Macroeconomic Effects of Income and Consumption Tax: Insights from a Life Cycle Model with New Zealand Microdata

Robert Kirkby Yao Yao Huong Ngoc Vu

June 17, 2024

Abstract

This paper investigates the macroeconomic effects of income and consumption taxes in New Zealand, using a life cycle model calibrated with administrative data from the Integrated Data Infrastructure (IDI). We document that mean earnings, Gini coefficient of earnings, and mean-to-median earnings ratio all increase over the working life cycle in New Zealand, aligning with patterns observed in the literature. By employing a human capital model with the appropriate distribution of initial human capital and learning ability, we effectively replicate these properties. Our analysis reveals that changes in consumption taxes primarily impact immediate consumption decisions, while changes in income tax rates have broader implications, influencing labor supply, savings, and human capital investment. We also emphasize the use of statutory tax rates in our model, highlighting their advantages over effective tax functions for policy analysis. Our findings provide valuable insights for policymakers, demonstrating the nuanced impacts of tax policies on economic behavior.

These results are not official statistics. They have been created for research purposes from the [Integrated Data Infrastructure (IDI) and/or Longitudinal Business Database (LBD)] which [is/are] carefully managed by Stats NZ. For more information about the [IDI and/or LBD] please visit https://www.stats.govt.nz/integrated-data/.

The results are based in part on tax data supplied by Inland Revenue to Stats NZ under the Tax Administration Act 1994 for statistical purposes. Any discussion of data limitations or weaknesses is in the context of using the IDI for statistical purposes, and is not related to the data's ability to support Inland Revenue's core operational requirements.

Copyright: This work is licensed under the Creative Commons Attribution 4.0 International licence. You are free to copy, distribute, and adapt the work, as long as you attribute the work to Kirkby, Yao, and Vu [2024] and abide by the other licence terms.
Liability: While all care and diligence has been used in processing, analysing, and extracting data and information in this publication, Stats NZ, Kirkby, Yao, and Vu [2024] gives no warranty it is error free and will not be liable for any loss or damage suffered by the use directly, or indirectly, of the information in this publication.

1 Introduction

This paper investigates the macroeconomic impacts of changes in income and consumption taxes in New Zealand, leveraging a life cycle model calibrated with administrative data from the Integrated Data Infrastructure (IDI). The primary focus is on understanding how these tax adjustments influence individuals' labor supply, savings, human capital investment, and earnings dynamics. Our approach draws on the established methodology of Huggett et al. [2011], which we adapt to the New Zealand context to capture local economic behaviors and outcomes.

In our study, we emphasize the importance of using statutory tax rates over effective tax functions in the New Zealand context. This choice is motivated by the simplicity and transparency of statutory rates, which facilitate more straightforward and comprehensible policy analysis. Statutory rates directly reflect the legal framework governing taxes, allowing us to model and simulate policy changes with greater accuracy and ease. Utilizing statutory tax rates also enables us to clearly illustrate the progressive nature of the New Zealand tax system and its implications for different income groups.

Our empirical analysis reveals that mean earnings, Gini coefficient of earnings, and mean-to-median earnings ratio all increase over most of the working life cycle in New Zealand. These findings align with those observed in the literature and underscore the relevance of initial conditions, such as human capital and learning ability, in shaping lifetime earnings trajectories. By calibrating our model to match New Zealand's microdata, we provide evidence that a human capital model can replicate these earnings dynamics, validating the model's applicability to the New Zealand economy.

We also find that changes in consumption taxes primarily affect immediate consumption decisions without significantly influencing labor supply, savings, or human capital investment. This is due to the direct impact of consumption taxes on the cost of consumption, which leads to a reduction in consumption levels. However, the relative prices of labor and capital remain unchanged, resulting in minimal adjustments in these areas. Conversely, changes in income tax rates have broader implications. Higher income taxes reduce the net wage from labor, thereby decreasing incentives for labor supply and human capital investment. This leads to lower overall earnings and disposable income, subsequently reducing consumption and savings. The progressive nature of income taxes also plays a redistributive role, addressing lifetime inequality by reallocating resources from higher to lower-income individuals.

This paper makes several contributions to the literature. First, we document the life-cycle earnings profiles for New Zealand, showing significant increases in mean earnings, dispersion, and skewness over time. Second, we demonstrate that a human capital model, when calibrated with appropriate initial conditions, can effectively capture these empirical patterns. Third, our use of statutory tax rates provides a novel approach to modeling tax policies, offering clear advantages for both theoretical analysis and practical policy evaluation. Finally, we reveal the macroeconomic effects of income and consumption taxes using this framework.

In the following sections, we detail the structure of our life cycle model, the calibration process using IDI data, and the findings from our policy simulations. Through this comprehensive analysis, we aim to shed light on the complex interactions between tax policies and macroeconomic variables, providing valuable insights for policymakers in New Zealand and beyond.

Related literature Our paper is closely aligned with Huggett et al. [2011] who delve into the sources of lifetime inequality, examining whether variations in lifetime outcomes primarily stem from disparities established early in life or from luck encountered throughout one's working years. Our paper differs from Huggett et al. [2011] in several important aspects. Firstly, we employ data encompassing both males and females to estimate age profiles and human capital shocks, aligning with the recommendation of Borella et al. [2018], who argue that incorporating data from both genders enhances the model's ability to match aggregate statistics. Secondly, we utilize administrative data instead of survey data for estimating age profiles and human capital shocks. This decision is supported by the insights of Johnson and Moore [2005], who highlight the advantages of administrative data, including broader population coverage, reduced nonresponse rates, and enhanced data quality. Finally, we focus on the impacts of tax changes on macroeconomic variables.

Numerous studies employ human capital models with intergenerational mechanisms to investigate lifetime inequality, such as Galor and Zeira [1993], Castaneda et al. [2003], Lee and Seshadri [2019], among others. These studies consistently demonstrate the pivotal role of the initial distribution of wealth in shaping human capital accumulation, thereby influencing later-life income and wealth distributions, as well as aggregate output and investment. The key feature of these models is that individuals inherit wealth from their parents, invest in human capital, engage in labor, and bequeath assets to their offspring. Higher levels of inherited wealth correspond to increased investment in human capital, leading to elevated future income levels. Our model departs from this intergenerational framework as we do not endogenize the distribution of initial conditions. We instead calibrate the distribution of initial conditions so that the model generates observed life-cycle earnings dynamics in New Zealand.

Our paper intersects with a strand of literature that estimates effective tax functions across various countries[1]. What sets us apart is our documentation that effective tax rates align with the statutory rates outlined

[1]For an overview of effective tax functions, see Borella et al. [2023].

by the New Zealand tax system. We therefore propose utilizing the statutory tax function, representing the relationship between taxes and income as prescribed by law, to model the effective income tax function in our study. Employing the statutory tax function not only simplifies the comprehension of different tax levels but also facilitates the analysis of changes in fiscal policy, offering a notable advantage.

Our paper also relates to the literature on tax structure. Several empirical studies, including Arnold [2008], Yanikkaya and Turan [2020], and Nguyen et al. [2021], suggest that shifting taxation from income to consumption may lead to a positive impact on economic growth, while Pestel and Sommer [2017] highlight that moving towards a consumption tax can reduce the excess burden of taxation, thereby enhancing aggregate efficiency. Using structural models, Chang et al. [1999] delve into the unemployment and wage effects of transitioning from an income tax to a consumption tax and show such a shift can significantly influence labor market dynamics depending on whether unemployment benefits are taxed or not, while Hansen and İmrohoroğlu [2018] demonstrate that replacing income taxation with consumption taxation can result in substantial increases in labor supply, investment, and output. Our research contributes to this body of literature by employing a human capital model that features idiosyncratic risk and heterogeneity in initial conditions, utilizes high-quality administrative data, and includes a rigorous calibration process to accurately reflect the economic context of New Zealand.

2 Model

This is a human capital model where human capital refers to the skills, knowledge, and abilities that an individual possesses, which enhance their productivity and earning potential. In this model, an individual seeks to maximize their lifetime satisfaction, considering their initial savings, human capital, and learning ability. They make decisions about consumption, savings, work, and learning throughout their life. At each age, their budget includes labor earnings and returns on savings minus taxes, and they decide on consumption and savings accordingly. Earnings depend on the wage rate, human capital, and work time, with earnings ceasing at retirement. Future human capital is influenced by current human capital, learning time, learning ability, and random shocks. The wage rate grows at a constant rate, and individuals cannot have negative savings, must deplete savings by the end of life, and must balance their time between work and learning. More technical details are provided below.

An agent maximizes expected lifetime utility, taking initial asset holding k_1, initial human captal h_1, and learning ability a^i as given. The decision problem for an agent born at time t is stated below.

$$\max_{\{c_j,k_j,h_j,l_j,s_j\}_{j=1}^{J}} E\Big[\sum_{j=1}^{J}\beta^{j-1}\frac{c_j^{1-\sigma}}{1-\sigma}\Big] \text{ subject to}$$

$$c_j + k_{j+1} = e_j + k_j(1+r_{t+j-1}) - T_{j,t+j-1}(e_j, c_j), \forall j; \tag{1}$$

$$e_j = w_{t+j-1}h_j l_j \text{ if } j < J_R, \text{ and } e_j = 0 \text{ otherwise}; \tag{2}$$

$$h_{j+1} = H(h_j, s_j, a^i, z_{j+1}); \tag{3}$$

$$w_{t+j-1} = (1+g)w_{t+j-2}; \tag{4}$$

$$k_{j+1} \geq 0;\ k_{J+1} = 0;\ \text{and } l_j + s_j = 1, \forall j.$$

where c_j, k_j, h_j are an age j agent's consumption, asset holding, and human capital, respectively. The agent is endowed with one unit of time which he/she can allocate in working activity (l_j) and learning activity (s_j). While working time directly influences the agent's earnings e_j, learning time is an input of the human capital production which, in turn, influences the agent's future earnings.

Equation (1) states that the agent's budget comprises labor earnings e_j and the value of assets $k_j(1+r_{t+j-1})$ less net taxes $T_{j,t+j-1}(e_j, c_j)$. Asset holdings pay a risk-free, real return r_{t+j-1} at time $t+j-1$. The agent pays net taxes $T_{j,t+j-1}(e_j, c_j)$ which equals labor income tax plus consumption tax less transfers. The agent then decide how much to consume c_j and how much to save for the next period k_{j+1} given this budget.

Equation (2) defines earnings e_j as a product of a rental rate w_{t+j-1} for human capital services, the agent's human capital h_j, and the fraction of time put into work l_j. When the agent reaches the retirement age J_R, his/her earnings are zero from then on.

Equation (3) states that the agent's future human capital h_{j+1} is a function of current human capital h_j, the fraction of time put into learning s_j, learning ability a^i where i denotes the agent's permanent type, and an idiosyncratic shock z_{j+1}. The human capital production function $H(h_j, s_j, a^i, z_{j+1})$ is assumed to have the following form

$$H(h, s, a, z') = \exp(z')\Big(h + a(hs)^{\gamma}\Big)$$

where γ is the elasticity parameter. Learning ability is fixed over an agent's lifetime and is exogenous. Idiosyncratic shocks are independent and identically distributed over time and follow a normal distribution $z_t \sim N(\mu, \sigma_z^2)$.

Equation (4) states that the rental rate w_{t+j-1} grows at a deterministic rate g. The rest of the agent's problem limits asset holdings to be non-negative in every period; and the total fraction of time for working and learning is one in every period.

3 Empirical analysis of the New Zealand economy

In this section, we use New Zealand administrative data on earnings to estimate age-conditional mean earnings, Gini coefficient of earnings, and the ratio of mean to median earnings. We then use data on those aged 55-65 to estimate the shocks to human capital. Finally, combining tax data with earnings we estimate effective tax rates, finding that they are close to the statutory rates.

3.1 Age profiles

Following Huggett et al. [2011], we estimate age profiles for mean earnings, earnings dispersion, and earnings skewness in New Zealand by using microdata on individuals' earnings. The dataset is obtained from the Inland Revenue tax database[2] for the period from 2006 to 2015.

We select data on annual earnings from wages and salaries (W&S) for individuals who satisfy two selection criteria. First, they must aged between 23 and 60. Second, they must earn at least \$9,500 (in 2006 prices) a year if they are 30 and older, or at least \$6,500 (in 2006 prices) a year if they are under 30. The motivations for these selection criteria are follows.

First, by selecting individuals at ages 23-60 and employing a five-year age bin[3], we have about 2,000 observations in each age-year bin for use to calculate earnings statistics. Second, the traditional retirement age in New Zealand is 65[4], and there is a significant decline in the labor force participation rate near this age due to reasons such as individual preference or health conditions. As these reasons are abstracted from in the model, we use the terminal age of 60 which is below the traditional retirement age. Third, individuals in the model allocate their time to either working or learning. We therefore select data on individuals who earn some income from labor work. For individuals age 30 and over, the lower bound for earnings is set below the annual earnings level of a full-time worker working at the New Zealand minimum wage[5]. For younger individuals, we reduce the minimum earnings to capture students doing part-time work while in school. Table 1 shows changes in the number of observations after every restriction criteria.

[2] i.e. the IR Restricted data for non-government researchers in IDI. We link the HES Income and IR Restrict data to collect information about individuals' earnings and demographics.

[3] For example, to form a bin of individuals age $j = 30$ in year $t = 2008$ we use data on individuals age 28-32 in 2008. We also use a five-year age bin centered at ages 23 and 60 - the minimum and maximum ages of selected individuals - by using data on individuals age 21-25 and 58-62.

[4] As there is no legal retirement age in New Zealand, 65 is commonly considered as the traditional retirement age because this is the superannuation qualification age.

[5] In 2006, the minimum wage rate was \$10.25. The annual earnings level of a full-time worker (working 40 hours per week and 52 weeks a year) working at the minimum wage is $\$10.25 \times 40 \times 52 = \$21,320$ in 2006 prices.

Sample	Number of observations
Income from Wages and Salaries (Linked IR data 2006-2015)	348,913
Select age range (21-62)	292,573
Select earnings range	256,308

Table 1: Number of observations

We convert nominal earnings to real earnings[6] and group them in specific age-year bins e_{jt} which is real earnings of all individuals in age bin j in year t[7]. We then calculate three statistics for every age-year bin namely the mean earnings, Gini coefficient of earnings, and earnings skewness measured by the ratio of mean earnings to median earnings[8].

We employ the classical Age-Period-Cohort (APC) model[9] to extract age effects on these earnings statistics:

$$stat_{j,t} = \alpha_c^{stat} + \beta_j^{stat} + \gamma_t^{stat} + \epsilon_{j,t}^{stat} \tag{5}$$

where $stat_{j,t}$ is each of the three statistics, α_c^{stat} is the cohort effects, β_j^{stat} is the age effects, γ_t^{stat} is the time effects, and $\epsilon_{j,t}^{stat}$ is error terms. We wish to estimate the age effects β_j^{stat}.

Model (6) separates three factors governing the evolution of the earnings statistic: cohort (c), age (j), and time (t). As birth cohort can be derived from time and age, which is $c = t - j$, the independent variables are colinear and their effects cannot be estimated without further restrictions. We follow Huggett et al. [2011] and use two alternative approaches. The first approach is the cohort effects model where we set $\gamma_t^{stat} = 0, \forall t$ [10]. The second approach is the time effects model where we set $\alpha_c^{stat} = 0, \forall c$ [11]. We use ordinary least squares to estimate the coefficients in each model.

Figure 1 plots the age effects, i.e. β_j^{stat} estimated from model (6), in mean earnings, earnings Gini, and earnings skewness. The age effects are normalized so that each profile runs through the mean value of each statistic across panel years at age 38.

Figure 1 highlights several facts about how mean earnings, earnings dispersion, and earnings skewness evolve with age. First, mean earnings are

[6] We collect annual CPI for New Zealand from the World Bank data and normalize so that CPI = 100 in 2006. Real earnings are equal to earnings divided by CPI.

[7] As we use five-year age bin, each age-year bin includes individuals age from $(j-2)$ to $(j+2)$ in year t.

[8] As the number of age is $J = 38$ and the number of year is $T = 10$ the number of observations for each statistic is equal to the number of age-year bins which is $J \times T = 380$.

[9] For readers who are unfamiliar with APC models, we suggest the review of Fosse and Winship [2019].

[10] In each regression the dependent variable is regressed on $J + T - 1$ cohort dummies and J age dummies. The intercept is not included.

[11] In each regression the dependent variable is regressed on J age dummies and $T - 1$ time dummies. The intercept is not included.

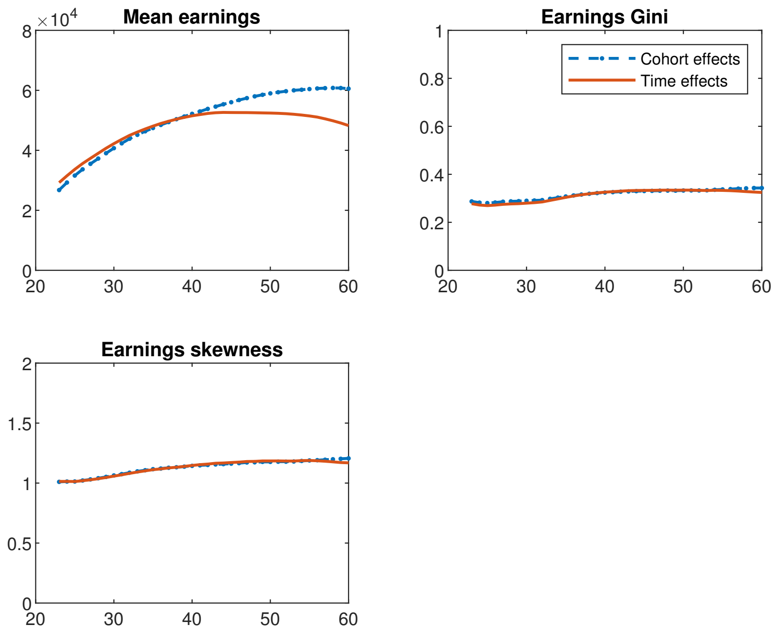


Figure 1: Mean, dispersion, and skewness of earnings by age. *Notes.* This figure plots the age effects in mean earnings, earnings dispersion measured by the Gini coefficient, and earnings skewness measured by the ratio of mean to median after controlling for time effects (red lines) or cohort effects (blue dashed lines) based on IR data, 2006-2015.

hump shaped, indicating that on average, earnings significantly increase from young ages to the 40s - 50s before slowing down and eventually declining at old ages. Second, earnings dispersion and skewness increase with age over most of the working lifetime in both the time and cohort effects model, with earnings dispersion increasing by about 0.1 and earnings skewness increasing by about 0.2 from age 23 to 60.

The estimated age effects from two models generally agree with each other. However, the time effects model implies lower age effects than the cohort effects model does, especially at older ages, which might suggest that there are patterns or characteristics associated with birth cohorts that influence the relationship between age and the corresponding earnings statistic. These patterns or characteristics might include economic events (such as recessions or periods of economic growth), technological advancements, or policy changes related to education, labor markets, or social welfare. We use the cohort effects model as the benchmark as the model provides a better fit with age profiles generated by the cohort effects model.

3.2 Human capital shocks

Following Huggett et al. [2011], we estimate the standard deviation of human capital shocks from wage data. The wage can be written as follows:

$$\begin{aligned} wage_t &= w_t h_t \\ &= w_t H(h_{t-1}, s_{t-1}, a^i, z_t) \\ &= w_t \exp(z_t)\left[h_{t-1} + a^i (h_{t-1} s_{t-1})^{\alpha}\right] \end{aligned}$$

The change in wage $wage_t$ is determined by rental rate w_t and human capital h_t which in turn is determined by the human capital shock process $\exp(z_t)$, the previous level of human capital h_{t-1}, the amount of time spent on learning in the previous period s_{t-1}, and the given learning ability a. The form of the human capital production implies that human capital is unchanged if there is no time spent on learning (i.e. $h_t = h_{t-1}$ when $s_{t-1} = 0$). Therefore, for individuals who spend no time on learning the change in wage is only determined by the rental rate and human capital shock process. This suggests that the human capital shock can be estimated from wage data including only individuals who spend no time on learning.

How to obtain a sample with only individuals investing no time on learning? From the model perspective, people spend time learning as a form of investment because they expect higher future earnings coming from human capital production. Young people have a higher incentive to learn because they have a long working lifetime ahead and can enjoy the returns from human capital investment. Older people, especially those approaching retirement, do not have such incentive to learn because they cannot enjoy the returns from human capital investment for long and their earnings are zero after retirement. It is thus reasonable to assume that older people do not spend time on learning, and we collect data on individuals age 55-65 for the purpose of estimating the human capital shock. Younger individuals are more likely to violate the assumption, and we use a slightly younger sample (age 50-60) for sensitivity analysis.

The following equations explain our estimation model. By assuming that an individual spends no time on learning from period t through $t+n$ the first equation defines the wage $wage_{t+n}$ as the product of the rental rate w_{t+n}, the shock process $z_{t+1}, \ldots, z_{t+n}$, and human capital h_t. The second equation takes logs of the first equation, and the third equation defines the n-period log wage differences $y_{t,n}$ as the sum of rental rate differences, human capital shocks, and measurement error differences $\epsilon_{t+n} - \epsilon_t$.

$$wage_{t+n} = w_{t+n}h_{t+n} = w_{t+n}H(h_{t+n-1}, 0, a^i, z_{t+n}) = w_{t+n}\prod_{i=1}^{n}\exp(z_{t+i})h_t$$

$$\ln wage_{t+n} = \hat{w}_{t+n} + \sum_{i=1}^{n} z_{t+i} + \hat{h}_t$$

$$y_{t,n} = \ln wage_{t+n} - \ln wage_t + \epsilon_{t+n} - \epsilon_t = \hat{w}_{t+n} - \hat{w}_t + \sum_{i=1}^{n} z_{t+i} + \epsilon_{t+n} - \epsilon_t$$

We assume that both human capital shocks z_t and measurement errors ϵ_t are independent and identically distributed over time and people, with $z_t \sim N(\mu, \sigma_z^2)$ and $\text{var}(\epsilon_t) = \sigma_\epsilon^2$. The log wage difference equation and these assumptions imply the three cross-sectional moment conditions below:

$$E[y_{t,n}] = \hat{w}_{t+n} - \hat{w}_t + n\mu$$
$$\text{var}(y_{t,n}) = n\sigma_z^2 + 2\sigma_\epsilon^2$$
$$\text{cov}(y_{t,n}, y_{t,m}) = m\sigma_z^2 + \sigma_\epsilon^2 \text{ for } m < n$$

We use the same dataset as section 3.1 to estimate human capital shocks. After converting nominal earnings to real earnings, we select data for individuals who (i) are aged from 55 to 65; and (ii) earn labor income of at least \$9,500 a year (in 2006 prices). We further restrict our sample by selecting individuals who have earnings recorded in at least four consecutive years, and then calculate three log earnings differences (i.e., $y_{t,n}$ for $n = 1, 2, 3$) in the period from 2006 to 2015. Table 2 shows how the number of observations changes after every restriction criteria.

Sample	Number of observations
Income from Wages and Salaries (Linked IR data 2006-2015)	348,913
Select age range (55-65)	58,429
Select earnings range	51,071
4 consecutive years	41,834

Table 2: Number of observations

From our sample data of log wage differences, we calculate the sample analog to five moments for each year:

$$\frac{1}{N_t}\sum_{i=1}^{N_t}(y_{t,n}^i - \mu_{t,n})^2; n = 1, 2, 3$$
$$\frac{1}{N_t}\sum_{i=1}^{N_t}(y_{t,n}^i - \mu_{t,n})(y_{t,1}^i - \mu_{t,1}); n = 2, 3$$

where $\mu_{t,n} = \frac{1}{N_t}\sum_{i=1}^{N_t} y^i_{t,n}; n = 1, 2, 3$. We stack the moments across years and use a two-iteration efficient General Method of Moments estimation with an identity matrix as the initial weighting matrix.

Model	Age range	Period	N	σ_z	SE(σ_z)	σ_ϵ	SE(σ_ϵ)
1	55-65	2006-2015	41,834	0.160	0.003	0.094	0.004
2	50-60	2006-2015	53,822	0.153	0.002	0.105	0.003
3	55-65	2003-2015	51,853	0.160	0.003	0.096	0.004
4	50-60	2003-2015	69,321	0.154	0.002	0.106	0.003
5	55-65	2006-2015	41,834	0.160	0.003	0.094	0.004
6	50-60	2006-2015	53,822	0.153	0.002	0.105	0.003
7	55-65	2003-2015	51,853	0.160	0.003	0.096	0.004
8	50-60	2003-2015	69,321	0.154	0.002	0.106	0.003

Table 3: Estimation of human capital shocks. *Notes:* The table reports the estimates for σ_z and σ_ϵ for various samples and model specifications. The second column provides the minimum and maximum age in the sample, the third column specifies the time period, and the fouth column (labeled N) provides the sample size. Columns labeled SE refer to standard errors.

Table 3 provides the estimation results for σ_z and σ_ϵ[12] using various samples and model specifications. Model 1 is our benchmark model which includes individuals age 55-65 in the period 2006-2015. The point estimate for human capital shocks is $\sigma_z = 0.160$ so that a one standard deviation shock moves wages by about 16 percent. When we alter the sample to include slightly younger individuals (model 2) the point estimate is $\sigma_z = 0.153$. As younger individuals might have incentive to learn so that they can have higher earnings in the future, this younger sample is more likely to violate the critical assumption which is that no time is spent on learning.

Table 3 also provides results for sensitivity analysis in two directions. First, by widening the time period, we obtain a bigger sample (model 3) and the point estimate of σ_z almost remains unchanged. Second, we control for education and age before calculating (log) wage differences (model 5, and model 7 with longer time period), and find that the point estimates of σ_z change negligibly [13]. The use of slightly younger samples (model 4, 6, and 8) leads to lower point estimates of σ_z and provides a comparison with the

[12] The measurement errors ϵ_t are not of interest, but the point estimates of σ_ϵ are significantly lower than the estimates by Huggett et al. [2011], which points to the difference between administrative data and survey data.

[13] As an individual's wage is highly likely to be influenced by his/her education and age, we employ an additional model to separate this influence:

$$\ln wage = \beta_1 + \beta_2.educ + \beta_3.age + \beta_4.\frac{age^2}{100} + \beta_5.\frac{age^3}{1000} + u$$

We then use residuals from this model as input to calculate (log) wage differences.

benchmark model.

3.3 Income tax function

This section discusses features of the New Zealand income tax system. We define the effective tax rate as the ratio of taxes actually paid to taxable income. An effective income tax function is the empirical relationship between the effective tax rate and economic income. We first document features of the effective income tax function in New Zealand using tax data from Inland Revenue.

To fit the model setup, which focuses on individuals receiving labor earnings and paying income tax, we limit our research to labor income and income tax only. We first select individuals who receive earnings from wages/salary and are subject to payroll tax, and then calculate the effective tax rate as follows:

$$\tau^{\text{eff}} = \frac{T}{E}$$

where τ^{eff} is effective tax rate, T is PAYE tax deductions on earnings from wages/salary only[14], and E is earnings from wages/salary.

We also calculate the statutory tax rate, which refers to the tax rates explicitly specified by the tax laws for different income brackets. Since each individual can have several statutory tax rates depending on how many income brackets their earnings fall within, we calculate the average statutory tax rate for each individual instead. The calculation has two steps. First, we calculate the statutory tax amount they owe based on the following formula:

$$T^{\text{stat}} = \sum_i \tau_i^{\text{stat}} \cdot \max\Big(0, \min(E, E_{i+1}) - E_i\Big)$$

where T^{stat} is the statutory tax amount, τ_i^{stat} is the statutory tax rate for the i-th income bracket, E is earnings from wages/salary, E_i and E_{i+1} are the lower and upper limits of the i-th income bracket, respectively. Second, we calculate the average statutory tax rate for each individual as following

$$\tau^{\text{stat}} = \frac{T^{\text{stat}}}{E}$$

where τ^{stat} is the (average) statutory tax rate for each individual.

[14]Note that the tax amounts have not been adjusted for Family Support Tax Credits, and thus the effective tax rate tends to overstate the actual tax burden for individuals eligible for Family Support Tax Credits. This overestimation is particularly relevant for individuals in low- and middle-income groups, leading to effective tax rates that appear higher than the corresponding statutory tax rates within those income brackets.

To avoid confusions from different tax reforms[15], we focus on the tax schedule effective from 1 October 2010 to 31 March 2021[16] and limit our study period from 2011 to 2015. We further restrict our sample by capping annual earnings at $500,000 (in 2011 prices) because more than 90 percent of the population earn less than that. Our final sample contains over 11 million of observations (annual frequency).

To capture the relationship between income and effective tax rates, we divide the sample into 20 income groups and calculate the average of effective tax rates for each group. We do the same for statutory tax rates and get the average of statutory tax rates for each group. Figure 2 plot these results along with boxplots which represent the distributional features of effective tax rates for each income group. The bottom of the box represents the first quartile ($Q1$), and the top of the box represents the third quartile ($Q3$). The line at the middle of the box is the median or second quartile ($Q2$). The distance between $Q3$ and $Q1$ is the interquartile range IR (i.e. $IR = Q3 - Q1$), and each whisker extends to $\pm 1.5 \times IR$.

To understand Figure 2, consider the first income group (0-25 thousand dollars) as an example. The first and third quartiles of effective tax rates are about 13 and 17 percent, respectively, giving an interquartile range of 4 percent. The median effective tax rate is about 15 percent, positioned in the middle of the box. The lowest whisker extends to about 7 percent, which equals $Q1 - 1.5 \times IR$, while the highest whisker reaches about 23 percent, which equals $Q3 + 1.5 \times IR$. The mean effective tax rate is just over 15 percent, represented by an asterisk, and the mean statutory tax rate is about 11 percent, represented by a red line. These numbers indicate that: (i) the mean effective tax rate is higher than the mean statutory tax rate; (ii) effective tax rates vary significantly between individuals (as shown by the long interquartile range and whiskers' lengths); and (iii) the mean and median effective tax rates are almost identical.

Figure 2 captures several key features of the New Zealand tax system. First, the average effective tax rates (asterisk points) closely track the average statutory tax rates (red lines). This alignment suggests that, on average, taxpayers experience effective tax rates in line with the statutory rates prescribed by the tax system[17]. Second, the data on effective tax rates exhibit distinctive patterns across income groups, with concentration for middle and high-income groups and greater dispersion for very low ($0-$25,000)

[15] which happened in 1 April 2000, 1 October 2008, 1 April 2009, and 1 October 2010, respectively.

[16] The tax schedule in place from 2010 to 2021 includes four levels of tax rate: 10.5 percent ($0-$14,000), 17.5 percent ($14,001-$48,000), 30 percent ($48,001-$70,000), and 33 percent ($70,001 and over).

[17] As the income range of $25,000 might be large, we plot another figure using a smaller income range for each group (i.e., $1,000) and find the same observation. This figure is in the Appendix.

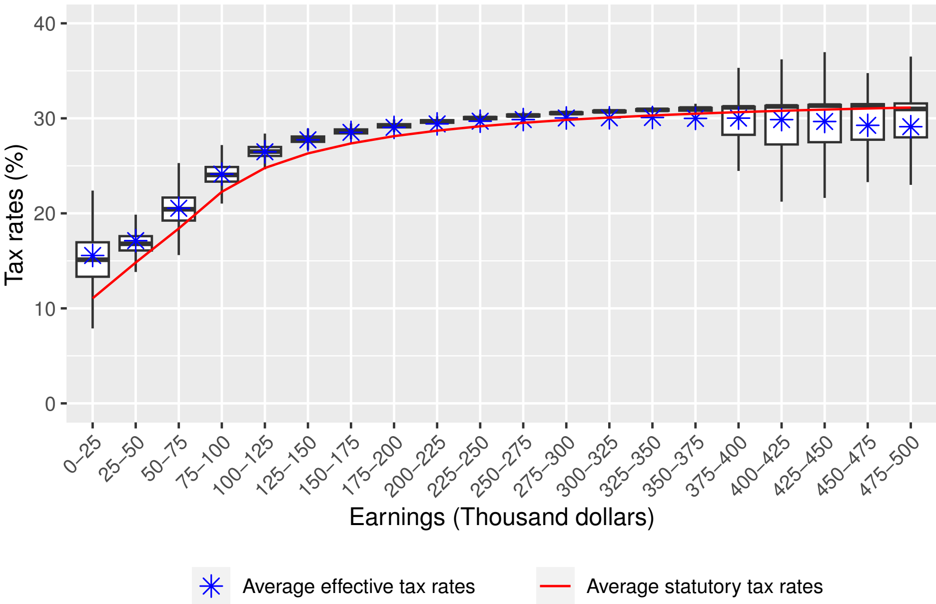


Figure 2: Average tax rates paid by tax payers in New Zealand. *Notes.* This figure plots the average effective tax rates (asterisk points) together with the average statutory tax rates (red lines) based on IR data, 2011-2015. Boxplots represent the distributional features of effective tax rates for each income group. The bottom of the box represents the first quartile ($Q1$), and the top of the box represents the third quartile ($Q3$). The line at the middle of the box is the median or second quartile ($Q2$). The distance between $Q3$ and $Q1$ is the interquartile range IR (i.e. $IR = Q3 - Q1$), and each whisker extends to $\pm 1.5 \times IR$.

and very high (over $400,000) income groups. While the concentration indicates a clustering of effective tax rates within a relatively narrow range, the dispersion signals a broader distribution of tax burdens within these extremities. Third, in terms of central tendency, mean and median closely align for most income groups, indicating symmetric or nearly symmetric distributions, but in very high-income groups (over $400,000), a leftward skew is observed, with the mean falling below the median and the median closer to the upper box limit.

In most countries there is a large difference between statutory and effective tax rates, and models therefore use effective tax rates. This has prompted the development of various methodologies for estimating the effective tax function[18]. But in New Zealand the effective and statutory tax rates are closely aligned. We therefore skip estimating an effective tax function and directly use the statutory taxes in our model. Using the statutory tax function also makes it straightforward to understand different tax levels

[18]See, e.g., Gouveia and Strauss [1994], Huggett and Parra [2010], or Guner et al. [2014] for the USA; García-Miralles et al. [2019] for Spain; Li and Ma [2017] for China; Lim and Hyun [2006] for Korea and several countries from the Luxembourg Income Study (LIS) dataset.

and changes in fiscal policy, which is an important advantage.

4 Setting model parameters

All parameters in our model are listed in Table 4. These parameters are divided into two groups. First, the parameters in group A are calibrated without needing to solve the model. Then the parameters in group B are jointly calibrated so that the model generates results which best match several moments of the data.

Category	Symbol	Parameter value
Group A		
Demographics	(J, J_R, n)	$J = 63; J_R = 43; n = 0.0124$
Preferences	$u(c) = \frac{c^{1-\sigma}}{(1-\sigma)}$	$\sigma = 2$
	$E\sum_{j=1}^{J} \beta^{j-1} u(c_j)$	$\beta = 0.981$
Wage growth	$w_{t+1} = w_t(1+g)$	$g = 0.01$
Tax system	$T_j = T_j^{inc} + T_j^{cons}$	T_j^{inc} - see text
		$T_j^{cons} = 0.15c_j$
Human capital shocks	$z_t \sim N(\mu, \sigma_z^2)$	$\mu = -0.016; \sigma_z = 0.160$
Group B		
Human capital technology	$h' = H(h, s, a, z')$	$\gamma = 0.4184$
	$H(h, s, a, z') = \exp(z')(h + a(hs)^\gamma)$	
Initial distribution of h and a	$\Phi = LN(\mu_x, \Sigma)$	$\mu_x = (\mu_{logh}, \mu_{loga}) = (7.3327, 0.4042)$
		$(\sigma_{logh}^2, \sigma_{loga}^2, \sigma_{loghloga}) = (0.2417, 0.9519, -0.4796)$

Table 4: Parameter values: Benchmark model

4.1 Group A

This group includes parameters governing demographics, preferences, technology, the tax system, and shock process.

Demographics The time period in the model is a year. An agent lives for $J = 63$ periods, which corresponds to a real-life age of 23 to 85. The agent retires at age $J_R = 43$ or a real-life age of 65. The population growth rate is set to $n = 0.0124$. This is the average population growth rate in New Zealand over 1991-2022[19] from Statistics New Zealand[20].

Preferences The risk aversion/intertemporal substitution parameter is set to $\sigma = 2$. The discount factor is set to $\beta = 0.981$ (following Huggett et al. [2011]).

Wage growth The growth rate of wage $g = 0.01$ is set to match the average growth rate of mean real earnings in IR data during 2006-2015.

[19]The average population growth rate over 2006-2015 is 1.15 percent.

[20]available at https://www.stats.govt.nz/topics/population

Tax system The tax revenue T_j comes from an income tax and a consumption tax: $T_j = T_j^{inc} + T_j^{cons}$. The income tax in the model captures the pattern of effective average income tax rates in New Zealand. Based on the empirical analysis in Section 3.3, we directly use the statutory tax rates as the income tax function. The tax rates and income brackets in the tax function come from the tax reform package which took effect from 1 October 2010. The consumption tax charges a flat rate of 15 percent on goods consumed, which captures the goods and services tax (GST) in New Zealand.

Shock process The human capital shock is assumed to follow a normal distribution: $z_t \sim N(\mu, \sigma_z^2)$. The standard deviation of human capital shocks is set to $\sigma_z = 0.160$ based on the estimate from Table 3. The mean human capital shock is set to $\mu = -0.016$ so that the model matches the average rate of decline of mean earnings for the cohorts of older workers (age 55-62) in IR data documented in Figure 1 [21].

4.2 Group B

This group includes parameters governing the human capital technology and the distribution of initial conditions. We choose these parameters to get the model to match the life-cycle profiles for mean earnings, the Gini coefficient of earnings, and the mean-to-median ratio of earnings, as seen in the New Zealand data in Section 3.1.

We set the parameters governing the elasticity parameter γ and the distribution Φ to minimize the squared distance of log model moments from log data moments. Given an initial guess of these six parameters $(\gamma, \mu_{logh}, \mu_{loga}, \sigma^2_{logh}, \sigma^2_{loga}, \sigma_{loghloga})$[22], we solve the life cycle model and simulate to find the model moments for mean earnings, variance of (log) earnings, and earnings skewness at age j. The corresponding data moments are obtained from Section 2.1.

The objective of the minimization problem is

$$\sum_{j=1}^{Jr-1} \left[\left(log(\frac{m_{1j}}{d_{1j}}) \right)^2 + \left(log(\frac{m_{2j}}{d_{2j}}) \right)^2 + \left(log(\frac{m_{3j}}{d_{3j}}) \right)^2 \right]$$

where (m_{1j}, m_{2j}, m_{3j}) denote model moments for mean earnings, variance of (log) earnings, and earnings skewness at age j; (d_{1j}, d_{2j}, d_{3j}) denote the corresponding data moments.

[21] The average rate of decline of mean earnings is highly sensitive to the selected age range. By selecting individuals aged 55-61 the average rate of decline of mean earnings for the cohorts of older workers is -1.4 percent (implying $\mu = -0.014$).

[22] We come up with a good initial guess after several calibrations. Using a good initial guess helps reduce the running time of the calibration code.

The calibrated values of $(\gamma, \mu_{logh}, \mu_{loga}, \sigma^2_{logh}, \sigma^2_{loga}, \sigma_{loghloga})$ are provided in Table 4.

5 Properties of the benchmark model

In this section, we analyze the ability of the benchmark model to generate the earnings facts documented in section 3.1.

Dynamics of the earnings distribution Figure 3 plots the age profiles of mean earnings, earnings dispersion and earnings skewness produced by the benchmark model. The model performs reasonably well in replicating earnings dynamics. For mean earnings, it produces an increasing earnings profile that closely aligns with the mean earnings observed during the mid-career phase, although it does not fully capture the characteristic hump shape. Regarding earnings dispersion, the model accurately reflects dispersion levels during early and mid-career stages but overestimates dispersion at older ages. In terms of earnings skewness, the model consistently generates higher skewness across all age groups.

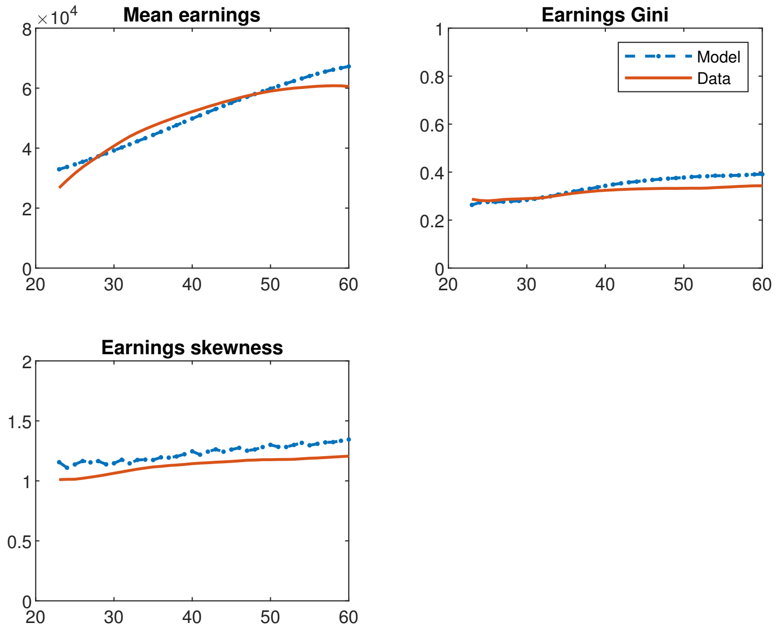
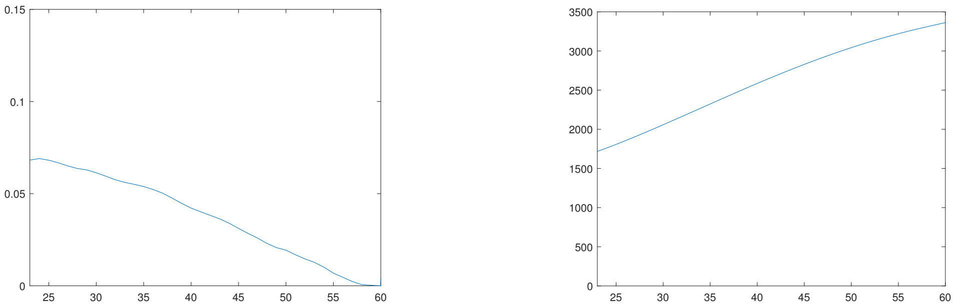


Figure 3: Mean, dispersion, and skewness of earnings by age: Model versus Data *Notes.* This figure plots the age effects in mean earnings, earnings dispersion measured by the Gini coefficient, and earnings skewness measured by the ratio of mean to median after controlling for cohort effects generated by IR data (red lines) or model (blue dashed lines).

Figure 4 presents the age profiles of the mean fraction of time allocated to human capital production and the mean human capital levels underpinning the earnings dynamics. Figure 4a illustrates that approximately 7 percent of available time is dedicated to human capital production at age 23, with almost no time allocated to this activity beyond age 58. Figure 4b indicates that mean human capital increases with age. The model implies that early in the working life cycle, individuals allocate more time to human capital production compared to later stages. Higher ability individuals tend to devote a greater fraction of their time to human capital production. These time allocation decisions facilitate human capital accumulation in the early part of the working life cycle. Subsequently, increases in human capital result from positive shocks and the time allocation decisions of high-ability individuals. Given the significant variation in ability, most low-ability individuals spend minimal time on human capital production and experience a decline in human capital over their careers. Conversely, a few extremely high-ability individuals invest about 20 percent of their time in human capital production, leading to substantial human capital accumulation. Consequently, the average human capital level increases with age.



(a) Mean time in human capital (b) Mean human capital

Figure 4: Properties of human capital by age. *Notes.* This figure plots the mean time spent on human capital (left figure), and the mean human capital (right fiture) by age. The figures plot the human capital properties for the model that fits the cohort effects view.

Risk versus Ability differences Figure 3 indicates that both the model and the data generate an increasing earnings dispersion. From the model's perspective, this rise in earnings dispersion arises from two primary sources. The first source is human capital shocks, which repeatedly impact individuals throughout their working life cycle. Since the human capital level at any given period is influenced by the cumulative shock process up to that period[23] the accumulated shocks play a significant role in generating increasing dispersion in human capital and earnings as a cohort ages. The

[23]The next period human capital is given by

$$\begin{aligned} h_{t+n} &= \exp(z_{t+n}) H(h_{t+n-1}, s_{t+n-1}, a) \\ &= \exp(z_{t+n}) H\big(\exp(z_{t+n-1}) H(h_{t+n-2}, s_{t+n-2}, a), s_{t+n-1}, a\big) \\ &= \dots \end{aligned}$$

second source is the variation in learning ability among individuals within an age group. Individuals with higher learning ability devote more time to human capital production and consequently accumulate more human capital over their careers, resulting in higher earnings later in life compared to their lower-ability counterparts. Mean earnings profiles of high-ability individuals are thus steeper than those of low-ability individuals.

We analyze the quantitative importance of risk and ability differences for generating the increase in earnings dispersion in the benchmark model by conducting two experiments. In these experiments, we only change either idiosyncratic shocks or ability while keeping other parameters constant.

First, we eliminate idiosyncratic shocks to human capital by setting $\sigma_z = 0$. Figure 5 illustrates that the model generates decreasing profiles for both mean earnings[24] and earnings dispersion. When idiosyncratic risk is removed, human capital accumulation becomes unattractive, leading all individuals to opt out of spending time on human capital production. Consequently, human capital depreciates deterministically as individuals age, resulting in declining earnings. This dynamic leads to lower mean earnings and reduced earnings dispersion later in life.

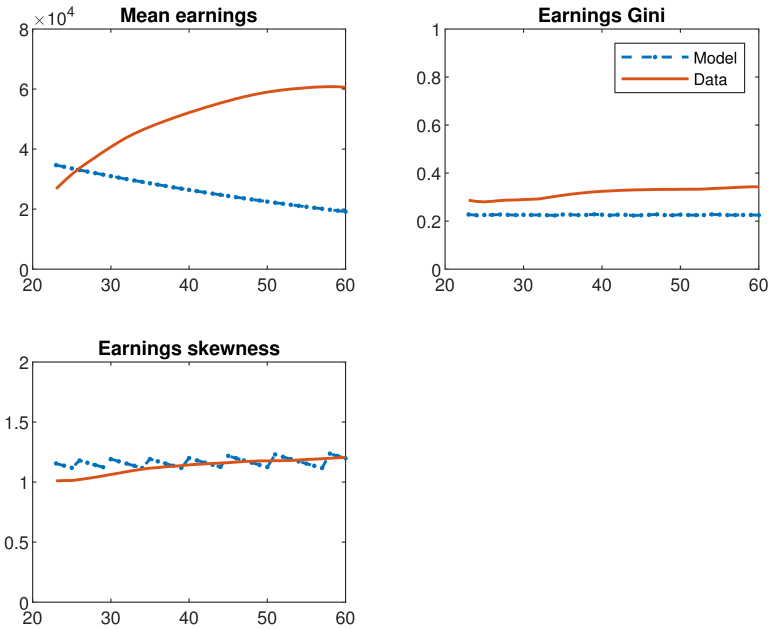


Figure 5: Mean, dispersion, and skewness of earnings by age: Model without idiosyncratic shocks versus Data *Notes.* This figure plots the age effects in mean earnings, earnings dispersion measured by the Gini coefficient, and earnings skewness measured by the ratio of mean to median after controlling for time effects generated by IR data (red lines) or model (blue dashed lines).

Second, we eliminate ability differences by changing the initial distribu-

[24] The expected value of $\exp(z)$ is $E[\exp(z)] = \exp(\mu + \frac{\sigma_z^2}{2})$, so $\sigma_z = 0$ is also imposing the negative drift, which contributes to the decreasing mean earnings.

tion so that all agents have the same learning ability, which we set equal to the mean ability. Figure 6 shows that the model generates too much earnings dispersion. As individuals now have the same learning ability, what makes their future earnings different from others is idiosyncratic risk. Individuals receiving lots of good shocks have higher human capital and higher earnings, while individuals receiving lots of bad shocks have lower human capital and lower earnings. As we show below, the model implies a negative correlation between learning ability and initial human capital which supports the relatively flat earnings dispersion profile. Absent differences in learning ability, there is no longer the force to prevent earnings dispersion from rising fast later in life. As the result, the model significantly overestimates earnings dispersions at older ages.

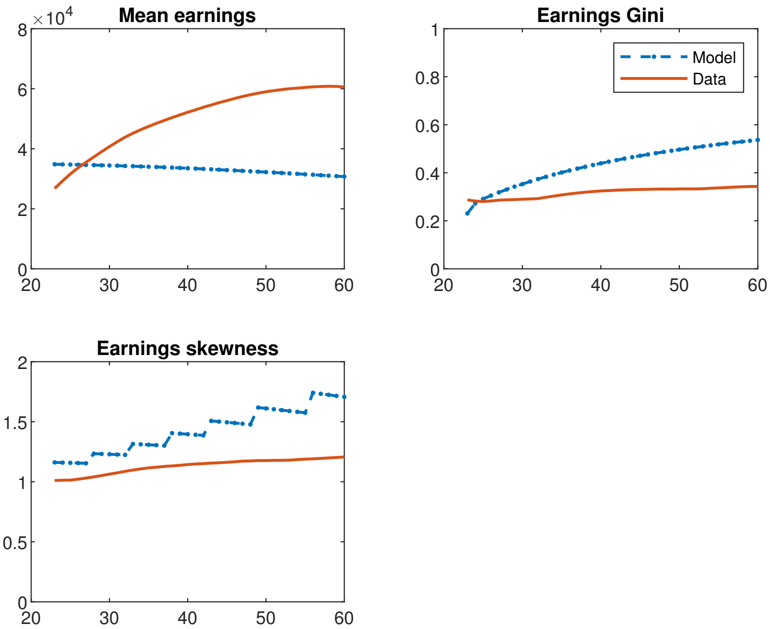


Figure 6: Mean, dispersion, and skewness of earnings by age: Model without ability differences versus Data *Notes.* This figure plots the age effects in mean earnings, earnings dispersion measured by the Gini coefficient, and earnings skewness measured by the ratio of mean to median after controlling for time effects generated by IR data (red lines) or model (blue dashed lines).

Properties of the initial distribution Table 5 summarizes the properties of the distribution of initial conditions. A notable feature of the model is that (log) human capital and (log) ability are negatively correlated at age 23, which contrasts with the findings of Huggett et al. [2011]. This negative correlation implies that high-ability individuals are born with low initial human capital, while low-ability individuals are born with high initial human capital. High-ability individuals, by devoting a larger fraction of their time to human capital production, accumulate more human capital and subsequently earn higher future earnings compared to their low-ability counterparts. If there were a zero or positive correlation between learning ability and human capital at age 23, the model would predict a much steeper dispersion profile, as high-ability individuals would quickly overtake the earnings of low-ability individuals. This increasing dispersion of earnings profile is precisely what is seen in U.S. data, and hence Huggett et al. [2011] found a positive correlation. Given that the earnings dispersion documented for New Zealand in Figure 1 does not support a steep dispersion profile over ages 23-60, the model accounts for this observation by incorporating a negative correlation between learning ability and human capital at age 23.

Statistic	Value
Mean log learning ability ($\ln a$)	0.4043
Coefficient of variation ($\ln a$)	2.4133
Mean log initial human capital ($\ln h_1$)	7.3327
Coefficient of variation ($\ln h_1$)	0.0671
Correlation ($\ln a, \ln h_1$)	-0.9998

Table 5: Properties of the distribution of initial conditions. *Note.* Entries show properties of the distribution of initial conditions for the parameters that best match the profiles of mean earnings, earnings dispersion, and skewness.

6 Effects of consumption and income tax changes

In this section, we explore the effects of changes in the level and the composition of taxes. The experiment involves altering the rates of consumption and income taxes, either individually or simultaneously, while holding other factors constant. This approach enables us to analyze how different tax policies influence individuals' incentives for labor supply, savings, and human capital investment as well as the impact of these policies on earnings dynamics and income inequality.

Changing consumption tax We derive the first order conditions to analyze effects of changing the consumption tax[25]. Consumption tax only enters the first order condition with respect to consumption

$$\beta^{j-1} c_j^{-\sigma} = \lambda_j (1 + \tau^c)$$

where τ^c is consumption tax rate, λ_j is the Lagrangian multiplier.

An increase in consumption tax rate τ^c directly affects the right-hand side of the equation. For the marginal utility of consumption $\beta^{j-1} c_j^{-\sigma}$ to balance the equation, consumption c_j must decrease. Therefore, higher consumption tax leads to a decrease in consumption. Since consumption tax

[25]Details are provided in Appendix

does not enter the first order conditions with respect to k_{j+1}, l_j, s_j, changes in consumption tax do not directly influence savings, labor supply, or human capital investment.

The interpretation is that consumption tax influences the immediate consumption decision by making it more expensive to consume, leading to a reduction in current consumption, while leaving the relative prices of labor, savings, or investments in human capital unchanged. Individuals may adjust their consumption patterns in response to changes in the consumption tax rate without altering their work or savings behavior.

Figure 7 plots the age profiles of asset holdings, consumption, labor supply, and earnings under three scenarios: (i) Benchmark model; (ii) High consumption tax; and (iii) Low consumption tax. It confirms that the primary adjustment mechanism is through consumption, while other economic activities remain unaffected.

Figure 7: Age profiles with different consumption tax rates *Notes.* This figure plots the age profiles of asset holdings, consumption, labor supply, and earnings at three levels of consumption tax: (i) Benchmark model; (ii) High consumption tax; (iii) Low consumption tax.

Changing income tax Recall that the income tax amount in our model is calculated as

$$T = \sum_i \tau_i^{\text{stat}} \cdot \max\Big(0, \min(E, E_{i+1}) - E_i\Big)$$

where τ_i^{stat} is the statutory tax rate for the i-th income bracket, E is earnings from wages/salary, E_i and E_{i+1} are the lower and upper limits of the i-th income bracket, respectively.

Utilizing the statutory tax function offers clear advantages for policy experiments. Adjusting the tax rates for higher income brackets while keeping other rates constant can create a more progressive income tax schedule. Conversely, reducing rates for higher brackets makes the tax schedule less progressive. Increasing or decreasing all tax rates by a uniform amount or proportion adjusts the overall tax burden while maintaining the progressivity of the tax system. Modifying the income brackets themselves can also change the level or progressivity of the tax schedule. These adjustments are straightforward to implement and understand, avoiding the complexities involved in estimated income tax functions, which typically are difficult to relate to actual tax rates.

The first order conditions with respect to time spent on working l_j and time spent on human capital production s_j[26] indicate that changes in income tax affect the net wage for labor income and the return on human capital investment, leading to changes in labor supply and human capital accumulation. Higher income tax typically reduces the net wage, decreasing the incentive to supply labor, thus lowering labor supply. Higher income tax also reduces the future after-tax return on human capital investment, leading to reduced time allocated to learning activities. Progressive income taxes accentuate this as the future tax rates are typically higher than the present one. With lower labor supply and human capital investment, overall earnings decrease, reducing disposable income, consumption, and savings.

Figure 8 plots the age profiles of asset holdings, consumption, labor supply, and earnings under three scenarios: (i) Benchmark model; (ii) High income tax; and (iii) Low income tax. By "high income tax," we mean a higher level and more progressive income tax, implemented by raising tax rates for the two highest income brackets while keeping other tax rates and all income brackets unchanged. Similarly, "low income tax" refers to a lower level and less progressive income tax.

Figure 8 indicates that changes in income tax typically affect labor supply and human capital investment, leading to changes in earnings, consumption, and savings. Note that in the benchmark model, the mean time spent on human capital (Figure 4a) is minimal[27], resulting in insignificant adjustments in labor supply and earnings. The responses of savings and consumption are more substantial, and their directions are as expected.

Simultaneous changes in consumption and income taxes To study effects of simultaneous changes in consumption and income taxes, we implement five policy combinations: (i) Benchmark model; (ii) High income

[26]Details are provided in Appendix for a model with a flat income tax rate.

[27]The mean time spent on human capital at age 23 is about 7 percent, compared to about 25 percent as in Huggett et al. [2011]

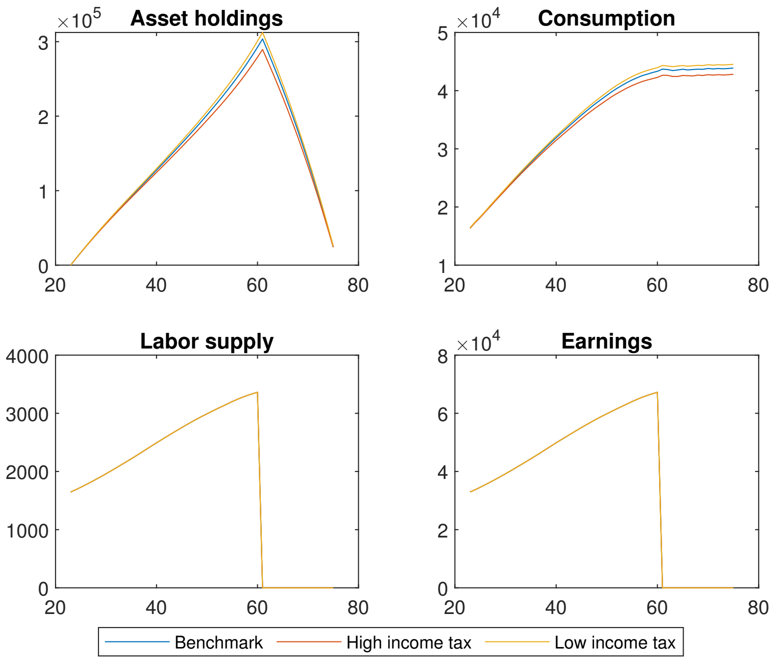


Figure 8: Age profiles with different income tax rates *Notes.* This figure plots the age profiles of asset holdings, consumption, labor supply, and earnings at three levels of income tax: (i) Benchmark model; (ii) High income tax; (iii) Low income tax.

tax, high consumption tax; (iii) Low income tax, low consumption tax; (iv) High income tax, low consumption tax; and (v) Low income tax, high consumption tax. The specifications of ”high” and ”low” are consistent with the previous experiments.

Figure 9 reveals that asset holdings, labor supply, and earnings are primarily influenced by changes in income tax, while consumption is responsive to changes in both consumption and income taxes. The highest level of consumption is achieved when both taxes are low, and the lowest level is observed when both taxes are high. This demonstrates the distinct and combined impacts of consumption and income taxes on economic behavior.

7 Conclusion

This paper provides a comprehensive analysis of the macroeconomic effects of changes in income and consumption taxes in New Zealand, utilizing a life cycle model calibrated with administrative data from the Integrated Data Infrastructure (IDI). We document that mean earnings, earnings dispersion, and earnings skewness all increase over most of the working life-cycle in New Zealand, aligning with patterns observed in the literature. We demonstrate that a human capital model can replicate these properties, given the appropriate distribution of initial human capital and learning ability.

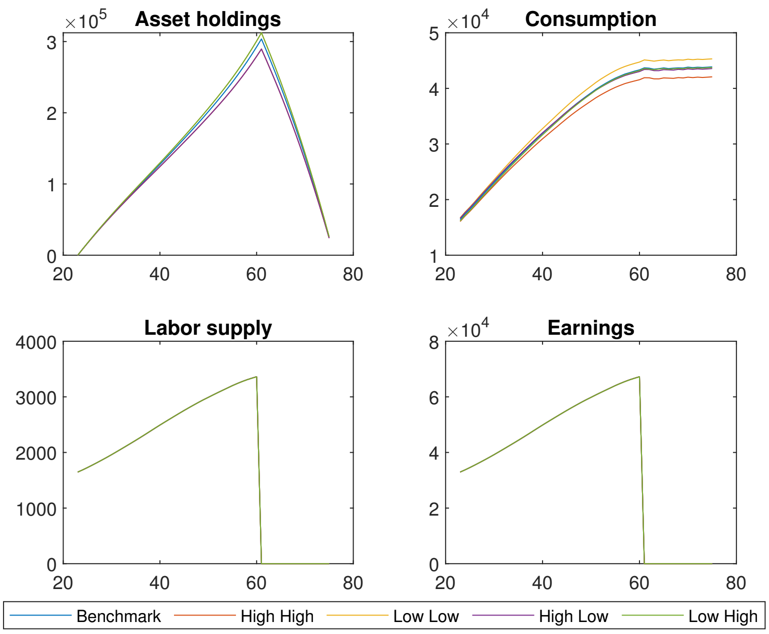


Figure 9: Age profiles with different income and consumption tax rates *Notes.* This figure plots the age profiles of asset holdings, consumption, labor supply, and earnings at five levels of income and consumption tax: (i) Benchmark model; (ii) High income tax, high consumption tax; (iii) Low income tax, low consumption tax; (iv) High income tax, low consumption tax; and (v) Low income tax, high consumption tax.

Our analysis reveals that changes in consumption taxes primarily affect immediate consumption decisions without significantly influencing labor supply, savings, or human capital investment. This is due to the direct impact of consumption taxes on the cost of consumption, which leads to a reduction in consumption levels. However, the relative prices of labor and capital remain unchanged, resulting in minimal adjustments in these areas.

Conversely, changes in income tax rates have broader implications. Higher income taxes reduce the net wage from labor, thereby decreasing incentives for labor supply and human capital investment. This leads to lower overall earnings and disposable income, subsequently reducing consumption and savings. The progressive nature of income taxes also plays a redistributive role, addressing lifetime inequality by reallocating resources from higher to lower-income individuals.

Our approach of using statutory tax rates provides clear advantages. Unlike effective tax functions, which require numerous parameters to capture the level and progressivity of taxation, statutory tax rates are straightforward to implement and understand. This simplicity enhances the model's usability for policy analysis, allowing for transparent evaluation of different tax scenarios.

Overall, this study highlights the importance of detailed data work and the benefits of using statutory tax rates in economic modeling. Our findings offer valuable insights for policymakers, emphasizing the nuanced impacts of tax policies on economic behavior.

References

J. M. Arnold. Do tax structures affect aggregate economic growth? *OECD Economics Department Working Papers*, (643), 2008. doi: https://doi.org/https://doi.org/10.1787/236001777843.

M. Borella, M. De Nardi, and F. Yang. The aggregate implications of gender and marriage. *The Journal of the Economics of Ageing*, 11:6–26, 2018.

M. Borella, M. De Nardi, M. Pak, N. Russo, and F. Yang. FBBVA Lecture 2023. The Importance of Modeling Income Taxes over Time: U.S. Reforms and Outcomes. *Journal of the European Economic Association*, 21(6): 2237–2286, 09 2023. ISSN 1542-4766. doi: 10.1093/jeea/jvad053.

A. Castaneda, J. Diaz-Gimenez, and J.-V. Rios-Rull. Accounting for the us earnings and wealth inequality. *Journal of political economy*, 111(4): 818–857, 2003.

J.-J. Chang, C.-c. Lin, and C.-C. Lai. The unemployment and wage effects of shifting to an indirect tax in an efficiency wage model. *Economic Record*, 75(2):156–166, 1999.

E. Fosse and C. Winship. Analyzing age-period-cohort data: A review and critique. *Annual Review of Sociology*, 45:467–492, 2019.

O. Galor and J. Zeira. Income distribution and macroeconomics. *The review of economic studies*, 60(1):35–52, 1993.

E. García-Miralles, N. Guner, and R. Ramos. The spanish personal income tax: facts and parametric estimates. *SERIEs*, 10:439–477, 2019.

M. Gouveia and R. P. Strauss. Effective federal individual income tax functions: An exploratory empirical analysis. *National Tax Journal*, 47(2): 317–339, 1994.

N. Guner, R. Kaygusuz, and G. Ventura. Income taxation of us households: Facts and parametric estimates. *Review of Economic Dynamics*, 17(4): 559–581, 2014.

G. Hansen and S. İmrohoroğlu. Replacing income taxation with consumption taxation in japan. *Journal of the Japanese and International Economies*, 48:15–28, 2018.

M. Huggett and J. C. Parra. How well does the us social insurance system provide social insurance? *Journal of Political Economy*, 118(1):76–112, 2010.

M. Huggett, G. Ventura, and A. Yaron. Sources of lifetime inequality. *American Economic Review*, 101(7):2923–2954, 2011.

B. Johnson and K. Moore. Consider the source: Differences in estimates of income and wealth from survey and tax data. *Special Studies in Federal Tax Statistics*, pages 77–99, 2005.

S. Y. Lee and A. Seshadri. On the intergenerational transmission of economic status. *Journal of Political Economy*, 127(2):855–921, 2019.

R. Li and G. Ma. Personal-income-tax reforms and effective-tax functions in china. *FinanzArchiv/Public Finance Analysis*, pages 317–340, 2017.

B. I. Lim and J. K. Hyun. Comparative analysis of the effective income tax function: empirical evidence using lis data. *Applied Economics Letters*, 13(2):117–121, 2006. doi: 10.1080/13504850500378262.

A. D. Nguyen, L. Onnis, and R. Rossi. The macroeconomic effects of income and consumption tax changes. *American Economic Journal: Economic Policy*, 13(2):439–466, 2021.

N. Pestel and E. Sommer. Shifting taxes from labor to consumption: more employment and more inequality? *Review of Income and Wealth*, 63(3): 542–563, 2017.

H. Yanikkaya and T. Turan. Tax structure and economic growth: Do differences in income level and government effectiveness matter? *The Singapore Economic Review*, 65(01):217–237, 2020.

Appendix

A Calibration using Time effects data

Similar to the benchmark model, this model generates an increasing earnings profile which closely matches mean earnings at the middle of working life cycle, but is unsuccessful in producing its hump shape (Figure A1).

B Age profiles using Household Economic Survey HES

In this section, we estimate age profiles of mean earnings, earnings dispersion, and earnings skewness in New Zealand using data from the Household Economic Survey (HES). HES is an annual survey which collects income,

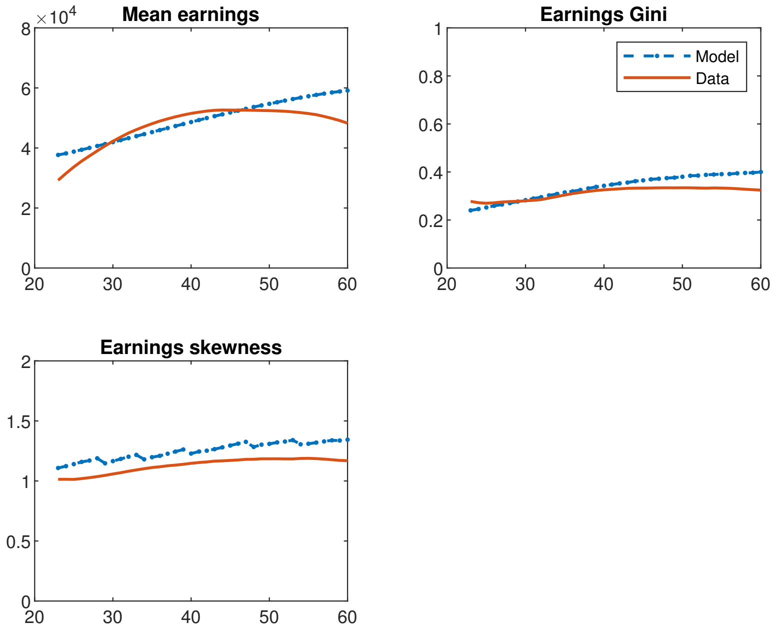
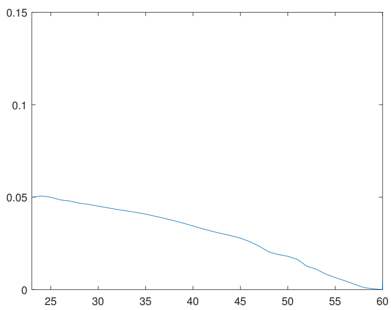
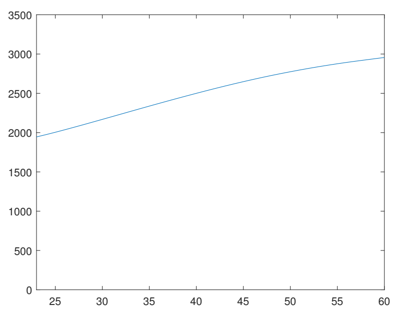


Figure A1: Mean, dispersion, and skewness of earnings by age: Model versus Data *Notes.* This figure plots the age effects in mean earnings, earnings dispersion measured by the Gini coefficient, and earnings skewness measured by the ratio of mean to median after controlling for time effects generated by IR data (red lines) or model (blue dashed lines).



(a) Mean time in human capital



(b) Mean human capital

Figure A2: Properties of human capital by age. *Notes.* This figure plots the mean time spent on human capital (left figure), and the mean human capital (right fiture) by age. The figures plot the human capital properties for the model that fits the cohort effects view.

expenditure, and demographic information on households and individuals in New Zealand. The survey covers people age 15 years and over, and the sample size is about 5,000 households or 10,000 individuals in a typical year.

From the HES Income data from 2006 to 2015, we select data on income from wages and salaries (W&S) for individuals who satisfy three selection criteria. First, they must aged between 23 and 60 at the interview time. Second, they must work at least 10 hours per week if they are 30 years old and over, or at least 5 hours per week if they are under 30. We limit

Statistic	Value
Mean log learning ability ($\ln a$)	0.5077
Coefficient of variation ($\ln a$)	1.9894
Mean log initial human capital ($\ln h_1$)	7.4739
Coefficient of variation ($\ln h_1$)	0.0609
Correlation ($\ln a, \ln h_1$)	-0.995

Table A1: Properties of the distribution of initial conditions. *Note.* Entries show properties of the distribution of initial conditions for the parameters that best match the profiles of mean earnings, earnings dispersion, and skewness.

the maximum hours worked at 100 hours per week for both age groups[28]. Third, they must earn at least \$9,500 (in 2006 prices) a year if they are 30 and older, or at least \$6,500 (in 2006 prices) a year if they are under 30. The motivations for these selection criteria are follows.

First, the HES Income data have relatively few observations at the beginning or end of the working life cycle. By selecting individuals at ages 23-60 and employing a five-year age bin[29], we have at least 94 observations in each age-year bin for use to calculate earnings statistics. Second, the traditional retirement age in New Zealand is 65[30], and there is a significant decline in the labor force participation rate near this age due to reasons such as individual preference or health conditions. As these reasons are abstracted from in the model, we use the terminal age of 60 which is below the traditional retirement age. Third, individuals in the model allocate their time to either working or learning. We therefore select data on individuals who work for at least several hours and earn some income. For individuals age 30 and over, the lower bound for hours worked is a quarter of full-time work hours, and the minimum earnings are below the annual earnings level of a full-time worker working at the New Zealand minimum wage[31]. For younger individuals, we reduce both the minimum hours and earnings to capture students doing part-time work while in school. Forth, the maximum hours worked is set at 100 hours per week to eliminate reporting errors and not-specified responses. Table A2 shows changes in the number of observations after every restriction criteria.

[28]Note that in section 3.1, we do not use this restriction. The reason is that data on hours worked only exist in HES which is a cross sectional data. The use of IR data brings us a much bigger sample, but we do not obtain data on hours worked over time.

[29]For example, to form a bin of individuals age $j = 30$ in year $t = 2008$ we use data on individuals age 28-32 in 2008. We also use a five-year age bin centered at ages 23 and 60 - the minimum and maximum ages of selected individuals - by using data on individuals age 21-25 and 58-62.

[30]As there is no legal retirement age in New Zealand, 65 is commonly considered as the traditional retirement age because this is the superannuation qualification age.

[31]In 2006, the minimum wage rate was \$10.25. The annual earnings level of a full-time worker (working 40 hours per week and 52 weeks a year) working at the minimum wage is $\$10.25 \times 40 \times 52 = \$21,320$ in 2006 prices.

Sample	Number of observations
Income from Wages and Salaries (HES Income 2006-2015)	45,823
Select age range (21-62)	38,837
Select (usual) hours worked range	31,766
Select earnings range	28,746

Table A2: Number of observations

We convert nominal earnings to real earnings[32] and group them in specific age-year bins e_{jt} which is real earnings of all individuals in age bin j in year t[33]. We then calculate four statistics for every age-year bin including mean earnings, earnings dispersion measured by variance of log earnings and Gini coefficient, and earnings skewness measured by the ratio of mean earnings to median earnings[34].

We employ the classical Age-Period-Cohort (APC) model[35] to extract age effects on these earnings statistics:

$$stat_{j,t} = \alpha_c^{stat} + \beta_j^{stat} + \gamma_t^{stat} + \epsilon_{j,t}^{stat} \tag{6}$$

where $stat_{j,t}$ is each of the four statistics, α_c^{stat} is the cohort effects, β_j^{stat} is the age effects, γ_t^{stat} is the time effects, and $\epsilon_{j,t}^{stat}$ is error terms. We wish to estimate the age effects β_j^{stat}.

Model (6) separates three factors governing the evolution of the earnings statistic: cohort (c), age (j), and time (t). As birth cohort can be derived from time and age, which is $c = t - j$, the independent variables are colinear and their effects cannot be estimated without further restrictions. We follow Huggett et al. [2011] to use two alternative approaches. The first approach is the cohort effects model where we set $\gamma_t^{stat} = 0, \forall t$ [36]. The second approach is the time effects model where we set $\alpha_c^{stat} = 0, \forall c$ [37]. We use ordinary least squares to estimate the coefficients in each model.

Figure A3 plots the age effects, i.e. β_j^{stat} estimated from model (6), in mean earnings, variance of log earnings, earnings Gini, and earnings skewness. The age effects are normalized so that each profile runs through the mean value of each statistic across panel years at age 38.

[32] We collect annual CPI for New Zealand from the World Bank data and normalize so that CPI = 100 in 2006. Real earnings are equal to earnings divided by CPI.

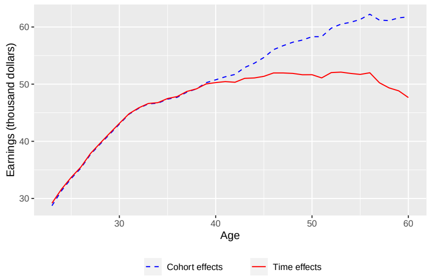
[33] As we use five-year age bin, each age-year bin includes individuals age from $(j-2)$ to $(j+2)$ in year t.

[34] As the number of age is $J = 38$ and the number of year is $T = 10$ the number of observations for each statistic is equal to the number of age-year bins which is $J \times T = 380$.

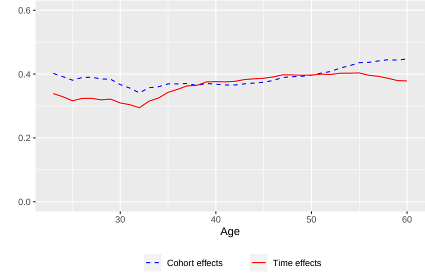
[35] For readers who are unfamiliar with APC models, we suggest the review of Fosse and Winship [2019].

[36] In each regression the dependent variable is regressed on $J + T - 1$ cohort dummies and J age dummies. The intercept is not included.

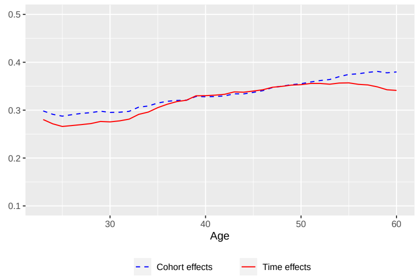
[37] In each regression the dependent variable is regressed on J age dummies and $T - 1$ time dummies. The intercept is not included.



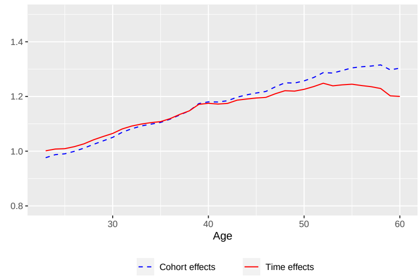
(a) Mean earnings



(b) Variance of log earnings



(c) Earnings Gini



(d) Earnings skewness

Figure A3: Mean, dispersion, and skewness of earnings by age. *Notes.* This figure plots the age effects in mean earnings, earnings dispersion measured by the variance of log earnings and the Gini coefficient, and earnings skewness measured by the ratio of mean to median after controlling for time effects (red lines) or cohort effects (blue dashed lines) based on HES data, 2006-2015.

C Average effective versus statutory tax rates

Figure A4 plots the average effective tax rates (asterisk points) together with the average statutory tax rates (red lines) based on IR data, 2011-2015 using the income range of $1,000 for each group. It can be seen that the average effective tax rates closely track the average statutory tax rates.

D First order conditions

This section provides details about the first order conditions. Let's consider the agent's maximization problem:

$$\max_{\{c_j,k_j,h_j,l_j,s_j\}_{j=1}^{J}} E\Big[\sum_{j=1}^{J}\beta^{j-1}\frac{c_j^{1-\sigma}}{1-\sigma}\Big] \text{ subject to}$$

$$c_j + k_{j+1} = e_j + k_j(1+r_{t+j-1}) - T_{j,t+j-1}(e_j, c_j), \forall j;$$
$$e_j = w_{t+j-1}h_j l_j \text{ if } j < J_R, \text{ and } e_j = 0 \text{ otherwise;}$$
$$h_{j+1} = \exp(z_{j+1})H(h_j, s_j, a);$$
$$w_{t+j-1} = (1+g)w_{t+j-2};$$
$$k_{j+1} \geq 0;\ k_{J+1} = 0;\ \text{and } l_j + s_j = 1, \forall j.$$

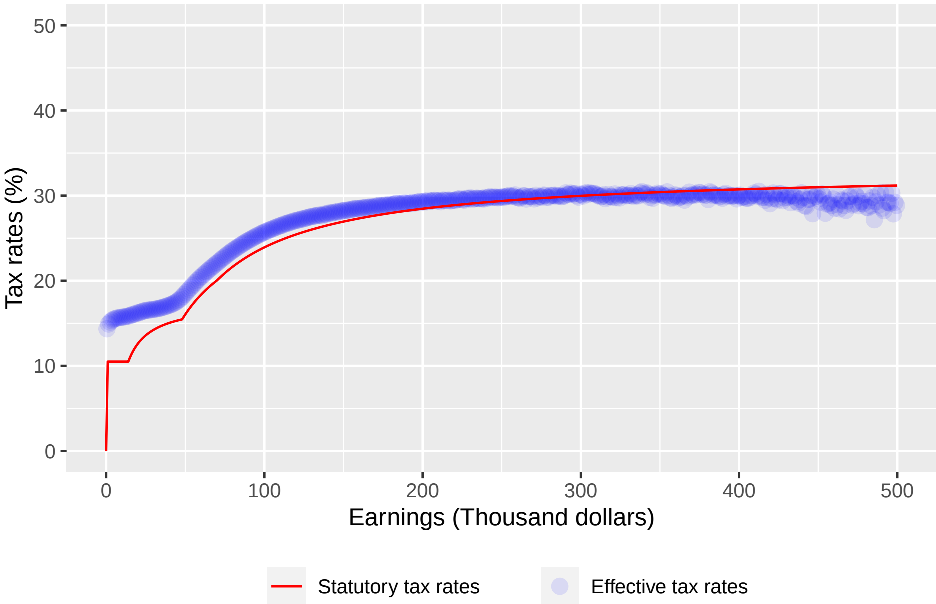


Figure A4: Average tax rates paid by tax payers in New Zealand. *Notes.* This figure plots the average effective tax rates (asterisk points) together with the average statutory tax rates (red lines) based on IR data, 2011-2015.

where c_j is consumption, k_j is asset holdings, h_j is human capital, l_j is time working, s_j is time studying, e_j is labor earnings, $T_{j,t+j-1}(e_j, c_j)$ is net taxes.

We can write $T_{j,t+j-1}(e_j, c_j) = \tau^e e_j + \tau^c c_j - \text{transfer}$ where τ^e, τ^c is income tax rate and consumption tax rate, respectively. Note that the income tax in this model is flat, meaning all agents face the same tax rate, τ^e. This simplification reduces the complexity of the notations while preserving the model's interpretative clarity.

The budget constraint can thus be expressed as:

$$c_j + k_{j+1} = e_j + k_j(1 + r_{t+j-1}) - (\tau^e e_j + \tau^c c_j - \text{transfer})$$
$$(1 + \tau^c)c_j + k_{j+1} = (1 - \tau^e)e_j + k_j(1 + r_{t+j-1}) + \text{transfer}$$

The Lagrangian is given by:

$$\begin{aligned}\mathcal{L} =& E\Big[\sum_{j=1}^{J} \beta^{j-1} \frac{c_j^{1-\sigma}}{1-\sigma}\Big] \\ &+ \sum_{j=1}^{J} \lambda_j \Big[(1 - \tau^e)e_j + k_j(1 + r_{t+j-1}) - (1 + \tau^c)c_j - k_{j+1} + \text{transfer}\Big] \\ &+ \sum_{j=1}^{J} \mu_j (1 - l_j + s_j)\end{aligned}$$

We take the first order conditions with respect to c_j, k_{j+1}, l_j, s_j:

FOC with respect to c_j:

$$\beta^{j-1} c_j^{-\sigma} = \lambda_j (1 + \tau^c)$$

FOC with respect to k_{j+1}:

$$\lambda_j = \beta \lambda_{j+1} (1 + r_{t+j})$$

FOC with respect to l_j:

$$\lambda_j (1 - \tau^e) w_{t+j-1} h_j = \mu_j$$

FOC with respect to s_j:

$$\lambda_{j+1} \Big(\frac{\partial (1-\tau^e) e_{j+1}}{\partial s_j} \Big) - \mu_j = 0$$
$$\lambda_{j+1} \Big[(1-\tau^e) w_{t+j} l_{j+1} \frac{\partial \exp(z_{j+1}) H(h_j, s_j, a)}{\partial s_j} \Big] = \mu_j$$