic Review* **82**(3): 393–408.
- Solon, G. (1999). Intergenerational mobility in the labor market, in D. Card & O. Ashenfelter (eds), *Handbook of Labor Economics*, Vol. 3B, North-Holland, Amsterdam, chapter 29, pp. 1761–1800.
- Solon, G. (2004). A model of intergenerational mobility variation over time and place, in M. Corak (ed.), *Generational Income Mobility in North America and Europe*, Cambridge University Press, chapter 3, pp. 38–47.
- Stevenson, D. L. & Baker, D. P. (1992). Shadow education and allocation in formal schooling: Transition to university in japan, *The American Journal of Sociology* **97**(6): 1639–1657.
- Suzuki, A. (ed.) (2005). *Gender and Career in Japan*, Stratification and Inequality Series, Trans-Pacific Press.
- Tachibanaki, T. (2009). *Confronting Income Inequality in Japan*, MIT Press, Cambridge (MA).
- Topel, R. (1991). Specific Capital, Mobility and Wages : Wages Rise With Seniority, *Journal of Political Economy* **99**(1): 145–176.
- Ueda, A. (2009). Intergenerational mobility of earnings and income in Japan, *The B.E. Journal of Economic Analysis & Policy* **9**(1): 54.
- Yano, M. (1982). Nyugaku to shushoku no keizaigaku [economics of entry school and getting job], in S. Ichikawa, J. Kikuchi & M. Yano (eds), *Kyowiku no Keizaigaku [Economics of Education]*, Dai-ichi Hoki, Tokyo, pp. 39–61.
- Yoshida, T. (2008). Inequality of opportunity from a perspective of intergenerational income mobility (in japanese), in T. Watanabe (ed.), *Intergenerational Mobility and Intragenerational Mobility (The 2005 SSM Research Series/ Volume 3)*, pp. 147–160.
- Zimmerman, D. (1992). Regression toward mediocrity in economic stature, *American Economic Review* **82**(1): 409–429.

Table 1: Descriptive statistics

	Fathers sample		Sons sample		Daughters sample	
	mean	sd	mean	sd	mean	sd
birth year	1941.127	14.957	1953.814	9.901	1956.203	9.702
survey year	1965 1975 1985 1995 2005	0.178 0.230 0.210 0.197 0.184	0.000 0.000 0.371 0.336 0.293	0.000 0.000 0.224 0.411 0.365	0.000 0.000 0.224 0.411 0.365	
social status (EGP class)	I- Higher service II- Lower service III- Routine clerical/sales IV- Small employer and Independent V- Manual foreman VI- Skilled manual VII- Semi-Unskilled manual VII- Farmers/Farm workers	0.118 0.183 0.105 0.167 0.081 0.100 0.153 0.094	0.145 0.229 0.130 0.134 0.095 0.085 0.149 0.033	0.028 0.181 0.428 0.050 0.005 0.073 0.220 0.015		
education	primary or lower secondary education upper secondary education higher education	0.344 0.419 0.237	0.129 0.501 0.370	0.099 0.601 0.300		
firmsize	>1000 government other	0.164 0.106 0.730	0.187 0.119 0.694	0.099 0.091 0.810		
self-employed		0.321	0.236	0.196		
residential area	6 major cities urban rural	0.137 0.584 0.279	0.125 0.643 0.232	0.134 0.641 0.225		
income	ln(individual income) ln(family income)	5.940 6.203	0.687 0.650	6.190 6.449	0.531 0.518	0.913 0.560
N		7170	2656	2598		

Notes: sd refers to the standard deviation; N refers to the maximum number of observation. Actual numbers may vary because of missing observations. Fathers and sons samples exclude individuals with missing or zero earnings. Individuals with zero or missing earnings are kept in the descriptive statistics for female except for computation of mean income.

Table 2: First-step estimates - dependent variable : own income

	(1)		(2)		(3)	
	coef.	std.err.	coef.	std.err.	coef.	std.err.
primary or lower sec. education						
upper secondary education	REF		REF		REF	
higher education	0.134*** (0.0238)		0.125*** (0.0278)		0.137*** (0.0252)	
	0.185*** (0.0293)		0.171*** (0.0341)		0.191*** (0.0318)	
EGP class I	REF		REF		REF	
EGP class II	-0.205*** (0.0239)		-0.203*** (0.0239)		-0.209*** (0.0239)	
EGP class III	-0.430*** (0.0275)		-0.428*** (0.0274)		-0.431*** (0.0275)	
EGP class IV	-0.422*** (0.0285)		-0.421*** (0.0284)		-0.418*** (0.0285)	
EGP class V	-0.292*** (0.0301)		-0.289*** (0.0301)		-0.293*** (0.0302)	
EGP class VI	-0.522*** (0.0290)		-0.518*** (0.0289)		-0.524*** (0.0290)	
EGP class VII	-0.545*** (0.0268)		-0.543*** (0.0268)		-0.546*** (0.0269)	
EGP class VIII	-0.680*** (0.0326)		-0.674*** (0.0326)		-0.684*** (0.0325)	
self-employed	0.174*** (0.0293)		0.194*** (0.0339)		0.170*** (0.0316)	
firm size : >1000	0.263*** (0.0177)		0.262*** (0.0176)		0.261*** (0.0177)	
firm size : government	0.131*** (0.0216)		0.130*** (0.0215)		0.133*** (0.0216)	
residential area : 6 major cities	REF		REF		REF	
residential area : urban	-0.0769** (0.0235)		-0.0270 (0.0288)		-0.109*** (0.0257)	
residential area : rural	-0.117*** (0.0264)		-0.0466 (0.0336)		-0.168*** (0.0295)	
age	0.0135*** (0.00297)		0.0136*** (0.00321)		0.0107*** (0.00275)	
age ²	-0.000414 (0.000231)		-0.000430 (0.000237)		-0.000487* (0.000218)	
age × upper secondary educ.	0.00215 (0.00281)		0.00214 (0.00281)		0.00182 (0.00279)	
age × higher educ.	0.00906** (0.00324)		0.00878** (0.00323)		0.00846** (0.00321)	
age ² × upper secondary educ.	-0.0000996 (0.000233)		-0.0000888 (0.000232)		0.0000308 (0.000231)	
age ² × higher educ.	-0.000151 (0.000273)		-0.0000857 (0.000271)		0.0000611 (0.000270)	
age × self-employed	-0.0000332 (0.00255)		0.00000242 (0.00255)		-0.000317 (0.00256)	
age ² × self-employed	-0.000445* (0.000207)		-0.000501* (0.000206)		-0.000505* (0.000206)	

continued on next page

Table 2 - continued from previous page

	(1)		(2)		(3)	
	coef.	std.err.	coef.	std.err.	coef.	std.err.
cohort × upper secondary educ.	-0.00227	(0.00159)	0.000543	(0.00298)		
cohort ² × upper secondary educ.	0.000149*	(0.0000679)				
cohort × upper secondary educ. × pre1945			-0.00628	(0.00378)	0.125***	(0.0361)
upper secondary education × pre1931					0.0222	(0.0476)
upper secondary education × post1950						
cohort × higher educ.	-0.00190	(0.00173)	0.00192	(0.00315)		
cohort ² × higher educ.	0.000179*	(0.0000784)				
cohort × higher educ. × pre1945			-0.00830*	(0.00422)		
higher education × pre1931					0.0929*	(0.0456)
higher education × post1950					0.0383	(0.0512)
cohort × self-employed	-0.00401***	(0.00119)	-0.00681**	(0.00250)		
cohort ² × self-employed	-0.0000867	(0.0000547)				
cohort × self-employed			0.00573	(0.00346)	0.0609	(0.0324)
self-employed × pre1931					-0.0492	(0.0353)
self-employed × post1950						
cohort × residential area : urban	0.00291*	(0.00134)	-0.00321	(0.00256)		
cohort ² × residential area : urban	-0.0000717	(0.0000672)				
cohort × residential area : urban × pre1945			0.0111**	(0.00370)		
residential area : urban × pre1931					-0.0526	(0.0439)
residential area : urban × post1950					0.0595	(0.0450)
cohort × residential area : rural	0.00607***	(0.00154)	-0.00302	(0.00310)		
cohort ² × residential area : rural	-0.000160*	(0.0000748)				
cohort × residential area : rural × pre1945			0.0172***	(0.00441)		
residential area : rural × pre1931					-0.160***	(0.0476)
residential area : rural × post1950					0.109*	(0.0512)

continued on next page

Table 2 - continued from previous page

	(1)		(2)		(3)	
	coef.	std.err.	coef.	std.err.	coef.	std.err.
cohort × age	0.0000983	(0.000189)	-0.000367	(0.000240)		
cohort ² × age	0.00000834	(0.00000553)				
cohort × age × pre1945			0.000126	(0.000350)		
cohort × age ²	0.00000894	(0.0000110)	-0.0000663**	(0.0000234)		
cohort ² × age ²	-0.000000460	(0.000000363)				
cohort × age ² × pre1945			0.0000917**	(0.0000284)		
age × pre1931					0.00489	(0.00476)
age ² × pre1931					-0.000734*	(0.000308)
age × post1950					0.00318	(0.00286)
age ² × post1950					-0.000520	(0.000306)
Observations	7170		7170		7170	
R-squared	0.459		0.460		0.458	
Adjusted R-squared	0.456		0.457		0.455	
AIC	10597.0		10581.8		10613.0	
BIC	10879.0		10863.8		10908.8	

Notes : The dependent variable is the log of respondent's own income; standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01; the age variable is centered at age 40; the cohort variable is centered around birth cohort 1945; preYY is a dummy variable equal to one for cohorts born before year YY; postYY is a dummy variable equal to one for cohorts born after year YY; All equations include intercept and year dummies.

Table 3 - continued from previous page

	(1)		(2)		(3)	
	coef.	std.err.	coef.	std.err.	coef.	std.err.
cohort × upper secondary educ.		-0.00399*	(0.00173)	-0.000211	(0.00321)	
cohort ² × upper secondary educ.		0.0000527	(0.0000726)			
cohort × upper secondary educ. × pre1945				-0.00578	(0.00395)	0.108** (0.0349)
upper secondary education × pre1931						0.00374 (0.0493)
upper secondary education × post1950						
cohort × higher educ.		-0.00190	(0.00186)	0.00532	(0.00335)	
cohort ² × higher educ.		0.000175*	(0.0000825)	-0.0129**	(0.00435)	
cohort × higher educ. × pre1945						0.108* (0.0443)
higher education × pre1931						0.0830 (0.0527)
higher education × post1950						
cohort × self-employed		-0.00266*	(0.00120)	-0.00410	(0.00254)	
cohort ² × self-employed		-0.0000434	(0.0000548)	0.00286	(0.00347)	
cohort × self-employed						0.0473 (0.0314)
self-employed × pre1931						-0.0299 (0.0352)
self-employed × post1950						
cohort × residential area : urban		0.00251	(0.00134)	-0.00324	(0.00259)	
cohort ² × residential area : urban		-0.000102	(0.0000671)	0.0107**	(0.00371)	
cohort × residential area : urban × pre1945						-0.0641 (0.0423)
residential area : urban × pre1931						0.0706 (0.0446)
residential area : urban × post1950						
cohort × residential area : rural		0.00736***	(0.00158)	-0.00220	(0.00318)	
cohort ² × residential area : rural		-0.000192*	(0.0000757)	0.0181***	(0.00445)	
cohort × residential area : rural × pre1945						-0.208*** (0.0460)
residential area : rural × pre1931						0.139** (0.0509)
residential area : rural × post1950						

continued on next page

Table 3 - continued from previous page

	(1)		(2)		(3)	
	coef.	std.err.	coef.	std.err.	coef.	std.err.
cohort × age						
cohort ² × age	-0.0000382	(0.000198)	0.00000141	(0.000251)		
cohort × age × pre1945	0.00000853	(0.00000565)				
cohort × age ²	-0.00000493	(0.0000110)	-0.0000376	(0.000351)		
cohort ² × age ²	-0.000000249	(0.000000368)	-0.0000477*	(0.0000241)		
cohort × age ² × pre1945			0.0000557	(0.0000290)		
age × pre1931					0.00392	(0.00458)
age ² × pre1931					-0.000109	(0.000298)
age × post1950					0.00469	(0.00286)
age ² × post1950					-0.000540	(0.000310)
Constant						
Observations	5.764***	(0.0384)	5.778***	(0.0382)	5.777***	(0.0431)
R-squared	6814		6814		6814	
Adjusted R-squared	0.452		0.454		0.451	
AIC	0.449		0.450		0.447	
BIC	9448.9		9435.7		9467.3	
	9728.8		9715.6		9760.9	

Notes : The dependent variable is the log of respondent's family income; standard errors in parentheses; * p<0.10, ** p<0.05, *** p<0.01; the age variable is centered at age 40; the cohort variable is centered around birth cohort 1945; preYY is a dummy variable equal to one for cohorts born before year YY; postYY is a dummy variable equal to one for cohorts born after year YY; All equations include intercept and year dummies.

Table 4: IGE estimates for sons and daughters : main results

	(1)	(2)	(3)
Child's income variable	own	own	family
Father's income variable	own	family	family
A- sons, linear regression			
base specification - quadratic trends (QT)	.3239 (.02918)	.3657 (.03397)	.2897 (.03440)
base specification - piecewise linear trends (PW)	.3368 (.02999)	.3738 (.03444)	.2969 (.03488)
base specification - stepwise trends (SW)	.3669 (.0317)	.4200 (.03706)	.3254 (.03772)
Observations	2265	2265	2106
B- daughters, linear regression			
base specification - quadratic trends (QT)	.2155 (.068555)	.2551 (.07925)	.3748 (.03897)
base specification - piecewise linear trends (PW)	.2273 (.07025)	.2660 (.08035)	.3759 (.03953)
base specification - stepwise trends (SW)	.2676 (.07462)	.3252 (.08679)	.3991 (.04272)
Observations	1555	1555	1921
C- daughters, heckman selection model			
base specification - quadratic trends (QT)	.3444 (.07603)	.4043 (.08799)	.3724 (.03925)
base specification - piecewise linear trends (PW)	.3544 (.07783)	.4136 (.08915)	.3738 (.03981)
base specification - stepwise trends (SW)	.4079 (.08275)	.4915 (.09643)	.3968 (.04302)
Observations	2456	2456	2456

Notes: standard errors in parentheses. Columns 1 to 3 report IGE estimates for sons and daughters based on equation 2 for different combinations of father's and child's income variables. For each pair of income measures, the different lines report the IGE estimates obtained for the three specifications of the first-step equation presented in columns 1, 2 and 3 of tables 2 and 3. Panels A and B are estimated using 2SLS. Panel C is estimated using the Heckman's sample selection model. The selection equation uses the following regressors : marital status, indicators of the number of children, spouse's income and education when married, age profile and year dummies.

Table 5: IGE estimates for sons : sensitivity to first-step specification

			(1)	(2)	(3)
Child's income variable			own	own	family
Father's income variable			own	family	family
Specification	Father's characteristics in first-step equation	Trends specification			
(A)	base	QT	.3239 (.02918)	.3657 (.03397)	.2897 (.03440)
		PW	.3368 (.02999)	.3738 (.03444)	.2969 (.03488)
		SW	.3669 (.03176)	.4200 (.03706)	.3254 (.03772)
(B)	(A)+ trends for EGP-class + EGP-class specific age profiles	QT	.3082 (.02822)	.3433 (.03213)	.2772 (.03248)
		PW	.3078 (.02803)	.3449 (.03186)	.2767 (.03219)
		SW	.3593 (.03107)	.4071 (.03556)	.3171 (.03617)
(C)	(B) - father's education variables	QT	.3066 (.02926)	.3366 (.03324)	.2944 (.03357)
		PW	.3268 (.03014)	.3732 (.03490)	.3040 (.03530)
		SW	.3914 (.03443)	.4443 (.03948)	.3475 (.04020)
(D)	(A) - father's EGP class	QT	.2871 (.03469)	.3060 (.03911)	.2378 (.03942)
		PW	.3014 (.03582)	.3159 (.03980)	.2455 (.04012)
		SW	.3561 (.04005)	.3821 (.04521)	.2791 (.04587)
(E)	(A) - job characteristics - regional information	QT	.3006 (.02869)	.3494 (.03348)	.2892 (.03376)
		PW	.2966 (.02840)	.3441 (.03303)	.2878 (.03329)
		SW	.3467 (.03167)	.4046 (.03705)	.3301 (.03754)
(F)	notrends		.4018 (.03420)	.4661 (.03991)	.3534 (.04077)
Observations			2265	2265	2106

Notes: standard errors in parentheses. Columns 1 to 3 report IGE estimates for sons and daughters based on equation 2 for different combinations of father's and child's income variables. For each pair of income measures, the different lines report the IGE estimates obtained for various specifications of the first-step equation used to predict father's income. Specifications differ in both the variables used in the first-step equation and the parametric form used for modeling changes cohort in the coefficient of various predictors in the wage equation. As far as trends specification is concerned, QT denotes quadratic cohort trends, PW denotes piece-wise linear trends, SW denotes step-wise changes across cohorts. See p.6 for a full discussion. The variables included in the first-step equation are : (A) : see variable list in table 2; (B)=(A) plus trends for EGP-class and EGP-class specific age profiles; (C)=(B) minus father's education variables and their interactions; (D)=(A) minus father's EGP class; (E)=(A) minus job characteristics and regional information; (F)=(A) without cohort trends.

Table 6: IGE estimates for daughters : sensitivity to first-step specification

			(1)	(2)	(3)
Child's income variable			own	own	family
Father's income variable			own	family	family
Specification	Father's characteristics in first-step equation	Trends specification			
(A)	base	QT	.3444 (.07603)	.4043 (.08799)	.3724 (.03925)
		PW	.3544 (.07783)	.4136 (.08915)	.3738 (.03981)
		SW	.4079 (.08275)	.4915 (.09643)	.3968 (.04302)
(B)	(A)+ trends for EGP-class + EGP-class specific age profiles	QT	.3368 (.07456)	.3719 (.08420)	.3485 (.03752)
		PW	.3313 (.07386)	.3795 (.08366)	.3418 (.03733)
		SW	.3957 (.08158)	.4574 (.09287)	.3744 (.04145)
(C)	(B) - father's education variables	QT	.3193 (.07621)	.3470 (.08586)	.3289 (.03869)
		PW	.3401 (.07852)	.3961 (.09021)	.3403 (.04074)
		SW	.4229 (.08915)	.4806 (.10126)	.3696 (.04589)
(D)	(A) - father's EGP class	QT	.2594 (.08791)	.2914 (.09861)	.3698 (.04451)
		PW	.2649 (.09045)	.2989 (.10034)	.3703 (.04533)
		SW	.3438 (.10159)	.3973 (.11440)	.4158 (.05154)
(E)	(A) - job characteristics - regional information	QT	.3513 (.07749)	.3814 (.08987)	.3546 (.03990)
		PW	.3400 (.07628)	.3787 (.08852)	.3478 (.03939)
		SW	.4022 (.08442)	.4518 (.09828)	.3856 (.04392)
(F)	notrends		.4404 (.08936)	.5523 (.10421)	.4391 (.04661)
Observations			2456	2456	2456

Notes: standard errors in parentheses. Columns 1 to 3 report IGE estimates for sons and daughters based on equation 2 for different combinations of father's and child's income variables. For each pair of income measures, the different lines report the IGE estimates obtained for various specifications of the first-step equation used to predict father's income. Specifications differ in both the variables used in the first-step equation and the parametric form used for modeling changes cohort in the coefficient of various predictors in the wage equation. As far as trends specification is concerned, QT denotes quadratic cohort trends, PW denotes piece-wise linear trends, SW denotes step-wise changes across cohorts. See p.6 for a full discussion. The variables included in the first-step equation are : (A) : see variable list in table 2; (B)=(A) plus trends for EGP-class and EGP-class specific age profiles; (C)=(B) minus father's education variables and their interactions; (D)=(A) minus father's EGP class; (E)=(A) minus job characteristics and regional information; (F)=(A) without cohort trends. All IGE's are estimated using the Heckman's sample selection correction.

Table 7: IGE estimates for sons : sensitivity to age effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
age interval for children	30-50	30-50	30-40	40-50	35-45	25-55	30-50	30-50	30-40	40-50	35-45	25-55
	A-child's own income regressed on father's own income											
	all cohorts						1945-1965 cohorts					
IGE	.3190 (.04355)	.3368 (.02999)	.3450 (.03998)	.3211 (.04248)	.3249 (.03955)	.3504 (.02665)	.3339 (.05087)	.3426 (.03498)	.3503 (.04392)	.3234 (.05161)	.3357 (.04613)	.3501 (.03273)
$\ln(\text{father's income}) \times \text{age}$	-0.02079 (.004928)											
$\ln(\text{father's income}) \times \text{age}^2$.0005315 (.0008942)											
Observations	2265	2265	1132	1236	1242	3138	1604	1604	843	863	900	1922
p-values	.769											
	B- child's family income regressed on father's family income											
	all cohorts						1945-1965 cohorts					
IGE	.3158 (.05057)	.2969 (.03488)	.3126 (.04992)	.2778 (.04699)	.3041 (.04362)	.3141 (.03160)	.3052 (.05919)	.3065 (.04051)	.3178 (.05311)	.2881 (.05817)	.3086 (.05158)	.3360 (.03897)
$\ln(\text{father's income}) \times \text{age}$	-.003875 (.005789)											
$\ln(\text{father's income}) \times \text{age}^2$	-.0004776 (.001053)											
Observations	2106	2106	1037	1164	1171	2890	1502	1502	800	796	848	1788
p-values	.716											

Notes: Standard errors in parentheses. Columns 1 to 12 report estimation results using various children age groups for sons. Estimates are based on equation 2 where in the first-step, father's earnings are predicted using the base specification of tables 2 and 3, using a piece-wise linear cohort trends specification. Columns 1 and 7 include interactions terms between child's age and father's earnings. The reported p-values are for the test of joint significance of these interactions terms. In panel A, the IGE is estimated for the regression of child's own income on father's own income. In panel B, the IGE is estimated for the regression of child's family income on father's family income. Estimates for daughter are based on the Heckman sample selection model.

Table 8: IGE estimates for daughters : sensitivity to age effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
age interval for children	30-50	30-50	30-40	40-50	35-45	25-55	30-50	30-50	30-40	40-50	35-45	25-55
	A-child's own income regressed on father's own income											
	all cohorts											
IGE	.2971 (.1100)	.35441 (.0778)	.5019 (.1308)	.2724 (.0921)	.3866 (.1185)	.3548 (.06267)	.3104 (.1301)	.3416 (.09006)	.4707 (.15278)	.2847 (.1068)	.4854 (.1414)	.3538 (.08197)
$\ln(\text{father's income}) \times \text{age}$	-.01525 (.01132)						-.01817 (.01423)					
$\ln(\text{father's income}) \times \text{age}^2$.002141 (.002165)						.001712 (.002673)					
Observations	2456	2456	1237	1321	1270	3526	1775	1775	808	1069	960	2147
p-values	.311						.4133					
	B- child's family income regressed on father's family income											
	all cohorts											
IGE	.3183 (.05873)	.3738 (.03981)	.3507 (.05782)	.3906 (.05194)	.3269 (.05344)	.3464 (.03523)	.3307 (.06732)	.3798 (.04629)	.3616 (.06365)	.3951 (.06162)	.3202 (.06151)	.3587 (.04391)
$\ln(\text{father's income}) \times \text{age}$.002874 (.006533)						.004040 (.007725)					
$\ln(\text{father's income}) \times \text{age}^2$.001449 (.001219)						.001301 (.001418)					
Observations	2456	2456	1237	1321	1270	3526	1775	1775	808	1069	960	2147
p-values	.401						.541					

Notes: Standard errors in parentheses. Columns 1 to 12 report estimation results using various children age groups for sons. Estimates are based on equation 2 where in the first-step, father's earnings are predicted using the base specification of tables 2 and 3, using a piece-wise linear cohort trends specification. Columns 1 and 7 include interactions terms between child's age and father's earnings. The reported p-values are for the test of joint significance of these interactions terms. In panel A, the IGE is estimated for the regression of child's own income on father's own income. In panel B, the IGE is estimated for the regression of child's family income on father's family income. All IGE's are estimated using the Heckman sample selection correction.

Table 9: Trends in the IGE : sons and daughters

	(1)	(2)	(3)	(4)	(5)	(6)
Child's income variable	own	own	family	own	own	family
Father's income variable	own	family	family	own	family	family
	A - all cohorts					
	sons			daughters		
ln(father's income)	0.3398 (0.03016)	0.3765 (0.03457)	0.2973 (0.03502)	0.3446 (0.07841)	0.4048 (0.08965)	0.3781 (0.04005)
ln(father's income) \times post1954	-0.006529 (0.007041)	-0.006151 (0.006827)	-0.0007549 (0.007016)	0.02123 (0.01753)	0.01946 (0.01697)	-0.007192 (0.007533)
Observations	2265	2265	2106	2456	2456	2456
95% confidence interval for IGE change	[-0.20 , .007]	[-0.20 , .007]	[-0.15 , .013]	[-0.13 , .056]	[-0.14 , .053]	[-0.22 , .008]
	B - 1945-1965 cohorts					
	sons			daughters		
ln(father's income)	0.3458 (0.03519)	0.3889 (0.04037)	0.3070 (0.04064)	0.3200 (0.09083)	0.3981 (0.1036)	0.3765 (0.04666)
ln(father's income) \times post1954	-0.007592 (0.009166)	-0.006951 (0.008885)	-0.001471 (0.009026)	0.03970 (0.02390)	0.03653 (0.02311)	0.006030 (0.01024)
Observations	1604	1604	1502	1775	1775	1775
95% confidence interval for IGE change	[-0.26 , .010]	[-0.24 , .010]	[-0.19 , .016]	[-.007 , .087]	[-.009 , .082]	[-0.14 , .026]

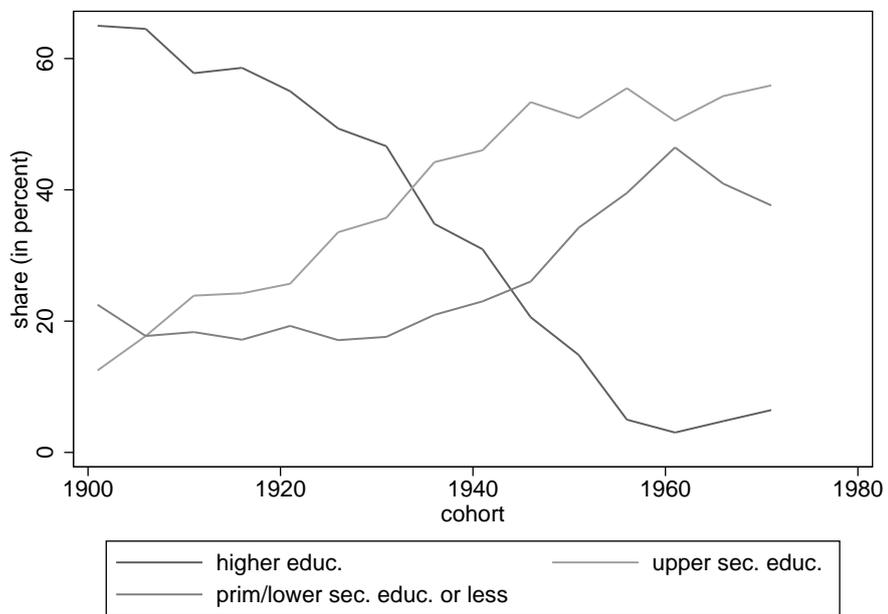
Notes: Standard errors in parentheses. Columns 1 to 6 report IGE estimates for sons and daughters based on equation 2 for different combinations of father's and child's income variables. Estimations include interactions between father's predicted income and a dummy variable equal to one for cohorts born after 1954 (post1954). The table only reports the main effect of father's predicted income and the interacted term. The estimation is performed for the whole sample of children (panel A) and for the subsample born between 1945 and 1965 (panel B). Coefficients for daughters are estimated using the Heckman sample selection correction.

Table 10: IGE and IGC: sons and daughters

	(1)	(2)	(3)	(4)
Child's income variable	own	own	family	own, full-time
Father's income variable	own	family	family	own
$\sigma(\tilde{Y}_f)$.5243	.5009	.5009	.5243
$\sigma(\tilde{Y}_f Z_f)$.264	.2181	.2181	.264
$\sigma(u_f)$.4574	.4555	.4555	.4574
	A- sons			
$\sigma(\tilde{Y}_c)$.4893	.4893	.4801	
$\sigma(\tilde{Y}_c Z_c)$.2592	.2592	.2126	
$\sigma(u_c)$.4238	.4238	.4357	
IGE	.3368	.3738	.2969	
IGC1	.3608	.3826	.3097	
IGC2	.3430	.3145	.3045	
	B- daughters			
$\sigma(\tilde{Y}_c X_c)$	1.0971	1.0971	.5728	.6976
$\sigma(\tilde{Y}_c)$.48878	.48878	.2406	.3486
$\sigma(u_c)$.98227	.98227	.5198	.6043
IGE	.3544	.4136	.3738	.3544
IGC1	.1693	.1888	.3268	.2663
IGC2	.1914	.1845	.3388	.2683

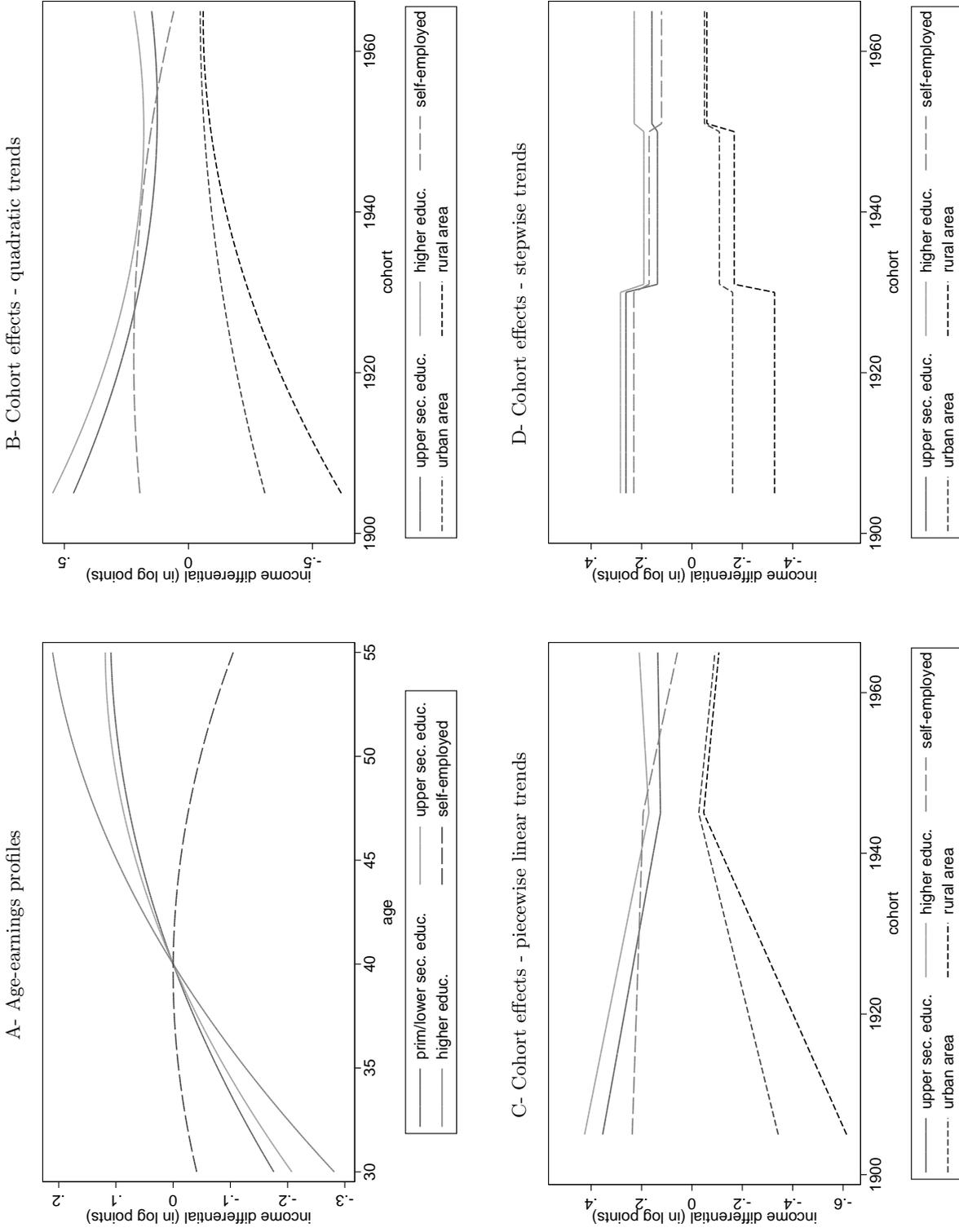
Notes: The table presents values of the IGE, the intergenerational correlation (IGC) and the standard-deviation of the log of income for fathers, sons and daughters and for different combinations of father's and child's income variables. u denotes the residual of the income regression presented in tables 2 and 3; $\tilde{y}|Z$ denotes the predicted log income at age 40 (see discussion p.25); $\tilde{y} = \tilde{y}|Z + u$. Subscript c indicates that standard-deviations are estimated on the sample of children; subscript f indicates that standard-deviations are estimated on a sample of individuals with birth year representative of the fathers of the children sample. For daughters, we re-estimated the model of tables 2 and 3, column 2, using Heckman sample selection procedure, where the selection equation included the following regressors : marital status, indicators of the number of children, spouse's income and education when married, age profile and year dummies. In column 4, own income is considered as censored when the daughter is not working full-time so Y refers to full-time potential income. IGEs are the same as in table 4. IGC1 is computed as $\text{IGE} \times \sigma(\tilde{y}_f) / \sigma(\tilde{y}_c)$. IGC2 is computed as $\text{IGE} \times \sigma(\tilde{y}_f|Z) / \sigma(\tilde{y}_c|Z)$.

Figure 1: Distribution of education by birth cohort



Notes: Source : authors' computation from SSM data.

Figure 2: Age and cohort effects, first-step income equation - own income



Notes: see table 2 for coefficient estimates.

Appendix : early fathers cohorts

Given the age restrictions we impose and the survey years available, our oldest cohort of children was born in 1935 and our oldest cohort of fathers was born in 1905. The tabulation of father's birth cohort by child's birth cohort, in table 11 below indicates that for the oldest child cohorts, a significant fraction of the fathers were born before 1905. This raises the question of how their income should be predicted given the limited information exploited in the first-step equation.

Table 11: Father's birth cohort by child's birth cohort

Child's birth Cohort	Father's birth cohort				
	mean	p10	p25	p75	p90
1935	1902.57	1892	1899	1906	1910
1936	1902.59	1893	1899	1908	1912
1937	1904.33	1897	1900	1909	1910
1938	1904.61	1896	1898	1912	1913
1939	1904.29	1895	1899	1910	1911
1940	1907.63	1902	1904	1912	1913

Notes: Source : authors' computations from SSM data. pxx designates the value of the xx th percentile.

We considered four options. The first, which we use in the rest of the paper, was to assume that the earnings equation estimated for the 1905 cohort was also valid for the cohorts born before 1905. The second one amounts to retropolate the cohort-trends in the earnings structure estimated after 1905. The third one was to drop all individuals whose father was born before 1905. The fourth one is to exclude all child's cohorts in which a significant share of the fathers was born before 1905.

All four procedures produce very similar results, as shown in the table below, in the case of the piece-wise linear trends specification.

Table 12: IGE estimates for various treatments of fathers born before 1905

	(1)	(2)	(3)
Child's income variable	own	own	family
Father's income variable	own	family	family
w/ trends extrapolation	.3321682 (.0297205)	.366469 (.0340147)	.2925363 (.0344422)
N	2265	2265	2106
no trends extrapolation	.3368869 (.0299921)	.3738406 (.0344416)	.2969808 (.0348861)
N	2265	2265	2106
excl. fathers born before 1905	.343647 (.0317294)	.3854528 (.036683)	.3001721 (.0373722)
N	2056	2056	1901
only children born after 1940	.3290154 (.030755)	.365947 (.0353181)	.2922087 (.0359411)
N	2092	2092	1937

Notes: Source : authors' computations from SSM data.