The Elasticity of Taxable Income in New Zealand

IRIS CLAUS, JOHN CREEDY and JOSH TENG

Appendix. The data

The database was constructed by randomly sampling Inland Revenue’s individual taxpayer population. It covers the period 1994 to 2008. The number of taxpayers in the random sample rises from 128,440 in 1994 to 162,651 in 2008. The sample is weighted to match the individual taxpayer population, which increased from 2,761,000 taxpayers in 1994 to 3,423,421 in 2008. The database includes people with wage/salary income (including taxable welfare benefits) and people who filed an IR5 or IR3 tax return or received a personal tax summary (PTS). It excludes people with no personal taxable income unless they filed.

The requirement for wage and salary earners to file an IR5 tax return was in part based on earnings over a threshold. This threshold increased from $20,000 in the early 1990s to $38,000 by 1999. However, in 2000, the IR5 tax return was replaced with the PTS, a pre-populated taxpayer square-up based on data collected from employers during the year. The income threshold was removed, with a consequential reduction in the number of taxpayers required to square up. This has caused a structural break in the income tax data collected, especially on dividend and interest income. Taxpayers who previously filed an IR5 were not required to square up via a PTS if their only income was from salaries and wages or from investments with the correct amount of tax deducted at source, or where the investment income was below a certain threshold.

Individual taxpayer information is gathered from: client registration; individual tax return IR3; PTS from 2000 onwards; salary/wage earner income tax return IR5 (pre-PTS) from 1994 to 1999; employer monthly schedule (EMS) from 2000 onwards; annual tax deduction certificate (TDC) from 1994 to 1999.

Taxpayers are broadly categorised into two groups based on a taxpayer’s ‘entity class’, a client registration feature: (a) salary/wage earner with a salary/wage (SW) entity class; and (b) other taxpayers with a non-SW entity.
class (for example, self-employed or salary/wage earners with other income from rental properties or overseas investments). The selection method is: a random 2 per cent of total salary/wage earners (the random sample is selected from the last two digits of the IRD number); a random 10 per cent of total other individual taxpayers – also based on the last two digits of the IRD number with the chosen range including the 2 per cent sample above.

A taxpayer generally has one IRD number and the same entity class. However, a minority of taxpayers could have a second IRD number due to bankruptcy or they might retain the same IRD number but change entity class over time. This means that the above selection method misses some taxpayers if they had an entity class change between non-SW and SW, or were issued a new IRD number because of bankruptcy.

The data set contains the following main variables: general variables – unique IRD number, date of birth, gender; income variables – salary/wage income, business income, estate or trust beneficiary income, interest income, dividend income, overseas income, rental income, shareholder-employee salary, partnership income, other income, taxable income; other variables – expenses, losses claimed, loss attributing qualifying company (LAQC) losses claimed.

The gender variable is determined based on the ‘title’ of a taxpayer (e.g. Mr, Miss), so some imputation is required. When no title is present, or the title is ambiguous, gender is randomly assigned. Variables for IR3 filers include: salary/wage income; interest income; dividend income; business income; estate or trust beneficiary income; overseas income; shareholder-employee salary; net rents; partnership income; other income; LAQC losses claimed; expenses; losses claimed; taxable income. For PTS/IR5 filers, variables are: salary/wage income; interest income; dividend income. For non-filer TDC/EMS individuals, only salary and wage earnings are relevant. Taxable income is the sum of all incomes less LAQC losses claimed, less expenses and less losses claimed (in that order). Taxable income is zero if a taxpayer has negative taxable income (i.e. a loss). As the focus is on income, rebates (such as for childcare, housekeeping and donations) are ignored.

The expenses variable is different from the business expenses that can be claimed in the general set of financial accounts. The expenses variable can be defined as ‘fees’ paid for professional services. This includes: a fee to someone for completing a tax return; commission on interest or dividend income; expenses incurred in earning income that has had withholding tax deducted; additional expenses incurred in earning partnership income (e.g. interest on capital borrowed to purchase a share in the partnership); interest on money borrowed to buy shares or to invest; premiums on loss of earnings insurance (income protection), provided the benefit from the insurance policy is taxable.