Forecasting Income Inequality with Demographic Projections

Chong, Terence Tai Leung and Ka, Yiu Tung

The Chinese University of Hong Kong

1 December 2019

Forecasting Income Inequality with Demographic Projections

Terence Tai Leung Chong and Yiu Tung Ka

Department of Economics, The Chinese University of Hong Kong

17/8/2019

Abstract:

This paper provides a first attempt in the literature to forecast the future evolution of income inequality with the demographic projections. The contribution of this paper is twofold. First, we establish a framework to quantify and analyze the effects of population ageing and the secular upward trend in educational attainment on income inequality. Second, we modify the human capital model and perform microsimulations to forecast a list of standard measures of income inequality of Hong Kong for the coming years of 2021, 2026 and 2031 based on the projected changes in the demographic structure of Hong Kong's working population. The pseudo out-of-sample forecasts are reasonably close to the corresponding realized values. Our true out-of-sample forecasts suggest that income disparity will be alleviated in the next 15 years, as a result of the increasingly equal spread of level of schooling across the workforce.

JEL Classification: D31; I24; J11

Keywords: Income Inequality; Demographic Projections; Population Ageing.

Corresponding Author: Terence Tai Leung Chong, email: chong2064@cuhk.edu.hk

1. Introduction

The economy and the demographic structure of Hong Kong have both experienced dynamic and dramatic changes over the past few decades. Back in the 1950s after World War II, thanks to the skills, capital and labor brought by foreign firms and refugees from Mainland China, Hong Kong transformed from a territory specializing in entrepôt trade into a labor-intensive manufacturing hub.[1] The fast-paced industrialization during the 1960s and 1970s gained Hong Kong the reputation as one of the Four Asian Tigers[2] in recognition of its exceptionally high economic growth and success in export-oriented light industries such as textile and clothing, toys and electronics. Benefiting from the tide of globalization and China's Open Door Policy in the 1980s, Hong Kong's economy underwent another major structural change. Factories relocated away from Hong Kong to take advantage of the lower labor and land costs at the north of the border (Wong, 1991). Meanwhile, there was a substantial expansion of the service sector. A huge number of high value-added jobs were created during the last three decades of the 20th century along with the boom of network technology and the increasing interaction between China and the Western world. In this "golden era" of unprecedented growth, a range of socioeconomic measures of living standard, including life expectancy, literacy rate and women's rights in the workplace, improved significantly.

Despite having a small population, Hong Kong is one of the advanced open economies globally in terms of real GDP per capita. However, in order to assess the overall living standard of a region, it is also essential to consider how economic resources are distributed to different strata of the society. In recent years, voices have emerged in many developed countries, claiming that "the rich are getting richer while the poor are getting poorer", suggesting that there might be a trade-off between economic growth and distributive

[1] The Cold War embargoes imposed by the United States and the United Nations on China in 1949 and 1951 also facilitated the industrialization of Hong Kong during the 1950s. See Schenk (2008) for details.

[2] The Four Asian Tigers, also known as Four Asian Dragons or Four Little Dragons, refer to the economies of Hong Kong, Singapore, South Korea and Taiwan. These small open East Asian economies all achieved an annual growth rate exceeding 7% in the 1970s.

equality. In response to rising public concerns about this phenomenon, a growing number of studies have investigated both the theoretical and the empirical relationship between inequality and growth.[3]

Although economists in general agree that inequality has been worsening in most developed countries since 1980 (Goldin and Katz, 2008; Piketty, 2014), there are many competing theories on its potential causes. Depending on the standard of measurement used, economic inequality can be driven by many underlying forces beyond the changes in the distribution of income among individuals. Thus, to design effective policies to pre-empt the possibility of a widening income gap in the future, we should start our analysis with three fundamental questions: (1) What aspects of inequality are we concerned with, (2) what are the corresponding determinants, and (3) how do these determinants evolve over time? While the first two questions have been well addressed in previous literature, there are only a limited number of studies attempting to forecast how the income distribution would look like in future years based on the projected trends of hypothesized determinants of inequality. After all, if foreseeable changes in the socioeconomic environment will reduce income inequality naturally in the future, then all we need to do is to simply wait for such events to occur. On the contrary, if we anticipate that ongoing demographic changes will inevitably widen the income gap in the future, then policymakers may have to think about how to ease the situation and come up with adequate measures at the current stage to prevent further deterioration in income disparity.

Given the strategic role of inequality projections in policy-making, once we have chosen the relevant measurement of inequality, we can proceed to the next step and figure out the factors which can give rise to economic inequality in future periods. Previous research suggests that both population ageing and changes in educational composition can affect labor market outcomes, which in turn alter the income distribution among individuals and households (De Gregorio and Lee, 2002; Zhong, 2011; Yang and Qiu, 2016). In other words, as long as we can project the trends of certain demographic variables which are believed to be influential and robust in explaining income inequality in an *ex post* manner, we might be able to construct a model to give precise *ex ante* forecasts on the future

[3] For a survey of some of the literature, see Kuznets (1955), Persson and Tabellini (1994), Aghion, Caroli and García-Peñalosa (1999), Panizza (2002), Banerjee and Duflo (2003) and Murphy and Topel (2016).

evolution of income distribution. It is therefore crucial to examine these relevant demographic traits in detail and investigate their relationships with income inequality.

Hong Kong is a world-class metropolis renowned for its economic freedom and competitive labor force. However, it is also characterized by a high Gini coefficient and serious population ageing compared with other economies in the Asia-Pacific region.[4] These special features make Hong Kong an interesting case to study. In addition, given that the unique demographic structure of Hong Kong caused by waves of mass migrations in the second half of the 20th century, the retirement of post-war baby boomers is expected to exert a significant impact on the composition of Hong Kong's workforce and its social security system.[5] The ever-rising elderly dependency ratio and the increasing fraction of young workforce receiving tertiary education in Hong Kong, among other demographic trends, are also typical like the socioeconomic environments in developed economies, enabling us to study the respective implications of these demographic changes on the future evolution of income distribution.

Using the 5%-sample raw data from Hong Kong Population Census and By-census, this paper estimates and analyzes the effects of population ageing and the proliferation of higher education institutes on income inequality among the working population in Hong Kong. On top of that, we modify the Mincerian earnings function and perform microsimulations to make forecasts of a list of standard measures of income inequality, including Gini coefficient, Theil index, Atkinson index, the variance of the log of income and the 90th percentile to 10th percentile income ratio for the years of 2021, 2026 and 2031 based on the projected demographic changes in the workforce. Pseudo out-of-sample forecasts generated for the years 2006, 2011 and 2016 are reasonably close to their corresponding realized values. This suggests that the evolution of the demographic structure, specifically the changes in the distribution of age and educational attainment of the workforce, possess predictive power for the changes in the overall distribution of labor income. Meanwhile, the true out-of-sample forecasts imply that income inequality will be alleviated gradually in the next 15 years, mainly driven by a less dispersed distribution of schooling(education) among the workforce.

[4] According to the Central Intelligence Agency World Factbook (2017), Hong Kong ranked the 9th out of the 150 economies studied in terms of family income Gini index.

[5] See Wong (2017) for details.

The rest of this paper is organized as follows. Section 2 provides an overview of literature concerning income inequality. Some leading hypotheses about the causes of economic disparity, as well as other studies which focus on the income distribution in Hong Kong will be reviewed. Section 3 presents the baseline human capital model, which serves as a benchmark for analyzing the effects of ageing and the secular upward trend in education on earnings disparity, measured by the variance of the log of main employment income. Section 4 describes the Hong Kong Population Census and By-census datasets and the extrapolation procedure adopted. A list of commonly used income inequality indices from 1981 to 2016 are also reported to outline some stylized facts of Hong Kong's income distribution. Section 5 reports and discusses the estimation and forecasting results obtained from the baseline model. Section 6 attempts to enrich the Mincerian earnings function to simulate the whole income distribution and generate income inequality forecasts for the years of 2021, 2026 and 2031. Section 7 concludes and offers directions for future research.

2. Literature Review

There is extensive literature documenting the trends of income inequality in different countries over different time horizons. A large body of research are also dedicated to testing and empirically verifying a great variety of hypotheses concerning the causes and consequences of economic inequality. In contrast, only a handful of studies have attempted to advance to the next level by commenting on how the distribution of income will evolve in future periods based on the established channels. We thereby briefly summarize the development of this field and review some of the relevant studies which focus on income disparity in Hong Kong to support the subsequent forecast.

The oldest hypothesis of income inequality can probably be traced back to those that relate the distribution of income to the distribution of individuals' "abilities" (Staehle, 1943; Mincer, 1958). Moore (1911) explicitly assumes that "industrial ability—general sagacity and energy—is distributed according to the normal or Gaussian law". Consequently, the difference in wages among people was rationalized by the consensus that different people have different abilities. While Schumpeter (1916) was once celebrated for this approach,[6] the logical weakness of the hypothesis soon led to its refutation. In particular, Pigou (1932) argued that empirically income usually follows the Pareto's law of income distribution[7] instead of the Gaussian distribution. A paradox thus follows: How can the normal distribution of abilities be reconciled with the sharply skewed distribution of income? The generally accepted answer to this question provided by Pigou himself is that there are other omitted factors which intervene and distort the relation between ability and income. Since then, plenty of studies have been motivated to develop the model specifications and to provide alternative hypotheses for the rise of income inequality.

Among all the theories that have been proposed,[8] skill-biased technological change (SBTC) is one of the most widely recognized and citied explanations for the worsening income inequality over the past few decades. SBTC refers to a shift in production technology that favors skilled over unskilled labor by increasing the relative productivity of the former over the latter, which in turn raises the relative demand for skilled labor and induces a rise in the skill premium (Violante, 2008). Initiated by Schultz (1975) and popularized by Johnson (1997), a tide of research had been triggered to debate on the role, impacts and duration of technological changes. For example, Acemoglu (1998) documented the increase in supply of skilled labor from 1970 onwards and divided its impacts into short run and long run. In the short run, he proposed that the abundance of skilled labor leads to a decrease in skill premium through a substitution effect (a downward movement along the demand curve for skilled labor). However, in the long run, since the increase in skilled labor in the workforce facilitates the advancement of skill-complementary technologies, the skill premium soars as a result of the disproportionate increase in productivity of the skilled labor over the unskilled (skilled labor demand curve shifts outwards), which in turn causes inequality in wages. In a follow-up paper, Acemoglu (2002) further expanded his theory by formalizing his insight that technological change

[6] Schumpeter wrote in 1916, "The great idea of investigating the relationship between wage differences and differences in ability opens a vast perspective. The new trail is steep and stony, but it must be followed."

[7] The Pareto distribution is also known and referred to as "80 – 20" rule, meaning that approximately 80% of the wealth or income of a society is held or earned by 20% of the population.

[8] Some examples include discriminations (Becker, 1971; Darity and Mason, 1998), financial and capital market imperfections (Banerjee and Newman, 1993; Galor and Zeira, 1993; Clarke, Xu and Zou, 2006), international trade (Burtless, 1995; Furusawa and Konishi, 2016) and superstar effect (Rosen, 1981).

was skill-biased in the 20th century in contrast to skill-replacing in the 19th century. His argument is supported by the fact that both the real wages of high-skilled labor and the unemployment rate of unskilled labor have been rising simultaneously since the 1970s.

Apart from fostering SBTC, human capital variables can also affect income inequality via other channels. One recent example is given by Murphy and Topel (2016). By introducing the concept of “equilibrium inequality”, they argued that human capital investment responses to skill prices at both the extensive margin and the intensive margin. While the former means that the increase return of education will induce more people to attend college and thus produce more skilled labor, akin to how the output of an industry is expanded by the entry of new firms; the latter refers to the situation in which skilled workers acquire more human capital and apply them more intensively in the labor market sector when the prices of skills increase, similar to an expansion of output by intramarginal firms when rising market demand raises prices in a competitive market. Hence, if investment and utilization of human capital at the intensive margin are more responsive to the rise in skill prices than that at the extensive margin, the latter of which would lead to the creation of skilled workers which means that wage inequality will be exacerbated by the polarization of skills.

In addition, skill-neutral structural reforms can also give rise to income inequality. For instance, some industries and positions require job candidates to obtain certain licenses or professional qualifications before they are eligible to apply. For management positions, it is often necessary to acquire a high level of human capital, usually by means of getting a college degree, or nowadays, a postgraduate one. With the existence of such signaling and screening devices, structural reforms would induce resource reallocation and naturally change the relative demand for labor across different industries or positions. However, labor supply might not be able to respond and adjust accordingly because retraining workers and human capital investment are costly processes (Blundell et al., 1999). In that case, income distribution could be altered. Furthermore, as the opportunity cost of human capital investment varies among individuals (Becker, 1967; Ben-Porath, 1967), it is also believed that the ease of gaining access to training and education, as well as the distribution of their returns, would affect the level of income inequality.

As one of the pioneers in applying the human capital approach to analyze this issue, Chiswick and Mincer (1972) used the United States Current Population Survey (CPS) data to explain and predict earnings inequality, measured by the variance of the log of individuals' personal income by estimating a relative inequality function derived from the human capital earnings function. They found that their model achieves a high explanatory power in the analysis of annual income inequality during the postwar period between 1949 and 1969, with the average error of prediction being less than 2% and individual errors never exceeding 5%. Their research concluded that income inequality was mainly caused by the dispersion of weeks of employment resulting from business cycle fluctuations, followed by changes in the distributions of schooling and age.

In Hong Kong, there are also some studies which seek to analyze the changing pattern of income distribution over the years. Chow and Papanek (1981) found that income disparity in Hong Kong did not deteriorate significantly during the period between 1957 and 1976. By investigating the data from the mid-1970s to the mid-1980s, a research by Turner et al. (1991) reported a convergence of income between different classes of employees, which narrowed the income differentials in Hong Kong. Chau (1994) also proposed that the expansion of public education system substantially promoted the upward mobility of low-income households since the 1970s. All these findings suggest that there was no trade-off between growth and equality in Hong Kong before its handover in 1997.

On the other hand, some research works look into the effects of Hong Kong's economic reform on income distribution. Hsia and Chow (1978) argued that in its early stage of economic development, rapid industrialization in Hong Kong contributed significantly to both the rise in its living standard and the decline in its household income inequality. In contrast, by examining the sectoral shifts of 25 broad industries in Hong Kong, Suen (1995) noted that the changes in industrial composition alone accounted for about 70% of the increase in income dispersion during the period between 1976 and 1991. Lam and Liu (1998) also used census data from 1981 and 1991 to show that shifts in the distribution of heterogeneous population groups caused by changes of the immigration policy could give rise to income inequality.

In the latest thematic report published by the Census and Statistics Department (2017), it is proposed that both the acceleration of population ageing and the shrinking household size over time have contributed to an increase in household income inequality. Regarding the widening income gap in post-handover Hong Kong, Lui (2013) wrote an entire book to discuss several hypotheses. By categorizing the changes in income inequality by industry and by occupation, he founded that economic restructuring only accounted for around one-eighth of the change in income dispersion during the post-handover period, refuting the argument that economic transformation itself constitutes a major source of post-handover income inequality. Instead, he proposed an alternative hypothesis, claiming that the expansion of higher education was the true primary factor leading to rising inequality.[9]

Recently, Wong (2017) studied issues related to income inequality in Hong Kong. He found that age, sex, education, marital status and immigration status were all relevant factors in explaining the changing pattern of individual income inequality between 1981 and 2011, with education being the most influential determinant. However, in contrast to Lui (2013), Wong (2017) argued that failing to adopt an appropriate population policy, especially the indecision to provide sufficient post-secondary education opportunities under the context of increasing returns to education in the 1990s, fueled the increase in Hong Kong's post-handover income dispersion. He also suggested that the lack of investment in human capital owing to financial market imperfections may have resulted in an adverse effect on intergenerational social mobility.

Existing literature concerning income inequality in Hong Kong has provided ample empirical evidence, with diversified *ex post* analyses to explain the historical trends of income dispersion. However, the lack of research conducted from an *ex ante* perspective, i.e. to make predictions or forecasts of future income inequality based on the projected trends of human capital variables, is puzzling.[10] We therefore try to fill this gap by

[9] Lui (2013) argues that higher education levels like college degrees and high-skilled occupations like managers and administrators usually contribute to a higher intra-group income inequality. With a larger share of people belonging to these two groups, the overall income inequality will increase.

[10] One of the very few attempts to forecast income inequality is made by Gindelsky (2016). Using historical data from the CPS, she performed forecasts for eight measures of income inequality in the United States

following Chiswick and Mincer's (1972) approach and performing microsimulations to forecast the future evolution of income inequality in Hong Kong.

3. Baseline Human Capital Model

The human capital approach interprets schooling and post-school training as a form of investment which augments the productivity of workers. As specified in Chiswick and Mincer (1972), the relation between potential earnings[11] and investment in human capital for the ith person in year j can be expressed as

$$E_{ij} = E_{i0} + \sum_{t=1}^{j-1} r_{it} C_{it}$$

where the potential earnings (E_{ij}) are decomposed into two parts: (i) the "original" endowment labor earnings (E_{i0}), and (ii) the sum of returns on previous human capital investment (C_{it}). In equation (1), r_{it} denotes individual i's average rate of return to human capital investment in the tth year.

Assuming that original endowment labor earnings are constant across years and individuals, we have $E_{i0} = E_0$. Moreover, human capital investment is assumed to be a fraction of one's potential earnings, i.e., $C_{it} = k_{it} E_{it}$, where $k_{it} \in [0, 1]$. Then, we can rewrite equation (1) as[12]

---

and predicted that while the top 1% share of income will rise slowly for households, the top 0.1% income share and inequality within the top 1% would fall over the period between 2015 and 2017.

[11] Potential earnings here refer to the main employment income that a person with a certain level of human capital can potentially earn, before subtracting any contemporaneous costs in training.

[12] For the ith person, when $j = 1$,

$$E_{i1} = E_0$$

When $j = 2$,

$$E_{i2} = E_0 + r_{i1} k_{i1} E_{i1} = E_0 (1 + r_{i1} k_{i1})$$

When $j = 3$,

$$E_{i3} = E_0 + r_{i1} k_{i1} E_{i1} + r_{i2} k_{i2} E_{i2} = E_0 (1 + r_{i1} k_{i1}) + r_{i2} k_{i2} E_0 (1 + r_{i1} k_{i1}) = E_0 (1 + r_{i1} k_{i1})(1 + r_{i2} k_{i2})$$

and so on.

$$E_{ij} = E_0 + \sum_{t=1}^{j-1} r_{it} k_{it} E_{it} = E_0 \prod_{t=1}^{j-1} (1 + r_{it} k_{it})$$

Since $r_{it} k_{it}$ is small, by taking the natural logarithm on both sides of equation (2), we have the approximation

$$\ln(E_{ij}) = \ln(E_0) + \sum_{t=1}^{j-1} r_{it} k_{it}$$

The $j-1$ periods of human capital investment can be further divided into S years of schooling and $j-S-1$ years of post-school training. For estimation purpose, it is also assumed that there are no part-time students, which means that the direct cost of formal schooling is the entirety of the student's income he could earn with his human capital level. Thus, $k_{it} = 1$ and $C_{it} = E_{it}$ for schooling years. Moreover, the return to post-school training is assumed to be constant, i.e. $r_{it} = r_i^T$ for all $t > S$. When these assumptions are incorporated into equation (3), we have

$$\ln(E_{ij}) = \ln(E_0) + r_i^S S_i + r_i^T \sum_{t=S+1}^{j-S-1} k_{it}$$

, where r_i^S denotes the average rate of return to schooling for individual i.

Although the interpretation of equation (4) is straightforward, potential earnings E_{ij} are not directly observable from the data. Hence, practical estimation requires the use of actual earnings (Y_{ij}) as the dependent variable. By definition, Y_{ij} is related to E_{ij} as

$$Y_{ij} = (1 - k_{ij}) E_{ij}$$

, which is equivalent to

$$\ln(Y_{ij}) = \ln(1 - k_{ij}) + \ln(E_{ij})$$

Nonetheless, k_{it} could not be directly observable either. It is therefore necessary to specify an explicit functional form for k_{it} to make the estimation feasible. Following the convention, k_{it} is assumed to be a linear decreasing function of years of post-school training:[13]

$$k_{it} = k_0 \left(1 - \frac{T_i}{T_i^*}\right) \qquad \text{for } t > S$$

, where k_0, T_i and T_i^* are the initial investment ratio, years of work experience and the last year of positive post-school training for individual i respectively. Now, the term $r_i^T \sum_{t=S+1}^{j-S-1} k_{it}$ in equation (4) is a parabolic function of the number of years of post-school training, and its maximum is reached when $T_i = T_i^*$ (i.e. $k_{it} = 0$).

Substituting equation (6) and (7) back to equation (4), we get[14]

$$\ln(Y_{ij}) = \ln(1 - k_{ij}) + \ln(E_0) + r_i^S S_i + r_i^T k_0 T_i - \frac{r_i^T k_0}{2T_i^*} T_i^2 \tag{8}$$

, and the term $\ln(1 - k_{ij})$ can be further evaluated by a second-order Taylor series expansion around T_i^*:

$$\ln(1 - k_{ij}) = -k_0 \left(1 + \frac{k_0}{2}\right) + \frac{k_0 T_i}{T_i^*}(1 + k_0) - \frac{k_0^2 T_i^2}{2T_i^{*2}} \tag{9}$$

Plugging equation (9) into equation (8) results in

---

[13] There are some reasons to rationalize the assumption that k_{it} decreases over one's career. First, as post-school training increases, wages received by the worker are expected to increase alongside his productivity. Hence, the opportunity cost of time invested in post-school training increases with additional experience, reducing the profitability of further investments. Second, the net present value of post-school human capital investment will be higher the earlier it is undertaken, vice versa.

[14] Converting to continuous time,

$$\sum_{t=s+1}^{j-S-1} k_{it} = \sum_{t=0}^{T_i} k_{it} \approx \int_0^{T_i} k_{it}\, dT_i = k_0 T_i - \frac{k_0}{2T^*} T_i^2$$

$$\ln(Y_{ij}) = \left[\ln(E_0) - k_0\left(1 + \frac{k_0}{2}\right)\right] + r_i^S S_i + \left[r_i^T k_0 + \frac{k_0}{T_i^*}(1 + k_0)\right] T_i - \left[\frac{r_i^T k_0}{2T_i^*} + \frac{k_0^2}{2T_i^{*2}}\right] T_i^2$$

$$= \beta_0 + \beta_{1i} S_i + \beta_{2i} T_i + \beta_{3i} T_i^2$$

Since data on workers' post-school training are not available, we follow the convention and assume that labor market experience equals to age minus years of schooling minus 5 ($T = A - S - 5$). This assumption enables the study of income distribution by age group rather than by experience group. Further simplifications are made with this standard earnings function before we derive the relative inequality function. First, following Chiswick and Mincer (1972), the squared term of experience is deleted because its inclusion would be computationally cumbersome, yet the additional explanatory power is not likely to be economically significant.[15] Secondly, it is assumed that the returns to schooling and experience are both random variables that are uncorrelated with each other and vary across individuals. This implies that the variances of the coefficients of S and T in equation (10) are both strictly greater than zero. Applying these modifications to equation (10), we obtain

$$\ln(Y_{ij}) = \beta_0 + \beta_{1i} S_i + \beta_{2i}(A_i - S_i - 5) + \varepsilon_i \tag{11}$$

, where ε_i is a residual term which represents the combined effect of other omitted variables and measurement errors.

Taking the variance operator on both sides of equation (11) results in

$$\begin{aligned}\sigma^2[\ln(Y)] = {} & [(\beta_1 - \beta_2)^2 + \sigma^2(\beta_1) + \sigma^2(\beta_2)]\sigma^2(S) + [{\beta_2}^2 + \sigma^2(\beta_2)]\sigma^2(A) \\ & + [2\beta_2(\beta_1 - \beta_2) - \sigma^2(\beta_2)]R_{AS}\sigma(A)\sigma(S) + \sigma^2(\beta_1){\mu_S}^2 \\ & + \sigma^2(\beta_2)(\mu_A - \mu_S - 5)^2 \\ & + \sigma^2(\varepsilon) \end{aligned} \tag{12}$$

[15] A caveat of imposing this assumption is that the slope coefficient of experience, i.e. β_2, would be biased downward. See Chiswick and Mincer (1972) for detailed elaboration.

, where $\sigma^2(.)$ denotes the variance of the corresponding variable, R_{AS} denotes the correlation between age and schooling, and μ_A and μ_S represent the mean level of age and schooling respectively.

In equation (12), the variance of log-earnings ($\sigma^2[\ln(Y)]$) is expressed in terms of both the level and the variance of age and schooling, as well as the correlation between them. Indeed, one advantage of measuring income inequality in terms of $\sigma^2[\ln(Y)]$ is that it can be decomposed into different components and their respective effects can be analyzed separately. By taking the partial derivatives of the relative inequality function with respect to the independent variables, Table 1 lists the effect of a unit change in each demographic determinant on the dispersion of income.

Table 1: Partial effects of the independent variables on income inequality

$$\frac{\partial\sigma^2[\ln(Y)]}{\partial\sigma(S)} = 2[(\beta_1-\beta_2)^2+\sigma^2(\beta_1)+\sigma^2(\beta_2)]\sigma(S) + [2\beta_2(\beta_1-\beta_2)-\sigma^2(\beta_2)]R_{AS}\sigma(A)$$

$$\frac{\partial\sigma^2[\ln(Y)]}{\partial\sigma(A)} = 2[{\beta_2}^2+\sigma^2(\beta_2)]\sigma(A) + [2\beta_2(\beta_1-\beta_2)-\sigma^2(\beta_2)]R_{AS}\sigma(S)$$

$$\frac{\partial\sigma^2[\ln(Y)]}{\partial\mu_S} = 2[\sigma^2(\beta_1)-\sigma^2(\beta_2)]\mu_S - 2\sigma^2(\beta_2)(\mu_A-5)$$

$$\frac{\partial\sigma^2[\ln(Y)]}{\partial\mu_A} = 2(\mu_A-\mu_S-5)\sigma^2(\beta_2)$$

$$\frac{\partial\sigma^2[\ln(Y)]}{\partial R_{AS}} = [2\beta_2(\beta_1-\beta_2)-\sigma^2(\beta_2)]\sigma(A)\sigma(S)$$

The first determinant of income inequality is the variance in schooling among the workforce. Since schooling is a major determinant of one's earnings, it is certain that income inequality also depends critically on the distribution of education. The quantitative effect of a unit change in the dispersion of education on income inequality is represented by the partial derivative $\frac{\partial\sigma^2[\ln(Y)]}{\partial\sigma(S)}$, as indicated by the first equation in Table 1. The magnitude of this effect depends on several factors. Firstly, when the returns to schooling (β_1) rises, $\frac{\partial\sigma^2[\ln(Y)]}{\partial\sigma(S)}$ will increase because education becomes more influential in determining one's personal income, which in turn enlarges the effect of the distribution of schooling on the distribution of income. Secondly, the higher the β_2, the smaller the $\frac{\partial\sigma^2[\ln(Y)]}{\partial\sigma(S)}$. This is because β_2 reflects the average effect of additional labor market experience, and this experience is proxied by $A - S - 5$. In other words, given that two workers of the same age, the worker who received one fewer year of schooling is assumed to have one additional year of work experience. Hence, when the workforce becomes more diverse in terms of educational attainment, the increase in β_2 can mitigate its effect on income inequality by raising the income of those with less schooling (but with more work experience). Thirdly, an increase in $\sigma(A)$ will also reduce the effect of variance of schooling on income inequality. The rationale behind is that age and schooling are usually negatively correlated (i.e. $R_{AS} < 0$), which implies that young (less experienced) workers are in general more educated than old (more experienced) workers. In this case, the overall wage differential between young workers and old workers is lower. Thus, with a more extreme age distribution among the workforce, the dispersion of schooling would favor the less experienced young workers, so the impact of $\sigma(S)$ on $\sigma^2[\ln(Y)]$ would be lower. Likewise, if R_{AS} decreases, meaning that age and schooling among workers become more negatively correlated, then $\frac{\partial\sigma^2[\ln(Y)]}{\partial\sigma(S)}$ will also decrease.

The second factor of income inequality is the variance of age. As a proxy for labor market experience, age reflects one's productivity to some extent and plays a role in determining one's income. Therefore, changes in the age composition of workforce can theoretically result in a change in income inequality. The second equation in Table 1 specifies the impact of a unit increase in the dispersion of age on the dispersion of income.

Again, the magnitude of $\frac{\partial\sigma^2[\ln(Y)]}{\partial\sigma(A)}$ depends on various factors in which their mechanisms and interpretations are similar to those mentioned for $\frac{\partial\sigma^2[\ln(Y)]}{\partial\sigma(S)}$ above.

Apart from their variances, the levels of schooling and age also exert some effects on the distribution of income. Although it is traditionally believed that income inequality can be alleviated by promoting education, the model derived here suggests otherwise. The increase in average education level among workers would exacerbate income disparity since jobs requiring a high education level are usually more heterogeneous in nature, resulting in higher dispersion in income (Ruiz-Tagle, 2007; Lui, 2013). This effect is captured by the interaction term $\sigma^2(\beta_1)\mu_S{}^2$ in equation (12). For example, suppose there are two college graduates with one majoring in finance and the other majoring in history. Although both of them have received 16 years of formal schooling, the income disparity between them can be huge. The same argument can be applied to two college graduates majoring in the same subject but graduating from different universities with different rankings. In contrast, the degree of heterogeneity in jobs requiring only primary and secondary education is smaller, therefore a society with lower average education may have a lower disparity in income.

Age affects income inequality in a similar fashion. In general, as the average age of workers goes up, a relatively larger proportion of them will fall into the old-age group where income inequality is inherently larger. This is because when workers become more experienced as they age, some of them will be promoted to senior executive or management positions with considerable remuneration, while the others may not be able to enjoy a substantial increase in salary over the course of their career. On the contrary, for young workers with less work experience, the market for their skills is usually more competitive, leading to a smaller variance in their wages. Therefore, earnings inequality among experienced workers is generally higher, and this effect is reflected by the interaction term $\sigma^2(\beta_2)(\mu_A - \mu_S - 5)^2$ in equation (12). In other words, population ageing may be accompanied by a widening income distribution.

Last but not least, the intercorrelation between age and schooling affects income inequality in a favorable way. In most developed countries, both the quality and the quantity of education have risen over time. This has given rise to a phenomenon called education inflation, with young workers nowadays receiving more formal schooling than their elder counterparts in general. As a result of the secular upward trend in education, the correlation between age and schooling is usually negative at any given point in time and is decreasing across generations, leading to a narrowing of earnings inequality. This inequality-easing channel is captured by the term $[2\beta_2(\beta_1 - \beta_2) - \sigma^2(\beta_2)]R_{AS}\sigma(A)\sigma(S)$ in equation (12).

In summary, the human capital approach views schooling and labor market experience as two fundamental drivers of labor productivity, which in turn determines one's labor market earnings. As a result, changes in the demographic structure, particularly the distribution of age and education, together with the returns to them, as well as the interaction between them are the keys to understanding the evolution of income inequality among the workforce of a society. In order to forecast the dispersion of income, it is essential to project how these variables are going to change in future.

4. Data and Extrapolation

This paper studies income inequality in Hong Kong at the individual level. Although the lack of an annual survey has hindered inequality forecasting and other related time series analyses, we make use of multiple sets of Population Census and By-census data to estimate the quantitative effects of demographic changes, specifically population ageing and education inflation among the workforce, on income disparity in Hong Kong over time. Based on the available cross-sectional data, we project and simulate the demographic structure for the years of 2021, 2026 and 2031 by extrapolation, from which we are able to predict future changes of the variance of log-income. We also generate pseudo out-of-sample forecasts and true out-of-sample forecasts for other standard measures of income inequality.

The Hong Kong Population Census has been conducted every ten years since 1961, covering all residents physically present in Hong Kong during the survey period; whereas the Hong Kong Population By-census is conducted halfway between two consecutive censuses, covering one-tenth of all quarters in Hong Kong and all households and individuals therein. After publishing the official findings on the demographic and socioeconomic characteristics of the population, the government would extract a random subsample from the full Census or By-census dataset to construct microdata files for scholars to conduct academic research. These unpublished census microdata files are compiled and processed by the Census and Statistics Department, thus the data source is regarded as highly reliable. Nevertheless, caution is required in comparing the results across census cohorts as statistical standards and the definition of some variables may have already changed over the years.[16] Moreover, to protect respondent confidentiality, the income data are censored. The monthly main employment income is reported in the data as 99,998 for those who earned more than this amount in 1981, 1986 and 1991. This cap is relaxed to HK$150,000 from 1996 onwards. While this treatment would have essentially led to an underestimation of income inequality, the bias was expected to be trivial since the proportion of workers with such a high main employment income was very small.[17]

This study applies the 5%-sample raw data from 1981 to 2016 to analyze income inequality among individuals aged between 25 and 64. Several conditions are set to enhance the accuracy of the estimation. Firstly, all income data are adjusted using the respective composite consumer price index (with 1981 as the base year) to facilitate comparison across years. Secondly, all foreign domestic helpers are excluded because their salaries are constrained by the law and usually not related to their education and experience. Thirdly, since the variance of the log of monthly main employment income is used as the measurement of income inequality in the following analysis and the logarithm of zero is undefined in mathematics, respondents who are not earning a positive main employment income are excluded. Finally, for the remaining observations, workers with main employment income below the 3rd percentile of the whole distribution are also excluded because their exceptionally low salaries are probably resulted from unreasonably short

[16] For example, since 2001, the "resident population" approach has been adopted to conduct the censuses. In 1996, the Population By-census was carried out under the de jure enumeration approach whereas the de facto enumeration approach was used in earlier censuses and by-censuses.

[17] Only 0.2% of the sampled workers reported a monthly main employment income equal to or exceeding HK$ 99,998 in 1991. The share of such high income samples with a monthly main employment income not less than HK$ 150,000 increased to 0.9% in 2016.

working hours. Table 2 presents various measures of income inequality over the sample period.

Because different income inequality measures show a different degree of sensitivity to changes in different parts of the income distribution, they may show different trends over the years. However, from the increasing pattern shown by the Gini coefficient and the variance of log-income, we can still safely conclude that income inequality has been rising in general among the working population in Hong Kong, even though the increment over the post-handover period is not substantial. The Theil index is an inequality measure which is more responsive to changes at the top of the distribution. As both the Theil index and the P90/P50 ratio have exhibited rising trends in recent years, there is a possibility that the income share attributed to top income earners has increased. In contrast, the Atkinson index is more sensitive to the bottom part of the income distribution. Since the Statutory Minimum Wage came into force in 2011, both the Atkinson index and the P50/P10 ratio have recorded a decline in the latest census.

Table 2: Income inequality among the working population in Hong Kong

	1981	1986	1991	1996	2001	2006	2011	2016
Gini coefficient	0.394	0.408	0.409	0.423	0.425	0.430	0.432	0.439
Theil index	0.362	0.366	0.347	0.361	0.344	0.351	0.348	0.362
Atkinson index	0.230	0.243	0.241	0.254	0.256	0.261	0.262	0.256
Variance of log-income	0.429	0.441	0.453	0.471	0.497	0.509	0.514	0.522
P90/P50	2.353	2.362	2.500	3.000	2.917	2.885	2.878	2.941
P50/P10	1.890	2.000	2.000	1.818	2.000	2.128	1.984	1.890

To make forecasts of future income inequality, it is necessary to project the future demographic structure and recognize how the composition of the workforce is going to change. Extrapolation is therefore needed to make the forecasting task feasible. The 2016 Population By-census 5%-sample dataset is used to serve as a base for this extrapolation.

First, the population is divided into several categories according to gender, age and educational group. Since the workforce participation rate for never-married women is significantly higher than those who are ever-married (including those currently married, widowed and divorced/separated), females are further classified by their marital status. After such classification (e.g. males aged 25 – 29 with college degree, never-married females aged 40 – 44 with upper secondary education), We calculate the workforce participation rate for each group and the marriage rate for each female age-education group.[18] The workforce participation rate is assumed to be constant over time, while the female marriage rate is projected based on an extrapolation of past trends. For individuals aged 30 or above, educational attainments are also assumed to be constant. The projected working population in each group is then extrapolated based on the change in the education proportions between 2011 and 2016. For instance, females aged 10 – 14 in 2016 will become 25 – 29 years old in 2031. At that time, some of them will have finished a college degree and gotten married. The percentage of college graduates among all females aged 25 – 29 in 2031 is extrapolated by the corresponding proportion in 2016, plus three times the change in this proportion between 2011 and 2016.[19] Multiplying this proportion by the projected population (i.e. the number of females aged 10 – 14 in 2016) and the corresponding marriage rate and workforce participation rate would result in the projected working population who are ever-married females aged 25 – 29 with a college degree in 2031. Finally, the whole composition of the workforce in 2031 can simply be obtained by applying the same projection strategy to all groups and summing them up.

In addition to simulating the composition of the working population for 2021, 2026 and 2031, the same extrapolation method can also be adopted using the 2001 Population Census 5%-sample dataset as the base for generating pseudo out-of-sample forecasts to check the usefulness of the model. The detailed projection figures of the workforce composition can be found in the Appendix.

---

[18] The workforce participation rate is defined as the ratio of working population to the total population for the corresponding age-education group. Similarly, the marriage rate here is defined as the ratio of ever-married females to total females for the corresponding age-education group.

[19] The change in the proportion between 2011 and 2016 is multiplied by 3 because 15 years will have passed in 2031. The extrapolated proportions are bounded by zero.

5. Estimation Results

The Mincerian earnings function and the income inequality function specified in Section 3 are used to estimate the effects of population ageing and the expansion of education system on income inequality quantitatively.[20] Using the sample of working population aged 25 – 64 from the 2016 Hong Kong Population By-census 5%-sample dataset, the estimated result for equation (11) is reported below:

$$\ln y_i = \underset{(0.0093)}{7.931} + \underset{(0.0005)}{0.129}S_i + \underset{(0.0001)}{0.014}T_i + \varepsilon_i$$

$$N = 145{,}400 \qquad R^2 = 0.334$$

Since β_{1i} and β_{2i} in equation (11) are random variables, which should vary across individuals, the sample is further divided into groups by gender and industry for the estimation of the variance of the two coefficients.[21] Assuming the returns to schooling and work experience differ for people of different gender and for those who work in different industries, it is estimated that $\widehat{\sigma^2(\beta_1)} = 0.00099$ and $\widehat{\sigma^2(\beta_2)} = 0.00003$. With these statistics, equation (12) can be expressed as

$$\sigma^2[\ln(Y)] = \left[\left(\widehat{\beta_1} - \widehat{\beta_2}\right)^2 + \widehat{\sigma^2(\beta_1)} + \widehat{\sigma^2(\beta_2)}\right]\widehat{\sigma^2(S)} + \left[\widehat{\beta_2}^2 + \widehat{\sigma^2(\beta_2)}\right]\widehat{\sigma^2(A)}$$

$$+\left[2\widehat{\beta_2}\left(\widehat{\beta_1} - \widehat{\beta_2}\right) - \sigma^2(\beta_2)\right]\widehat{R_{AS}}\widehat{\sigma(A)}\widehat{\sigma(S)} + \widehat{\sigma^2(\beta_1)}\bar{S}^2$$

$$+\widehat{\sigma^2(\beta_2)}(\bar{A} - \bar{S} - 5)^2$$

$$+ \widehat{\sigma^2(\varepsilon)}$$

$$= 0.0142\widehat{\sigma^2(S)} + 0.0002\widehat{\sigma^2(A)} + 0.0032\widehat{R_{AS}}\widehat{\sigma(A)}\widehat{\sigma(S)} + 0.001\bar{S}^2$$

$$+ 0.00003(\bar{A} - \bar{S} - 5)^2 + \widehat{\sigma^2(\varepsilon)}$$

[20] This typical Mincerian earnings function may suffer from a selection bias, which may lead to an over estimation of the coefficient of years of schooling. Some remedies for the bias have been discussed and proposed by Card (1999, 2001), Carneiro et al. (2011), Hanushek and Zhang (2009), Heckman et al. (2008), among others. For the purpose of this research, we follow the treatment presented in Chiswick and Mincer (1972) and assume the coefficient of education to be a random variable across individuals.

[21] The estimation results of equation (11) by gender and industry group are reported in the Appendix.

$$= 0.202 + 0.027 - 0.056 + 0.145 + 0.019 + \widehat{\sigma^2(\varepsilon)}$$

$$= 0.337 + \widehat{\sigma^2(\varepsilon)}$$

The observed variance of the log of main employment income in 2016 is 0.522, as reported in Table 2. This human capital model suggests that 65% of income inequality among the workforce is attributed to its distribution of schooling and age. Moreover, both the level and the variance of schooling contribute to a significant portion of earnings disparity. Assuming all estimated parameters to be constant over the years and equal to their corresponding 2016 estimates, to assess the effect of each independent variable, the five partial derivatives derived in Table 1 are calculated based on the estimated results of equation (11).

$$\frac{\partial\sigma^2[\ln(Y)]}{\partial\sigma(S)} = 0.092 \qquad \frac{\partial\sigma^2[\ln(Y)]}{\partial\sigma(A)} = -0.00002 \qquad \frac{\partial\sigma^2[\ln(Y)]}{\partial\mu_S} = 0.021$$

$$\frac{\partial\sigma^2[\ln(Y)]}{\partial\mu_A} = 0.001 \qquad \frac{\partial\sigma^2[\ln(Y)]}{\partial R_{AS}} = 0.132$$

Schooling affects income inequality through three channels. First, with the expansion of higher education, more workers have received tertiary education nowadays. This would result in a more unequal distribution of labor earnings as a higher level of schooling is associated with a larger diversity (of something?) in terms of curriculum and quality. Given that there are different learning abilities and returns to education of workers, the impact on income inequality would be greater with an increasing level of schooling. From the partial derivative calculated above, a unit increase in schooling will raise income inequality by 0.021 points. On the other hand, as observed from the data, the increase in average schooling over the years is remarkable. It had risen by 4.58 years, from 7.52 years in 1981 to 12.1 years in 2016. According to our model, this would translate to an increase in income inequality over the period by 0.096 points, or about 22.4%.

Although education inflation itself would lead to a higher level of income inequality, compulsory education launched by the Hong Kong government since 1971 may have offered another channel to narrow the income gap. Promoting mandatory education not only guarantees students a minimum level of education but also helps to reduce the dispersion of years of schooling received by workers, and thereby narrows the distribution of productivity. As a result, income inequality can be eased by subsidizing free and compulsory education up to a certain level. In 1978, the government extended the universal basic education from primary school level (6 years) to junior secondary school level (9 years). 30 years later, education services have been named as one of Hong Kong's six priority industries with the aim to enhance Hong Kong's position as a regional education hub. The government therefore further extended the free education to senior secondary school level (12 years) in 2008 to show its commitment towards nurturing talents to support the growth of the economy and reinforcing Hong Kong's competitiveness. With these two waves of expansion, the standard deviation of schooling has dropped from 4.21 years in 1981 to 3.97 years in 1996. By 2016, it further declined to 3.78 years. This observed decrease in the standard deviation of schooling from 1981 to 2016 has contributed to a decrease in the variance of the log of income by 0.039 points. Given that the net change in the variance of log-income between 1981 and 2016 is 0.093 points, this channel has played an important role in alleviating income inequality.

The third channel for schooling to affect income inequality comes from the negative correlation between age and schooling. Since the expansion of schooling benefits younger and less experienced workers, the stronger the secular upward trend in schooling, the greater its impact is on narrowing the income gap. Over the sample period, the correlation has decreased from -0.344 in 1981 to -0.427 in 2016, indicating a moderate secular upward trend in schooling. In fact, this channel is responsible for a decrease of 0.011 points in the variance of log-income, which is about 11.8% of the net change in income inequality over the period.

On the other hand, population ageing also affects earnings disparity among the workforce via two channels. Firstly, since age is a proxy for one's potential work experience, the older the worker, the more potential post-school training he has acquired, holding the level of schooling constant. In this sense, population ageing itself would exacerbate income inequality because an older age group is usually associated with a higher within-group inequality than a younger one. For young workers, their distribution of income is in general more compressed as they do not have much work experience to differentiate their productivity from each other. The labor market of the young workforce is therefore more competitive, resulting in a less dispersed wage distribution. In contrast, even though the starting salaries among young workers are similar regardless of job choice, the subsequent career paths and promotion opportunities differ significantly from job to job. Inequality among the elderly and more experienced workers is thus higher as the trajectories of their lifetime income diverge. Hence, given that the return to post-school training varies across workers of different gender and those working in different industries, its effect on income inequality would be larger as the average age of the workforce goes up. Quantitatively, the mean age of the workforce has increased from 38.51 years in 1981 to 43.62 years in 2016. This has led to a slight increase in income inequality by 0.007 points.

Secondly, the age composition of the workforce as reflected by the variance of age, in principle can also affect the extent of income disparity. Intuitively, a society which consists of a 50-year-old experienced worker and a 20-year-old young worker should have a more unequal income distribution as compared to a society with two 30-year-old workers. This is because age is positively correlated with work experience, and work experience determines one's labor market earnings. Therefore, with a more extreme age distribution, the level of income inequality should also be higher. However, as mentioned above, young people in general have more schooling than their elder counterparts. The upward secular trend of education causes a negative correlation between age and schooling, resulting in a negative impact of the variance of age on income inequality. With these two counteracting forces, the resulting partial effect of a unit change in the standard deviation of age on the variance of log-income is -0.00002. In other words, the two effects have completely offset each other, leading to a trivial impact of the variance of age on income distribution.

Tables 3A and 3B show the predicted change in income inequality among the workforce for the pre-handover period (1981 – 1996) and the post-handover period (2001 – 2016) respectively. By looking into the contribution of each component to the change in income inequality, it is revealed that the majority of the predicted change is contributed by the increase in the average level of schooling, while the decrease in the variance of schooling has offset some of its effect. Although the age composition of the workforce also plays a role in determining income inequality, the effects are relatively small. Before the transfer of sovereignty over Hong Kong to Mainland China, the model predicts that income inequality would rise by 0.032 points, from 0.429 in 1981 to 0.461 in 1996. Compared to the realized value of 0.471, the gap between the observed and the predicted income inequality is reasonably small. Similarly, given the actual changes of the demographic variables, the model also predicts accurately over the post-handover period. The predicted change in the variance of the log of earnings between 2001 and 2016 is 0.017. Comparing it with the corresponding actual change during the same period, the model accounts for about 68% of the changes.

Table 3A: Comparison of predicted income inequality (Pre-handover period)

	1981	1996	1996 – 1981	Contribution to change in $\sigma^2[\ln(y)]$
$\sigma(S)$	4.21	3.97	-0.24	-0.0218
$\bar{S}$	7.52	10.10	2.58	0.0548
$\sigma(A)$	10.95	9.54	-1.42	0.0003
$\bar{A}$	38.51	39.37	0.85	0.0013
R_{AS}	-0.344	-0.364	-0.02	-0.0026
Predicted difference				+0.032
$\sigma^2[\ln(y)]_{1981}$				0.429
$\sigma^2[\widehat{\ln(y)}]_{1996}$				0.461

Table 3B: Comparison of predicted income inequality (Post-handover period)

	2001	2016	2016 – 2001	Contribution to change in $\sigma^2[\ln(y)]$
$\sigma(S)$	3.94	3.78	-0.16	-0.0148
$\bar{S}$	10.50	12.09	1.59	0.0337
$\sigma(A)$	9.36	10.85	1.50	-0.0004
$\bar{A}$	40.09	43.62	3.53	0.0052

R_{AS}	-0.376	-0.427	-0.05	-0.0067
Predicted difference				+0.017
$\sigma^2[\ln(y)]_{2001}$				0.497
$\sigma^2[\widehat{\ln(y)}]_{2016}$				0.514

With an established quantitative relationship between the demographic variables and income inequality, we proceed to project the workforce structure for the years of 2021, 2026 and 2031, and forecast the changes in income inequality for the next 15 years. Before making the true out-of-sample forecasts, pseudo out-of-sample forecasts for the years of 2006, 2011 and 2016 are made using the 2001 Population Census 5%-sample dataset as a base to extrapolate the changes in the composition of the workforce. Applying the procedures outlined in Section 4, Table 4 reports the pseudo out-of-sample forecasts of the variance of the log of main employment income.

Table 4: Pseudo out-of-sample forecasts on the variance of log-income

	2006		2011		2016	
	Projected value	Contribution to change in $\sigma^2[\ln(y)]$	Projected value	Contribution to change in $\sigma^2[\ln(y)]$	Projected value	Contribution to change in $\sigma^2[\ln(y)]$
$\sigma(S)$	3.96	0.0018	3.93	-0.0028	3.83	-0.0092
$\bar{S}$	10.90	0.0085	11.43	0.0112	12.05	0.0132
$\sigma(A)$	9.67	-0.0001	10.28	-0.0001	10.70	-0.0001
$\bar{A}$	41.71	0.0024	42.65	0.0014	43.11	0.0007
R_{AS}	-0.383	-0.0009	-0.403	-0.0027	-0.416	-0.0024
Predicted difference		+0.0117		+0.0071		+0.0022
$\sigma^2[\widehat{\ln(y)}]_{T-5}$		0.497		0.509		0.516

$\sigma^2\widehat{[\ln(y)]}$	0.509	0.516	0.518

The observed values of the variance of the log of main employment income for 2006, 2011 and 2016 are 0.509, 0.514 and 0.522 respectively. Comparing them with the corresponding forecasted values reported in Table 4, we found that the human capital model is able to generate accurate forecasts over a 15-year horizon. While the magnitude of the forecast errors increases with the length of the forecast horizon, the forecasts of the variance of log-income generated based on the demographic projection for 2016 are still very close to their realized values, with errors below 1%. The accuracy achieved by the model suggests that the effects on income inequality contributed by determinants other than the workforce structure are relatively stable. On the other hand, it also justifies the use of this methodology to make true out-of-sample forecasts to shed light on future evolutions of income inequality. Table 5 reports the true out-of-sample forecasts of the variance of log-income using the same extrapolation method.

Table 5: True out-of-sample forecasts on the variance of log-income

	2021		2026		2031	
	Projected value	Contribution to change in $\sigma^2[\ln(y)]$	Projected value	Contribution to change in $\sigma^2[\ln(y)]$	Projected value	Contribution to change in $\sigma^2[\ln(y)]$
$\sigma(S)$	3.65	-0.012	3.39	-0.0239	3.08	-0.0285
$\bar{S}$	12.61	0.011	13.25	0.0136	13.80	0.0117
$\sigma(A)$	11.06	-0.0001	10.86	-0.0001	10.50	-0.0001
$\bar{A}$	44.14	0.0008	44.40	0.0004	44.92	0.0008
R_{AS}	-0.439	-0.0016	-0.449	-0.0012	-0.456	-0.0011
Predicted difference		-0.0018		-0.0111		-0.0171
$\sigma^2\widehat{[\ln(y)]}_{T-5}$		0.522		0.520		0.509
$\sigma^2\widehat{[\ln(y)]}$		0.520		0.509		0.492

Based on the extrapolated workforce structure, the variance of log-income is expected to decline gradually in the next 15 years. In 2021, the average level of schooling of the workforce will increase to 12.61 years, as compared to 12.09 years in 2016. This would lead to an increase in the variance of log-income by 0.011 points. On the other hand, the standard deviation of schooling is also expected to decrease from 3.78 years in 2016 to 3.65 years in 2021. This would lead to a decrease in income inequality by 0.012 points. In other words, the positive effect on income inequality caused by the increase in the average level of schooling will be completely offset by the negative effect exerted by the decrease in the dispersion of schooling. While an ageing working population will lead to a slight increase of income inequality by 0.0008 points, this effect is also canceled out by the corresponding negative effects caused by the decrease in the standard deviation in age and the correlation between age and schooling. As a result, it is forecasted that there will be a small decrease in income inequality from 0.522 points in 2016 to 0.520 points in 2021.

In 2026, although the average level of schooling will rise to 13.25 years, its effect on income inequality will be overweighed by the compression of the distribution of schooling. Moreover, the correlation between age and schooling among the workforce will decrease further to -0.449, which also strengthens the downward pressure on earnings disparity. This suggests that the government can expand higher education and narrow the income gap at the same time, given that the rise in education is universal and benefits young workers more. As the age structure of the workforce does not exert a significant impact on the distribution of earnings, the educational structure will dominate the change in the variance of the log-income, resulting in a net decrease in income inequality to 0.509 points in 2026.

The extrapolated workforce structure reveals that the demographic trends described above will continue till 2031. The average level of schooling will increase to 13.8 years, while its standard deviation will decrease to 3.08 years. With the secular upward trend in education, the correlation between age and education will keep on decreasing to -0.456, but the pace of the decrease will be slower as compared to that in previous years. This can be reconciled by the fact that most of the less educated workers will be retired by 2031. Therefore, old and experienced workers by that time will be relatively highly educated. Moreover, the fertility rate in Hong Kong has been low during the entire post-handover period.[22] This will eventually result in a relatively small proportion of young people joining the workforce by the year 2031. Hence, the correlation between age and schooling among the workforce may eventually converge and fluctuate around a certain level, and its effect of easing earnings disparity will vanish in the long run. Again, population ageing will continue, and the average age among the workforce will rise to 44.92 years, with a standard deviation of 10.5 years. Although the combined effect of the change in age structure will exacerbate income inequality, its impact will be overweighed by the downward pressure caused by the expansion of educational structure. Thus, the forecasted variance of log-income will decrease to 0.492 points in 2031, which is a level comparable to that in 2001.

6. Forecasts of Other Income Inequality Indices

Instead of the variance of the log of main employment income, income inequality is usually reported in terms of the Gini coefficient. In this section, the baseline Mincerian earnings function derived in Section 3 is used to forecast a set of standard income inequality indices, including those reported in Table 2, for the next 15 years up to 2031. Pseudo out-of-sample forecasts for the years of 2001, 2006 and 2011 are reported to check the performance of the model. In addition, a modified version of the earnings function is also estimated to compare with the baseline model and determine if the accuracy can be improved by adding other control variables. The enhanced earnings function is specified below:

$$\begin{aligned}\ln(Y_i) &= \delta_0 + \delta_1 S + \delta_2 T + \delta_3 T^2 + \delta_4 US \times S + \delta_5 nondeg \times S + \delta_6 College \times S \\ &\quad + \delta_7 PG \times S + \delta_8 Fem + \delta_9 Fem \times S + \delta_{10} Fem \times T \\ &\quad + u_i \qquad (13)\end{aligned}$$

where US, $nondeg$, $College$ and PG are dummy variables with 1 indicating the highest educational attainment being upper secondary school, non-degree post-secondary

[22] The total fertility rate in Hong Kong has fluctuated at around 1.2 in the post-handover period, which is consistently below the replacement level of 2.1. In particular, the economic recession from 1998 to 2003 discouraged women from giving birth to children, leading to the record low total fertility rate of 0.931 in the early 2000s, as revealed by the 2001 Hong Kong Population Census.

education, college degree and postgraduate degree respectively. Fem is a gender dummy with 1 representing female and 0 indicating male.

Compared to equation (10), equation (13) allows for a non-linear effect of years of schooling on one's employment earnings. Besides, a gender dummy is also incorporated to account for the earnings gap between males and females. Two interactive terms, $Fem \times S$ and $Fem \times T$, enable males and females to have different returns to education and returns to work experience. The estimation results of equation (10) and equation (13) are reported in Table 6.

With all the coefficients being statistically significant at the 1% level, the estimation results derived from equation (10) show that the coefficients of the independent variables were stable over the period between 2001 and 2016, except schooling, which increased slightly from 0.1191 to 0.1251. On the other hand, the enhanced earnings function implies that returns to education increase with one's level of schooling. In 2016, the average returns to an additional year of schooling for male high school graduates, non-degree post-secondary education diploma holders, college graduates and postgraduate degree holders were 4.94%, 6.9%, 7.85% and 9.51% respectively. Moreover, it is also revealed that females were in general earning less than males in the labor market. In equation (13), the female dummy affects $\ln(Y)$ via three channels. Firstly, via the direct effect of the gender dummy, the $\ln(Y)$ of a female worker is on average lower than a male worker by 0.212, which is equivalent to a decline in main employment income by 19.1%, ceteris paribus. Secondly, females are more efficient school learners than males, as reflected by the positive coefficient of the interactive term $Fem \times S$. The better quality of formal schooling received by females would enable them to earn more in the labor market. Thirdly, females do not benefit as much as males in terms of work experience, despite the fact that the difference has narrowed over time. In 2001, the coefficient of work experience for males is 0.0545, but that for females is 0.049, with a difference of 0.0055. In 2016, those for males and females rose to 0.0549 and 0.0516 respectively, with the difference declining to 0.0033. The lower returns to work experience for females can probably be explained by the fact that there are chances for women to be pregnant. Not only pregnant women are eligible for a paid maternity leave in Hong Kong, they often spend greater portions of their time at home rather than engaging in marketplace production activities after giving birth. Thus, women on average are paid less in a competitive market. Nevertheless, as women usually have a comparative advantage in the service sector, the rise of the tertiary sector since the post-handover period has facilitated the increase in relative wage and market work hours of women, which in turn mitigates the earnings disparity between males and females (Ngai and Petrongolo, 2017).

After estimating the coefficients of equation (10) and equation (13), the next step is to obtain the fitted value and the error term for each observation in the base year sample. The 5%-samples of 2001 Hong Kong Population Census and 2016 Hong Kong Population By-census are adopted to be the base year sample dataset for the pseudo out-of-sample forecasts and the true out-of-sample forecasts respectively. The error terms are assumed to follow a normal distribution where the mean and variance are constant within but different across age and educational groups over the forecast horizon.[23] After that, we apply the procedures specified in Section 4 to project and simulate the workforce structure for the forecasted years. With the simulated datasets, equation (10) and equation (13) are used separately to predict the log-income earned by each simulated individual. An error term is then drawn from the respective distribution and added to the predicted log-income to arrive at the simulated log-income. Finally, by taking the exponential of the simulated log-income for each simulated observation, the whole simulated distribution of labor earnings can be obtained, and the income inequality indices can be calculated accordingly.

Tables 7 and 8 report the pseudo out-of-sample forecasts and true out-of-sample forecasts for the Gini coefficient, Theil index, Atkinson index, the variance of log-income, the 90th percentile to 50th percentile income ratio and the 50th percentile to 10th percentile income ratio. By comparing the pseudo out-of-sample forecasts with their corresponding realized values shown in Table 2, it reveals that the forecasted indices simulated by the enhanced earnings function (13) tend to deviate less from the observed figures. Nevertheless, all the pseudo out-of-sample forecasts generated by the baseline earnings function (10) are still reasonably close to the corresponding realized values. This implies that the projection and simulation methods adopted are useful in simulating the evolution of income distribution in future years. For the true out-of-sample forecasts, it is again predicted that the overall earnings disparity among the working population will be alleviated in the next 15 years. Based on the forecasts derived from the enhanced model, the Gini coefficient will decrease from 0.443 in 2021 to 0.435 in 2026, and it will further decrease to 0.429 in 2031. The Theil index and Atkinson index will also decrease from 0.368 to 0.341 and 0.276 to 0.253 respectively during the same period. For the 90th

[23] The distributions of the error terms can be obtained from the authors upon request.

percentile to 10th percentile income ratio, it is forecasted that it will decrease from 6.24 in 2021 to 5.84 in 2031. All the decreasing trends of income inequality indices are consistent with the patterns exhibited by the forecasts yielded by the baseline earnings function. In other words, the human capital approach suggests that earnings disparity among the workforce will ease naturally as the demographic structure evolves. (In particular the decrease in the dispersion of schooling)

Table 6: Estimation of equation (10) and equation (13)

Dependent variable: $\ln(Y)$	2001 ($N = 125{,}081$)		2016 ($N = 145{,}400$)	
	Equation (10)	Equation (13)	Equation (10)	Equation (13)
Constant	7.8246	8.4036	7.8226	8.6699
	(0.0112)	(0.0134)	(0.0098)	(0.0138)
S	0.1191	0.0151	0.1251	0.0163
	(0.0005)	(0.0011)	(0.0005)	(0.0013)
T	0.0307	0.0545	0.0297	0.0549
	(0.0006)	(0.0007)	(0.0005)	(0.0005)
T^2	-0.0004	-0.0009	-0.0003	-0.0008
	(0.00001)	(0.00001)	(0.00001)	(0.00001)
$US \times S$		0.0319		0.0331
		(0.0005)		(0.0005)
$nondeg \times S$		0.0528		0.0527
		(0.0007)		(0.0006)
$College \times S$		0.063		0.0622
		(0.0007)		(0.0007)
$PG \times S$		0.0758		0.0788
		(0.0008)		(0.0007)

Fem		-0.262		-0.212
		(0.0189)		(0.0178)
$Fem \times S$		0.074		0.0096
		(0.0011)		(0.001)
$Fem \times T$		-0.0055		-0.0033
		(0.0004)		(0.0003)
Adjusted R^2	0.3302	0.4089	0.3396	0.4137

Table 7: Pseudo out-of-sample forecasts of income inequality measures

Inequality Index	Equation (10)			Equation (13)		
	2006	2011	2016	2006	2011	2016
Gini coefficient	0.438	0.443	0.452	0.434	0.439	0.446
Theil index	0.359	0.367	0.378	0.345	0.352	0.366
Atkinson index	0.263	0.276	0.289	0.255	0.269	0.281
Variance of log-income	0.504	0.511	0.515	0.512	0.520	0.527
P90/P50	3.124	3.158	3.196	3.015	3.052	3.116
P50/P10	2.245	2.266	2.289	2.024	2.053	2.088

Table 8: True out-of-sample forecasts of income inequality measures

Inequality Index	Equation (10)			Equation (13)		
	2021	2026	2031	2021	2026	2031
Gini coefficient	0.448	0.441	0.436	0.443	0.435	0.429

Theil index	0.373	0.362	0.356	0.368	0.355	0.341
Atkinson index	0.281	0.272	0.264	0.276	0.268	0.253
Variance of log-income	0.517	0.505	0.486	0.523	0.512	0.494
P90/P50	3.146	3.118	3.072	3.045	2.980	2.955
P50/P10	2.183	2.167	2.125	2.048	2.027	1.975

7. Conclusion

Economic inequality has long been a polarizing issue in many countries, often drawing community-wide debates and special attention from policymakers. While plenty of studies have been conducted over the last few decades to provide ex post analyses on the causes of economic inequality, little effort has been devoted to ex ante forecasts of how income inequality will evolve over time. This paper provides a first attempt in the literature to forecast the future evolution of income inequality with the demographic projections. Since the lack of suitable time series data has prohibited the application of a conventional time series forecasting approach, we follow the human capital approach presented by Chiswick and Mincer (1972) to analyze the change in income inequality among the workforce aged 25 – 64 with a positive main employment income (excluding foreign domestic helpers). The model relates income inequality measured by the variance of log-income to the distribution of schooling, age and the intercorrelation between them. Results show that the rise in income inequality during the period between 1981 and 2016 is mainly caused by an increase in the average level of schooling, whereas population ageing has only played a limited role in influencing the income distribution.

While education inflation itself would exacerbate income inequality among the working population, it is counteracted by a fall in the dispersion of schooling over time stemmed from the implementation of the compulsory education scheme by the government. The latter turned out to be an important tool to mitigate earnings disparity. By projecting and extrapolating the demographic structure of the workforce, the model employed in this study forecasts that the standard deviation of schooling will decrease from 3.78 years in

2016 to 3.08 years in 2031. This will lead to an effect large enough to reverse the increasing trend of income inequality from 2016 onwards.

In addition to analyzing the change in income inequality in terms of the variance of the log-income, this paper also provides forecasts of a list of standard income inequality indices, including the Gini coefficient, Theil index, Atkinson index and the 90th percentile to 10th percentile income ratio. By simulating the whole income distribution for the years of 2021, 2026 and 2031 based on the projected workforce structure, we find that all forecasted indices derived from the baseline earnings function and the enhanced earnings function exhibit decreasing trends. In particular, based on the enhanced model, the Gini coefficient on individuals' earnings are forecasted to be 0.443, 0.435 and 0.429 in 2021, 2026 and 2031 respectively. These results suggest that earnings disparity among the working population will ease naturally with the evolution of demographic structure in Hong Kong.

Economic inequality is a complex topic that requires comprehensive analysis. While this paper only focuses on the quantitative effects of education expansion and population ageing on income inequality among the working population, future studies may extend the investigation to the effects of policy changes, such as tax reforms, social welfare and immigration policies on economic inequality. The forecast of household-level income inequality will also be an interesting extension of this paper, given that the decreasing trends in both marriage rate and household size in Hong Kong and other developed economies. If suitable longitudinal data are available in the future, one may also study and forecast the changes in intergenerational mobility over time. In short, this paper provides a first attempt to make use of demographic projections for analyzing income inequality. To alleviate economic inequality and promote social mobility, the government should spend more resources on conducting longitudinal surveys and developing more informative datasets, so that future socioeconomic policies can be more evidence-based and better informed.

The authors declare that they have no conflict of interest.

References

Acemoglu, Daron (1998), "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality", *The Quarterly Journal of Economics* 113(4), pp. 1055 – 1089.

– (2002), "Technical Change, Inequality, and the Labor Market", *Journal of Economic Literature* 40(1), pp. 7 – 72.

Aghion, Philippe, Eve Caroli, and Cecilia García-Peñalosa (1999), "Inequality and Economic Growth: The Perspective of the New Growth Theories", *Journal of Economic Literature* 37(4), pp. 1615 – 1660.

Banerjee, Abhijit V. and Andrew F. Newman (1993), "Occupational Choice and the Process of Development", *Journal of Political Economy* 101(2), pp. 274 – 299.

Banerjee, Abhijit V. and Esther Duflo (2003), "Inequality and Growth: What Can the Data Say?" *Journal of Economic Growth* 8(3), pp. 267 – 299.

Becker, Gary S. (1967), *Human Capital and the Personal Distribution of Income: An Analytical Approach*, W. S. Woytinsky Lecture no. 1, Ann Arbor: Institute of Public Administration, University of Michigan.

– (1971), *The Economics of Discrimination* (2nd ed.), Chicago: University of Chicago Press.

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

Blundell, Richard, Lorraine Dearden, Costas Meghir, and Barbara Sianesi (1999), "Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy", *Fiscal Studies* 20(1), pp. 1 – 23.

Burtless, Gary (1995), "International Trade and the Rise in Earnings Inequality", *Journal of Economic Literature* 33(2), pp. 800 – 816.

Card, David (1999), "Causal Effect of Education on Earnings", in: *Handbook of Labor Economics*, ed. O. Ashenfelter and D. Card, Amsterdam: North-Holland, pp. 1801 – 1863.

– (2001), "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica* 69(5), pp. 1127 – 1160.

Carneiro, Pedro, James J. Heckman and Edward J. Vytlacil (2011), "Estimating Marginal Returns to Education", *The American Economic Review* 101(6), pp. 2754 – 2781.

Census and Statistics Department (2017), *Hong Kong 2016 Population By-census – Thematic Report: Household Income Distribution in Hong Kong*, Hong Kong: Census and Statistics Department.

Central Intelligence Agency (2017), *The World Factbook 2017*, Washington: Central Intelligence Agency.

Chau, Leung-chuen (1994), "Economic Growth and Income Distribution in Hong Kong", in: *25 Years of Social and Economic Development in Hong Kong*, ed. B. K. P. Leung and T. Y. C. Wong, Hong Kong: Centre of Asian Studies, University of Hong Kong.

Chiswick, Barry R. and Jacob Mincer (1972), "Time-Series Changes in Personal Income Inequality in the United States from 1939, with Projections to 1985", *Journal of Political Economy* 80(3), pp. S34 – S66.

Chow, Steven C. and Gustav F. Papanek (1981), "Laissez-Faire, Growth and Equity - Hong Kong", *The Economic Journal* 91(362), pp. 466 – 485.

Clarke, George R. G., Lixin Colin Xu, and Heng-fu Zou (2006), "Finance and Income Inequality: What Do the Data Tell Us?", *Southern Economic Journal* 72(3), pp. 578 – 596.

De Gregorio, José and Jong-wha Lee (2002), "Education and Income Inequality: New Evidence from Cross-Country Data", *The Review of Income and Wealth* 48(3), pp. 395 – 416.

Darity, William A. Jr. and Patrick L. Mason (1998), "Evidence on Discrimination in Employment: Codes of Color, Codes of Gender", *Journal of Economic Perspectives* 12(2), pp. 63 – 90.

Furusawa, Taiji and Hideo Konishi (2016), "International Trade and Income Inequality", *Boston College Working Paper in Economics* (849).

Galor, Oded and Joseph Zeira (1993), "Income Distribution and Macroeconomics", *The Review of Economic Studies* 60(1), pp. 35 – 52.

Gindelsky, Marina (2016), "Will Inequality Continue to Rise? Forecasting Income Inequality in the United States", *Working Paper*.

Goldin, Claudia and Lawrence F. Katz (2008), *The Race between Education and Technology*, Cambridge: Belknap Press of Harvard University.

Hanushek, Eric and Lei Zhang (2009), "Quality-Consistent Estimates of International Schooling and Skill Gradients", *Journal of Human Capital* 3(2), pp. 107 – 143.

Heckman, James, Lance Lochner and Petra Todd (2008), "Earnings Functions and Rates of Return", *Journal of Human Capital* 2(1), pp. 1 – 31.

Hsia, Ronald and Larry Chau (1978), "Industrialisation and Income Distribution in Hong Kong", *International Labor Review* 117(4), pp. 465 – 479.

Johnson, George E. (1997), "Changes in Earnings Inequality: The Role of Demand Shifts", *Journal of Economic Perspectives* 11(2), pp. 41 – 54.

Kuznets, Simon (1955), "Economic Growth and Income Inequality", *The American Economic Review* 45(1), pp. 1 – 28.

Lam, Kit-chun and Pak-wai Liu (1998), "Immigration, Population Heterogeneity, and Earnings Inequality in Hong Kong", *Contemporary Economic Policy* 16(3), pp. 265 – 276.

Lui, Hon-kwong (2013), *Widening Income Distribution in Post-Handover Hong Kong*, New York: Routledge.

Mincer, Jacob (1958), "Investment in Human Capital and Personal Income Distribution", *Journal of Political Economy* 66(4), pp. 281 – 302.

Moore, Henry L. (1911), *Laws of Wages: An Essay in Statistical Economy*, New York: Macmillan.

Murphy, Kevin M. and Robert H. Topel (2016), "Human Capital Investment, Inequality, and Economic Growth", *Journal of Labor Economics* 34(S2), pp. S99 – S127.

Ngai, L. Rachel and Barbara Petrongolo (2017), "Gender Gap and the Rise of the Service Economy", *American Economic Journal: Macroeconomics* 9(4), pp. 1 – 44.

Panizza, Ugo (2002), "Income Inequality and Economic Growth: Evidence from American Data", *Journal of Economic Growth* 7(1), pp. 25 – 41.

Persson, Torsten and Guido Tabellini (1994), "Is Inequality Harmful for Growth?", *The American Economic Review* 84(3), pp. 600 – 621.

Pigou, Arthur C. (1932), *The Economics of Welfare* (4th ed.), London: Macmillan.

Piketty, Thomas (2014), *Capital in the Twenty-First Century*, Cambridge: Belknap Press if Harvard University Press.

Ruiz-Tagle, Jaime (2007), "Forecasting Wage Inequality", *Estudios de Economía* 34(2), pp. 141 – 162.

Schenk, Catherine R. (2008), "Economic History of Hong Kong", in: EH.Net Encyclopedia, ed. R. Whaples.
URL: http://eh.net/encyclopedia/economic-history-of-hong-kong/

Schumpeter, Joseph (1916), "Das Grundprinzip der Verteilungstheorie" *Archiv für Sozialwissenschaft und Sozialpolitik* 42, pp. 1 – 89.

Schultz, Theodore W. (1975), "The Value of the Ability to Deal with Disequilibria", *Journal of Economic Literature* 13(3), pp. 827 – 846.

Staehle, Hans (1943), "Ability, Wages and Income", *The Review of Economics and Statistics* 25(1), pp. 77 – 87.

Suen, Wing (1995), "Sectoral Shifts: Impacts on Hong Kong Workers", *The Journal of International Trade & Economic Development* 4(2), pp. 135 – 152.

Rosen, Sherwin (1981), "The Economics of Superstars", *The American Economic Review* 71(5), pp. 845 – 858.

Turner, H.A., Patricia Fosh, and Shek-hong Ng (1991), *Between Two Societies: Hong Kong Labour in Transition*, Hong Kong Centre of Asian Studies, University of Hong Kong.

Violante, Giovanni L. (2008), "Skill-Biased Technological Change", in: *The New Palgrave Dictionary of Economics* (2nd ed.), ed. S. N. Durlauf and L. E. Blume, New York: Palgrave Macmillan.

Wong, Yue Chim Richard (1991), "Hong Kong in Transition: Economic Transformation in the Eighties", in: *Proceedings of the Conference on the Changing World Community: Development Consequences and Adjustments in Asia and the Pacific, Ninth General Meeting of the Association of Development Research and Training Institutes of Asia and the Pacific*, pp. 1 – 18.

– (2017), *Fixing Inequality in Hong Kong*, Hong Kong: Hong Kong University Press.

Yang, Juan and Muyuan Qiu (2016), “The Impact of Education on Income Inequality and Intergenerational Mobility”, *China Economic Review* 37, pp. 110 – 125.

Zhong, Hai (2011), “The Impact of Population Aging on Income Inequality in Developing Countries: Evidence from Rural China”, *China Economic Review* 22(1), pp. 98 – 107.

Appendix

A1. Estimation results of equation (11) by gender and industry

Male	β_0	β_1	β_2	R^2	N
Manufacturing	8.238 (0.0565)	0.104 (0.0031)	0.014 (0.0009)	0.241	3,695
Construction	8.485 (0.0263)	0.092 (0.0015)	0.012 (0.0004)	0.235	12,763
Import/export trade and wholesale	8.075 (0.0414)	0.116 (0.0023)	0.020 (0.0006)	0.226	8,881
Retail, accommodation and food services	8.438 (0.0294)	0.086 (0.0018)	0.013 (0.0005)	0.165	11,183
Transportation, storage and courier services	8.582 (0.0292)	0.085 (0.0017)	0.005 (0.0004)	0.223	11,132
Information and communication	7.560 (0.0615)	0.145 (0.0035)	0.026 (0.0009)	0.305	4,338
Financing, insurance real estate, professional and business services	7.395 (0.0389)	0.172 (0.0021)	0.018 (0.0005)	0.341	13,872
Public administration, social and personal services	7.952 (0.0298)	0.135 (0.0016)	0.014 (0.0004)	0.355	14,460
Other industries	8.234 (0.1248)	0.115 (0.0062)	0.014 (0.0021)	0.373	665

A1. Continued

Female	β_0	β_1	β_2	R^2	N
Manufacturing	8.249 (0.0701)	0.098 (0.0038)	0.009 (0.0011)	0.269	2,402
Construction	8.520 (0.0802)	0.088 (0.0043)	0.007 (0.0014)	0.286	1,530
Import/export trade and wholesale	7.924 (0.0453)	0.119 (0.0026)	0.015 (0.0007)	0.231	7,154
Retail, accommodation and food services	8.594 (0.0252)	0.072 (0.0015)	0.004 (0.0004)	0.190	13,959
Transportation, storage and courier services	7.986 (0.0631)	0.115 (0.0036)	0.012 (0.0009)	0.272	3,163
Information and communication	7.314 (0.105)	0.155 (0.0060)	0.027 (0.0014)	0.300	1,702
Financing, insurance real estate, professional and business services	7.433 (0.0420)	0.163 (0.0023)	0.016 (0.0006)	0.303	12,549
Public administration, social and personal services	7.843 (0.0233)	0.134 (0.0012)	0.012 (0.0004)	0.426	21,714
Other industries	8.163 (0.2408)	0.104 (0.0122)	0.012 (0.0042)	0.285	238

A2. Data on the mean and standard deviation of age and schooling and their correlation

Year	A	$\sigma(A)$	S	$\sigma(S)$	R_{AS}
1981	38.51	10.95	7.52	4.21	-0.344
1986	38.62	10.71	8.22	4.10	-0.362
1991	38.98	10.21	9.15	4.09	-0.363
1996	39.37	9.54	10.10	3.97	-0.364
2001	40.09	9.36	10.50	3.94	-0.376
2006	41.17	9.56	11.03	3.82	-0.379
2011	42.53	10.30	11.58	3.80	-0.419
2016	43.62	10.85	12.09	3.78	-0.427

A3. Workforce participation rate by gender, age and education group, 2001 – 2016

2001	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	77.37%	80.23%	82.02%	82.84%	80.12%	75.34%	62.50%	38.19%
Upper Secondary	87.61%	89.77%	90.45%	89.47%	87.67%	80.98%	70.47%	43.85%
Non-degree post-secondary	94.33%	94.31%	94.24%	94.30%	91.99%	86.36%	71.89%	47.88%
College	90.14%	93.36%	93.11%	94.62%	92.80%	86.88%	74.69%	49.69%
Postgraduate	85.89%	95.72%	93.85%	93.88%	96.34%	90.36%	81.82%	67.61%

2001	Ever-married Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	37.33%	33.98%	34.40%	40.04%	43.90%	35.17%	22.73%	9.21%
Upper Secondary	73.94%	69.07%	62.85%	57.10%	55.36%	45.75%	29.23%	13.22%
Non-degree post-secondary	83.78%	80.22%	77.42%	73.52%	63.32%	61.02%	44.50%	14.74%
College	81.08%	79.84%	73.75%	68.42%	63.97%	61.30%	45.85%	23.13%
Postgraduate	72.29%	82.67%	79.92%	82.04%	77.67%	82.46%	42.86%	25.00%

A3. Continued

2001	Never-married Female							
	25 – 29	30 – 34	35 – 39	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	64.03%	65.42%	68.59%	62.98%	60.67%	43.78%	38.71%	18.00%
Upper Secondary	87.64%	86.62%	86.71%	85.23%	79.63%	69.46%	40.82%	23.81%
Non-degree post-secondary	93.74%	90.53%	93.33%	85.94%	84.34%	90.00%	76.47%	16.67%
College	92.61%	92.56%	90.66%	88.07%	88.62%	66.67%	75.00%	33.33%
Postgraduate	86.92%	91.94%	92.48%	90.00%	95.65%	92.31%	100.00%	33.33%

2006	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	77.67%	79.90%	79.73%	81.44%	79.29%	74.69%	58.42%	36.31%
Upper Secondary	84.08%	88.94%	88.42%	89.01%	86.38%	80.51%	63.08%	41.65%
Non-degree post-secondary	88.73%	91.53%	91.42%	90.93%	87.42%	84.36%	63.33%	37.15%
College	89.82%	92.67%	91.44%	91.68%	87.92%	83.43%	66.89%	45.02%
Postgraduate	84.76%	94.59%	93.42%	93.40%	89.45%	86.79%	68.42%	55.45%

A3. Continued

2006	Ever-married Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	37.93%	40.62%	39.68%	41.92%	43.10%	38.52%	24.89%	10.13%
Upper Secondary	64.77%	67.96%	62.61%	59.96%	55.84%	47.30%	28.65%	15.02%
Non-degree post-secondary	74.04%	81.42%	69.74%	72.48%	60.38%	57.34%	31.94%	16.49%
College	80.59%	79.31%	73.49%	72.32%	63.52%	61.30%	42.30%	23.81%
Postgraduate	74.76%	83.05%	78.87%	80.59%	77.27%	70.83%	37.35%	61.11%

2006	Never-married Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	65.26%	62.53%	57.81%	64.30%	55.16%	49.89%	31.05%	19.35%
Upper Secondary	84.40%	84.62%	81.62%	80.11%	72.27%	69.69%	48.42%	27.91%
Non-degree post-secondary	91.94%	86.69%	81.40%	80.07%	83.56%	67.16%	43.75%	38.46%
College	90.38%	90.24%	88.74%	80.52%	80.33%	65.18%	48.98%	32.14%
Postgraduate	86.95%	90.72%	90.31%	91.30%	91.09%	68.57%	55.00%	41.67%

A3. Continued

2011	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	67.07%	75.25%	77.20%	79.42%	80.96%	78.33%	67.09%	45.64%
Upper Secondary	82.59%	87.25%	87.45%	86.84%	86.00%	84.49%	70.37%	47.34%
Non-degree post-secondary	87.30%	92.83%	91.34%	90.78%	91.74%	87.21%	76.16%	43.19%
College	88.94%	93.87%	93.50%	90.78%	89.59%	88.46%	73.05%	50.71%
Postgraduate	84.20%	94.01%	95.74%	94.29%	93.03%	89.25%	83.27%	63.92%

2011	Ever-married Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	27.35%	31.24%	39.63%	45.41%	47.93%	41.29%	31.25%	15.85%
Upper Secondary	59.72%	63.80%	62.15%	60.80%	59.24%	53.53%	41.01%	19.49%
Non-degree post-secondary	74.52%	72.09%	71.12%	71.01%	69.33%	62.44%	50.24%	16.43%
College	77.84%	79.08%	77.13%	74.03%	72.00%	68.87%	53.70%	21.03%
Postgraduate	74.23%	83.63%	83.02%	79.60%	79.50%	79.58%	62.33%	33.33%

A3. Continued

2011	Never-married Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	49.73%	54.28%	52.75%	57.93%	58.35%	54.28%	47.19%	27.85%
Upper Secondary	80.91%	85.11%	82.30%	78.38%	75.52%	71.83%	52.70%	23.43%
Non-degree post-secondary	87.38%	88.70%	87.08%	84.54%	76.50%	75.18%	54.79%	32.43%
College	91.80%	91.78%	91.77%	89.33%	87.13%	75.00%	65.75%	25.00%
Postgraduate	86.99%	90.48%	92.71%	92.75%	89.27%	81.25%	81.82%	46.15%

2016	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	74.49%	78.64%	80.49%	82.24%	81.90%	78.92%	73.50%	57.59%
Upper Secondary	81.56%	85.61%	86.61%	88.19%	85.68%	85.00%	77.67%	55.45%
Non-degree post-secondary	83.68%	90.05%	88.97%	89.13%	86.65%	85.59%	78.83%	55.92%
College	86.80%	92.24%	92.45%	91.59%	89.02%	86.90%	79.12%	57.19%
Postgraduate	80.70%	92.62%	94.17%	93.67%	90.27%	89.47%	82.40%	61.80%

A3. Continued

2016	Ever-married Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	25.80%	29.11%	36.86%	48.21%	51.77%	47.73%	38.80%	23.19%
Upper Secondary	45.84%	53.65%	57.16%	58.88%	60.73%	57.90%	46.83%	29.28%
Non-degree post-secondary	60.47%	67.90%	64.34%	62.85%	67.51%	66.87%	48.73%	28.96%
College	74.69%	76.67%	76.38%	70.27%	72.29%	67.76%	56.65%	31.58%
Postgraduate	81.08%	81.42%	78.71%	79.10%	79.68%	73.74%	63.42%	45.75%

2016	Never-married Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	47.64%	60.37%	57.94%	59.66%	47.76%	55.15%	48.47%	36.94%
Upper Secondary	78.94%	79.67%	79.90%	77.81%	72.01%	69.13%	57.65%	35.57%
Non-degree post-secondary	83.50%	84.28%	88.45%	79.01%	74.40%	72.39%	66.29%	31.52%
College	85.03%	89.57%	87.86%	87.64%	80.92%	71.24%	52.51%	31.18%
Postgraduate	82.57%	85.66%	90.42%	88.48%	82.57%	77.21%	65.59%	45.65%

A4. Projected and realized female marriage rate by age and education group, 2006 – 2031

Projected 2006	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	63.43%	85.37%	89.22%	92.09%	92.44%	95.25%	98.47%	99.14%
Upper Secondary	37.98%	65.68%	77.29%	82.07%	84.79%	87.45%	94.24%	94.97%
Non-degree post-secondary	26.49%	59.58%	70.04%	76.23%	77.28%	84.88%	88.12%	91.21%
College	20.81%	54.35%	69.32%	73.30%	77.41%	84.31%	89.96%	94.23%
Postgraduate	22.50%	53.19%	70.25%	69.63%	71.04%	78.36%	79.23%	81.60%

Realized 2006	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	62.63%	84.70%	89.82%	91.08%	92.99%	95.01%	96.25%	98.60%
Upper Secondary	34.42%	65.68%	78.46%	83.40%	86.31%	88.99%	90.69%	95.23%
Non-degree post-secondary	24.11%	58.35%	72.83%	75.00%	79.97%	84.60%	85.71%	93.72%
College	19.35%	54.59%	69.69%	74.53%	75.39%	78.79%	87.11%	90.00%
Postgraduate	21.19%	55.00%	69.24%	69.74%	70.55%	77.42%	80.58%	80.00%

A4. Continued

Projected 2011	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	59.82%	84.60%	88.03%	90.82%	90.31%	93.63%	98.36%	99.28%
Upper Secondary	32.89%	61.62%	75.04%	79.42%	81.87%	84.28%	93.73%	93.46%
Non-degree post-secondary	23.36%	58.25%	68.18%	73.79%	75.20%	83.24%	84.42%	88.36%
College	17.93%	51.12%	68.30%	71.30%	75.99%	79.89%	86.65%	92.55%
Postgraduate	22.05%	50.32%	70.86%	68.80%	69.95%	75.29%	77.69%	79.00%

Realized 2011	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	60.59%	84.08%	90.91%	92.84%	92.60%	94.00%	95.61%	96.90%
Upper Secondary	34.13%	62.25%	75.72%	82.37%	84.40%	87.27%	89.83%	91.17%
Non-degree post-secondary	20.47%	54.55%	71.17%	77.70%	80.85%	82.80%	85.31%	90.51%
College	16.75%	52.85%	68.74%	75.32%	76.39%	77.60%	85.16%	88.27%
Postgraduate	20.95%	53.54%	68.21%	68.60%	71.27%	74.74%	76.84%	78.69%

A4. Continued

Projected 2016	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	56.22%	83.83%	86.85%	89.56%	88.19%	92.02%	98.25%	99.42%
Upper Secondary	27.80%	57.56%	72.78%	76.77%	78.96%	81.10%	93.21%	91.96%
Non-degree post-secondary	20.23%	56.92%	66.32%	71.35%	73.13%	81.60%	80.72%	85.52%
College	15.06%	47.88%	67.28%	69.29%	74.57%	75.48%	83.34%	90.88%
Postgraduate	21.60%	47.45%	71.47%	67.96%	68.86%	72.21%	76.15%	76.39%

Realized 2016	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	59.04%	83.21%	90.88%	94.09%	93.03%	92.56%	93.41%	95.59%
Upper Secondary	33.73%	60.70%	74.13%	81.88%	83.57%	85.80%	86.99%	89.02%
Non-degree post-secondary	20.36%	52.44%	70.24%	77.30%	80.18%	80.73%	81.02%	88.77%
College	16.13%	51.00%	68.12%	74.38%	76.45%	75.67%	82.90%	86.06%
Postgraduate	20.40%	49.85%	67.87%	67.37%	70.32%	72.78%	73.43%	76.88%

A4. Continued

Projected 2021	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	57.48%	82.35%	90.85%	95.34%	93.46%	91.13%	91.21%	94.27%
Upper Secondary	33.33%	59.16%	72.55%	81.38%	82.75%	84.34%	84.14%	86.87%
Non-degree post-secondary	20.24%	50.33%	69.31%	76.90%	79.50%	78.66%	76.73%	87.03%
College	15.52%	49.14%	67.50%	73.44%	76.52%	73.74%	80.63%	83.85%
Postgraduate	19.85%	46.16%	67.53%	66.13%	69.37%	70.82%	70.02%	75.08%

Projected 2026	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	55.93%	81.48%	90.83%	96.59%	93.88%	89.69%	89.00%	92.96%
Upper Secondary	32.93%	57.62%	70.96%	80.89%	81.92%	82.88%	81.30%	84.72%
Non-degree post-secondary	20.13%	48.21%	68.38%	76.50%	78.83%	76.59%	72.43%	85.29%
College	14.90%	47.29%	66.88%	72.50%	76.58%	71.82%	78.37%	81.63%
Postgraduate	19.30%	42.47%	67.19%	64.90%	68.43%	68.86%	66.60%	73.28%

A4. Continued

Projected 2031	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	54.38%	80.62%	90.80%	97.84%	94.31%	88.25%	86.80%	91.65%
Upper Secondary	32.53%	56.07%	69.37%	80.40%	81.10%	81.42%	78.46%	82.57%
Non-degree post-secondary	20.01%	46.10%	67.45%	76.10%	78.15%	74.52%	68.14%	83.55%
College	14.28%	45.43%	66.26%	71.56%	76.64%	69.89%	76.11%	79.42%
Postgraduate	18.75%	38.78%	66.84%	63.66%	67.48%	66.90%	63.19%	71.47%

A5. Projected and realized education structure by gender and age group, 2006 – 2031

Projected 2006	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	19.40%	23.25%	28.90%	36.32%	49.17%	59.18%	62.07%	65.43%
Upper Secondary	38.11%	38.19%	39.36%	38.12%	32.60%	27.22%	25.59%	21.10%
Non-degree post-secondary	12.93%	10.09%	8.90%	7.44%	5.63%	4.07%	3.79%	3.68%
College	25.08%	23.26%	18.38%	13.40%	9.23%	7.05%	6.77%	8.34%
Postgraduate	4.47%	5.21%	4.47%	4.72%	3.37%	2.49%	1.79%	1.46%

Realized 2006	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	18.39%	24.53%	29.50%	35.68%	48.49%	57.45%	62.36%	66.76%
Upper Secondary	41.66%	36.21%	37.36%	37.23%	32.00%	26.42%	25.30%	20.80%
Non-degree post-secondary	11.82%	9.94%	8.69%	7.57%	6.12%	5.16%	4.39%	3.96%
College	23.68%	22.72%	17.31%	13.28%	8.79%	7.66%	5.65%	6.91%
Postgraduate	4.46%	6.61%	7.15%	6.24%	4.59%	3.31%	2.31%	1.58%

A5. Continued

Projected 2006	Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	13.23%	23.61%	28.24%	37.92%	54.02%	67.45%	71.83%	74.95%
Upper Secondary	42.51%	39.25%	43.46%	44.28%	34.67%	23.87%	20.89%	15.97%
Non-degree post-secondary	12.97%	10.04%	8.63%	5.75%	3.73%	3.08%	2.76%	3.55%
College	27.20%	22.33%	16.35%	9.73%	6.11%	4.45%	3.85%	5.07%
Postgraduate	4.09%	4.77%	3.31%	2.32%	1.47%	1.14%	0.68%	0.44%

Realized 2006	Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	13.71%	22.87%	29.21%	38.18%	54.48%	67.15%	72.07%	75.72%
Upper Secondary	42.38%	40.87%	43.07%	42.75%	32.94%	24.36%	19.93%	15.43%
Non-degree post-secondary	12.77%	9.19%	8.40%	6.48%	4.49%	3.30%	3.28%	3.54%
College	26.95%	21.63%	14.98%	9.52%	5.98%	4.01%	3.71%	4.79%
Postgraduate	4.20%	5.43%	4.33%	3.06%	2.11%	1.18%	1.01%	0.51%

A5. Continued

Projected 2011	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	14.63%	17.60%	23.25%	28.90%	36.32%	49.17%	59.18%	62.07%
Upper Secondary	36.67%	37.03%	38.19%	39.36%	38.12%	32.60%	27.22%	25.59%
Non-degree post-secondary	14.63%	11.29%	10.09%	8.90%	7.44%	5.63%	4.07%	3.79%
College	28.83%	28.14%	23.26%	18.38%	13.40%	9.23%	7.05%	6.77%
Postgraduate	5.24%	5.95%	5.21%	4.47%	4.72%	3.37%	2.49%	1.79%

Realized 2011	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	11.97%	18.94%	25.14%	30.39%	35.66%	48.85%	59.66%	63.30%
Upper Secondary	37.27%	34.32%	33.88%	35.21%	36.71%	31.18%	24.56%	23.55%
Non-degree post-secondary	15.12%	12.67%	10.28%	9.36%	8.69%	6.99%	5.35%	4.98%
College	29.31%	24.40%	20.75%	15.85%	11.11%	7.96%	6.63%	5.39%
Postgraduate	6.32%	9.68%	9.95%	9.19%	7.82%	5.02%	3.79%	2.78%

A5. Continued

Projected 2011	Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	10.08%	18.97%	23.61%	28.24%	37.92%	54.02%	67.45%	71.83%
Upper Secondary	37.28%	35.04%	39.25%	43.46%	44.28%	34.67%	23.87%	20.89%
Non-degree post-secondary	14.15%	11.45%	10.04%	8.63%	5.75%	3.73%	3.08%	2.76%
College	33.25%	28.31%	22.33%	16.35%	9.73%	6.11%	4.45%	3.85%
Postgraduate	5.24%	6.23%	4.77%	3.31%	2.32%	1.47%	1.14%	0.68%

Realized 2011	Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	9.55%	17.51%	26.24%	31.40%	38.45%	54.85%	68.61%	73.18%
Upper Secondary	35.09%	35.65%	37.17%	40.79%	41.98%	32.56%	22.34%	18.98%
Non-degree post-secondary	16.48%	12.77%	10.03%	8.83%	7.00%	5.02%	3.82%	3.74%
College	32.57%	24.80%	19.10%	13.28%	9.05%	5.24%	3.78%	2.94%
Postgraduate	6.31%	9.27%	7.46%	5.70%	3.53%	2.32%	1.46%	1.17%

A5. Continued

Projected 2016	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	9.86%	11.94%	17.60%	23.25%	28.90%	36.32%	49.17%	59.18%
Upper Secondary	35.24%	35.87%	37.03%	38.19%	39.36%	38.12%	32.60%	27.22%
Non-degree post-secondary	16.33%	12.48%	11.29%	10.09%	8.90%	7.44%	5.63%	4.07%
College	32.58%	33.02%	28.14%	23.26%	18.38%	13.40%	9.23%	7.05%
Postgraduate	6.00%	6.69%	5.95%	5.21%	4.47%	4.72%	3.37%	2.49%

Realized 2016	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	8.82%	13.63%	17.96%	25.58%	29.53%	37.32%	48.96%	59.52%
Upper Secondary	30.56%	29.52%	30.62%	31.81%	34.21%	34.17%	29.33%	23.59%
Non-degree post-secondary	17.71%	15.65%	12.83%	11.15%	10.44%	9.60%	7.82%	6.02%
College	35.93%	30.13%	26.52%	21.22%	16.67%	12.67%	8.70%	7.51%
Postgraduate	6.98%	11.07%	12.06%	10.25%	9.15%	6.24%	5.18%	3.36%

A5. Continued

Projected 2016	Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	6.94%	14.33%	18.97%	23.61%	28.24%	37.92%	54.02%	67.45%
Upper Secondary	32.04%	30.83%	35.04%	39.25%	43.46%	44.28%	34.67%	23.87%
Non-degree post-secondary	15.33%	12.85%	11.45%	10.04%	8.63%	5.75%	3.73%	3.08%
College	39.30%	34.29%	28.31%	22.33%	16.35%	9.73%	6.11%	4.45%
Postgraduate	6.39%	7.69%	6.23%	4.77%	3.31%	2.32%	1.47%	1.14%

Realized 2016	Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	7.08%	12.88%	19.30%	27.50%	32.23%	41.56%	54.73%	68.24%
Upper Secondary	28.90%	29.86%	33.86%	34.87%	38.22%	37.24%	30.90%	21.57%
Non-degree post-secondary	17.63%	15.68%	12.70%	10.92%	9.91%	8.35%	5.97%	4.24%
College	37.88%	31.09%	24.54%	19.10%	14.19%	9.18%	6.14%	4.36%
Postgraduate	8.50%	10.49%	9.60%	7.62%	5.45%	3.67%	2.27%	1.58%

A5. Continued

Projected 2021	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	5.66%	8.32%	13.63%	17.96%	25.58%	29.53%	37.32%	48.96%
Upper Secondary	23.85%	24.72%	29.52%	30.62%	31.81%	34.21%	34.17%	29.33%
Non-degree post-secondary	20.31%	18.64%	15.65%	12.83%	11.15%	10.44%	9.60%	7.82%
College	42.55%	35.86%	30.13%	26.52%	21.22%	16.67%	12.67%	8.70%
Postgraduate	7.63%	12.46%	11.07%	12.06%	10.25%	9.15%	6.24%	5.18%

Projected 2021	Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	4.62%	8.26%	11.88%	19.30%	27.50%	32.23%	41.56%	54.73%
Upper Secondary	22.71%	24.07%	30.86%	33.86%	34.87%	38.22%	37.24%	30.90%
Non-degree post-secondary	18.78%	18.59%	15.68%	12.70%	10.92%	9.91%	8.35%	5.97%
College	43.19%	37.37%	31.09%	24.54%	19.10%	14.19%	9.18%	6.14%
Postgraduate	10.70%	11.72%	10.49%	9.60%	7.62%	5.45%	3.67%	2.27%

A5. Continued

Projected 2026	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	2.51%	3.01%	8.32%	13.63%	17.96%	25.58%	29.53%	37.32%
Upper Secondary	17.13%	19.92%	24.72%	29.52%	30.62%	31.81%	34.21%	34.17%
Non-degree post-secondary	22.90%	21.62%	18.64%	15.65%	12.83%	11.15%	10.44%	9.60%
College	49.17%	41.59%	35.86%	30.13%	26.52%	21.22%	16.67%	12.67%
Postgraduate	8.28%	13.85%	12.46%	11.07%	12.06%	10.25%	9.15%	6.24%

Projected 2026	Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	2.15%	3.64%	8.26%	11.88%	19.30%	27.50%	32.23%	41.56%
Upper Secondary	16.52%	18.28%	24.07%	30.86%	33.86%	34.87%	38.22%	37.24%
Non-degree post-secondary	19.94%	21.49%	18.59%	15.68%	12.70%	10.92%	9.91%	8.35%
College	48.50%	43.66%	37.37%	31.09%	24.54%	19.10%	14.19%	9.18%
Postgraduate	12.90%	12.94%	11.72%	10.49%	9.60%	7.62%	5.45%	3.67%

A5. Continued

Projected 2031	Male							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	0.00%	0.00%	3.01%	8.32%	13.63%	17.96%	25.58%	29.53%
Upper Secondary	9.77%	12.82%	19.92%	24.72%	29.52%	30.62%	31.81%	34.21%
Non-degree post-secondary	25.49%	24.61%	21.62%	18.64%	15.65%	12.83%	11.15%	10.44%
College	55.80%	47.32%	41.59%	35.86%	30.13%	26.52%	21.22%	16.67%
Postgraduate	8.94%	15.24%	13.85%	12.46%	11.07%	12.06%	10.25%	9.15%

Projected 2031	Female							
	25 – 29	30 – 34	35 – 59	40 – 44	45 – 49	50 – 54	55 – 59	60 – 64
Lower Secondary or below	0.00%	0.00%	3.64%	8.26%	11.88%	19.30%	27.50%	32.23%
Upper Secondary	10.01%	11.50%	18.28%	24.07%	30.86%	33.86%	34.87%	38.22%
Non-degree post-secondary	21.09%	24.40%	21.49%	18.59%	15.68%	12.70%	10.92%	9.91%
College	53.81%	49.94%	43.66%	37.37%	31.09%	24.54%	19.10%	14.19%
Postgraduate	15.09%	14.16%	12.94%	11.72%	10.49%	9.60%	7.62%	5.45%