Australian Government
The Treasury

A model to understand university student debts and repayments

s 22 [redacted]

Treasury Working Paper[2]

2025–##

Date created: May 2025

1 Macroeconomic Analysis, Policy and Research Division, The Treasury.

2 The views expressed in this paper are those of the authors and do not necessarily reflect those of The Australian Treasury or the Australian Government.

ISBN XXXXX

Manager
Media Unit
The Treasury
Langton Crescent
Parkes ACT 2600
Email: media@treasury.gov.au

In the spirit of reconciliation, the Treasury acknowledges the Traditional Custodians of country throughout Australia and their connections to land, sea and community. We pay our respect to their Elders past and present and extend that respect to all Aboriginal and Torres Strait Islander peoples.

A model to understand university student debts and repayments

s 22

2025–##

Abstract

Introduced in 1989, Australia's Higher Education Loan Program (HELP) is now a $75 billion program which continues to evolve over time. This paper presents a new microsimulation model to estimate HELP debt repayments, built using the rich microdata available through the Person Level Integrated Data Asset (PLIDA).

First, we provide an overview of the model and our methodology to estimate students' debts, lifetime incomes, and repayments. We then use the model to simulate student debt repayments under a recent policy change: the 2021 Job-Ready Graduate (JRG) package. Compared to pre-existing policies, we find that JRG increases the median time students take to clear their HELP debt. Under JRG, we estimate that the range of repayment times between different fields of study is wider than under previous policies. Although JRG increases the total pool of bachelor debt by 12-13 per cent, more debt is held by students with lower lifetime incomes who have worse repayment prospects. After accounting for repayments, we find that the fair value of debt owed to the Government increases by 6-7. Under the parameters used in the model, we estimate that subsidy rates are about 2 per cent higher under JRG policies.

JEL Classification Numbers: ###, ###, ###
Keywords: [Click and enter keywords]

s 22

Macroeconomic Policy, Analysis and Research
Macroeconomic Group
The Treasury
Langton Crescent
Parkes ACT 2600

1. Introduction

In Australia, the cost of studying a university bachelor's degree is shared by students, who pay a student contribution, and the Federal Government, who provides a subsidy known as a Commonwealth contribution. At a bachelor level, most Australian university students do not pay their student contribution fee upfront. Instead, most students choose to defer the cost of their study through the Higher Education Loan Program (HELP).

The HELP system is a federal loan program which helps student finance their higher education studies. Under the HELP system, the federal government provides students with income-contingent, real-interest free loans. This paper focuses on modelling students' HELP loan repayments. We develop a dynamic microsimulation model of student lifetime income, which can be used to estimate their HELP debt repayments under different policy settings and assumptions.

Students who take out HELP loans only begin repaying their debt once their annual income is above a certain threshold. For most students, compulsory repayments only begin after leaving university. The current compulsory repayment schedule is progressive, and repayments are calculated as a proportion of all income.[3]

Because HELP loans are real-interest free, they provide a subsidy to students. The value of the subsidy to students (and a cost to government) is the difference between inflation, and the Government's borrowing rate and is often called the deferral rate. Further, because loans are income-contingent and written off at death, not all loans are repaid. Accounting for both these factors – the deferral subsidy, and debt which is not repaid – gives the 'fair value' of HELP loans, which is a better indication of what it is worth to government. As of 2023, the total pool of outstanding HELP debt was more than $75 billion, with an estimated fair value of about $50 billion (Department of Education, 2023).The difference between the 'face value' and 'fair value' of debt can be expressed as a percentage, often called the total subsidy rate, which indicates the fraction of debt government loses by providing HELP loans.

The HELP system (formerly known as HECS) has evolved considerably since its introduction in 1989. Successive governments have introduced changes to student contributions, compulsory repayment thresholds, the repayment schedule, how loans are indexed, and other aspects of the system (Norton, 2022). Further changes to the HELP system were proposed in 2023 by a government commissioned review, the Universities Accord. The Government has committed to implement several changes to the HELP system.

Given the size and widespread use of the HELP system, policy changes can have substantial implications for students and government. Because policy changes affect how much debt different students accrue, they can affect the total amount of loans, subsidy rates (and therefore government costs), and repayment outcomes for students. Understanding how changes to the HELP system effect these outcomes is therefore essential for effective policy development.

3 As opposed to other parts of the tax system, which are determined on a marginal basis. As taxable income increases so too does the proportion of their income debtors are required to pay. Debtors earning between $54,435 and $62,85 repay 1 per cent of their income, and those earning above $159,664 repay 10 per cent. The Federal Labour party has proposed to change this repayment schedule, from a fixed proportion of income to a marginal scheme.

Policy simulation models are a tool which can play a role to help understand the effects of changes to policy, as well as to develop reform options. This rest of this paper outlines the methodology and early results from a new policy simulation model: the Higher Education Loan Payments and Earnings Model (HELPEM).

HELPEM is a dynamic microsimulation model designed to explore how changes in higher education policy affect outcomes for bachelor students and governments. It can be used to explore common questions about policy changes, including how changes to fees or repayment rules are likely to affect students' repayment times, how repayments differ across the distribution and by field of study, and how changes affect the costs to government of providing HELP loans.

The rest of this paper is structured as follows. First, we provide an overview of the key challenges in modelling HELP repayments and some common modelling approaches used in the literature. Next, we describe the data and methodology of HELPEM. Finally, using the introduction of the 2021 Job-Ready Graduate policy changes as an example case, we demonstrate how HELPEM can be used to simulate the effects of policy changes. We explore how the policy change affected the amount of debt students hold, expected repayment times, and costs to government of providing HELP loans, compared to previous policy settings. We then briefly conclude.

2. Modelling HELP debt repayments – challenges and approaches

A key challenge of simulating HELP debt repayments is the need to estimate repayments over a very long period and across the entire distribution. Recent research has often relied on microsimulation models to estimate debt repayments because these models are well suited to estimate outcomes distributionally and over long period of time. Within this framework, there are two key requirements to estimate repayments: estimating how much debt students accrue, and how much students earn throughout their life. The subsections below provide a brief overview of microsimulation modelling, and approaches to estimating students' debts and incomes.

Dynamic microsimulation models have become the dominant modelling framework

Historically, three main types of models have been used to estimate repayments under income contingent loan programs such as the HELP system. Early work often relied on representative debtor models and cell-based models, while more current work typically relies on microsimulation approaches (Higgins 2011).

A representative debtor model includes one or few individuals who are assumed to be 'representative' of the broader population. An amount of debt is estimated for each person, and their income level is projected over their life, often using percentiles from cross-sectional data. In a similar form, 'cell-based' models group similar individuals together into 'cells', for example by age and sex. The models then make estimates of the average debt and incomes for each cell in the model to estimate repayments.

Although both frameworks are instructive, they are limited because they cannot capture the full variety of outcomes individuals face. It can also be difficult to capture the dynamic nature of earnings variation over time when people are modelled in groups, or through 'representative' individuals.

Microsimulation models are computer programs where the unit in the model is a single person (or other unit such as a business). These models combine the benefits of representative debtor and cell-based models, by simulating a cohort of individuals who are representative of the broader population at the person level. Microsimulation models can be used to estimate how different characteristics of each individual change over their lives, tracing the outcomes of each simulated person over time. Because these models can capture heterogeneous distributions of outcomes, they have become a key tool for policymakers and academics and have been used to analyse policies concerning education loans, retirement incomes, health outcomes and many others (Harding, 2007).

Microsimulation models are often separated into two classes: 'static', and 'dynamic' models. The main difference is how individuals are allowed to change and age over time. Static approaches tend to hold individuals' characteristics constant and simulate ageing by changing how much weight is placed on each person. Dynamic models allow each person's characteristics to change over time depending on the individual's history and the likelihood of different outcomes.

Although both approaches have been used to model repayments of income contingent loans, static approaches have some significant drawbacks. Recent work has mainly focused on dynamic approaches (for example, see Dearden 2015, Higgins and Sinning 2013).

Estimating university costs

Under current policies, students' university fees are charged per subject. The cost of each subject depends on its study load and the field of study. While transparent, this structure can make it challenging to estimate how much any student's bachelor's degree costs. The eventual cost of a student's degree depends on the number and mix of subjects they take. Because students often switch courses, change majors, fail units and study electives outside their home field of study, even students who are enrolled in the same course can accrue different amounts of debt.

It is not always obvious how the cost of a degree will change when there are changes to policy. For example, a fixed increase to the cost of humanities subjects, holding all else equal, will increase the total cost for individuals studying an arts degree. But the size of this increase depends on how many humanities subjects each person studies, which will vary if the student fails and repeats subjects and on their particularly pattern of study.

There are two main considerations when modelling students' debts: realism and flexibility. Estimates should be realistic, accurately reflecting the size and mix of real students' debts. But the estimation process should also be flexible enough to estimate how debts might change under different policy settings.

To our knowledge, two approaches have been used in past Australian research to estimate student debts. First, models focused on costings generally use rich administrative tax data of the HELP debt balances for everyone in the model (Antcliff and O'Neill 2009). This approach is accurate, as students' actual debts are used, but it is not flexible to test the effects of policy changes without making strict assumptions.[4]

Second, models intended to simulate the effects of policy changes have historically been limited by data availability and have instead made strict assumptions about student study patterns. For example, Higgins et al. assume that students' debts and study patterns are homogenous for people who study the same field (Higgins and Khemka 2024). Similar studies generally exclude students who do not

4 This second point is generally less of a concern for costings, but is more important for a policy simulation model.

complete their degrees. While this approach is flexible, it is unlikely to account for the realistic mix of debts students incur at university.

Estimating student incomes

Because HELP debt repayments are calculated directly from an individual's income, estimating incomes over the lifecycle is the core of any repayment model.

A key characteristic of the HELP system is that the compulsory repayment schedule is non-linear. This makes it important for a model to accurately capture not only the total amount individuals earn over their life, but also how earnings vary over time. It also means incomes cannot be averaged over years (or individuals) without introducing error.[5]

There is a large literature and long history of research dedicated to modelling earnings dynamics and earnings over the lifecycle (Hu et al 2019). Previous work has often focused on how differences in incomes between individuals, and over time, arise from temporary changes and permanent income shocks (Erjnaes and Browning 2014). These changes are often assumed to be independent of one another, but there is debate as to how these components should be modelled.

Australian literature focused on simulating estimating the earnings of university students has generally taken one of three approaches to estimating incomes in a microsimulation framework: econometric models, which are generally the most common approach and often rely on autoregressive models (i.e. ARMA) or similar approaches; simulations using functions such as copulas; and stochastic approaches based on empirical data and probability tables, for example using Markov chains (Britton et al 2019, Dearden 2019, Antcliff and O'Neill, 2009a, Higgins 2011).

All these approaches aim to capture both transient and more permanent income shocks over time, as well as how these vary across the population. These earnings functions are usually paired with models of transitions between labour force states over time. A wide variety of methods has also been explored in international literature and there is no clear consensus on how earnings should be modelled or projected, with some recent models becoming particularly complex (Altonji et al 2023). Although econometric methods are broadly favoured by recent work, it is also not clear than any one method is most appropriate and the ideal method is likely to vary depending on the population under study and the constraints of the model.

3. Data and methodology

Data

HELPEM is built in the Australian Bureau of Statistics secure Datalab environment with data from the Person level Integrated Data Asset (PLIDA). We link higher education (Hied, 2005-2021), personal income tax (PIT, 2002-2022), and 2021 Census microdata records to create two longitudinal datasets.

The first dataset includes linked higher-education records, incomes, and demographic characteristics of students who were enrolled in a bachelor's degree and left university between 2005 and 2021. This

5 For example, an individual that earns $100,000 over two years could make no compulsory repayments if $50,000 is earned in each year, both falling below the lowest repayment threshold. Alternatively, if they earned $100,000 in one year and none in the other, the individual would repay about $6000 toward their debt.

dataset provides rich detail about recent students, including demographic information and the subjects they studied at university. However, it only contains information on recent university students. We rely on this dataset to inform assumptions about university study patterns and incomes in the early years after leaving university.

The second dataset includes individuals who completed the 2021 Census questionnaire and answered questions about basic demographic information such as sex, age, and education. We restrict the sample to Australian citizens who self-reported completing a university degree at a bachelor's level or higher, and we take their field of study as the self-reported value. This dataset does not contain rich information about study patterns, because it relies on self-reported Census data instead of higher education records. We rely on this dataset to inform our assumptions about income trajectories later in life.

We restrict both datasets to Australian citizens who submitted at least one tax return between 2002 and 2022. We use taxable income as our income measure (inflated to $2025 in line with WPI). Where taxable income data is missing, we impute this as $0.

For the results presented in this paper, we categorise the field students' study at university into 13 fields (see appendix for details). Where higher-education records are available, this definition is based on the course an individual studied; where information is from the Census, we rely on the self-reported field of study.

Methodology

The model is built from four modules:

1. An entry module, which determines the number of students leaving university in each year;
2. A loans module, which provides each simulated individual in the model with a HELP debt;
3. An income module, which is used to estimate each individual income over their lifecycle, and;
4. The repayments module, which uses the students' debts, incomes, and a set of policy rules to estimate their annual HELP debt repayments and remaining HELP balances.

The entry module

The entry module estimates how many bachelor students leave university with a HELP debt each year. We estimate counts for each combination of year, by sex, field of study, and whether the student completed their degree. These estimates are used to create the base cohort of simulated individuals used in a model run.

To estimate counts, we fit simple OLS regressions to historical counts of students who left university between 2010 and 2021. We estimated trends separately for each combination of field of study, sex and whether a student completed their degree by running regressions on subsets of the data. These trends are extrapolated to the base year to give total counts for each subgroup.

A key assumption in the entry model is that students do not change their choice of university course in response to changes in university fees. Although this is a strong assumption, past research has found that price sensitivity is very low for students eligible for HELP loans, and very few students change their study patterns or choice of field even in the face of large changes to fees (Yong et al. 2023, Aungles 2002, Andrews 1997). This research is consistent with student surveys, which find that students make decisions about university based on their interest, with little regard to fees under the HELP system (DEWR 2009, Cherastidtham and Norton 2018, Beavis and Elsworth 1998).

The loans module

After estimating the aggregate number of students in the base year by sex, field of study and graduation status, we use these counts to generate an individual-level dataset of past students. This dataset is created by sampling the historical records of bachelor's student who left university between 2015 and 2020. We choose recent, pre-pandemic years as our sample so that student study patterns are more representative of current students. We draw random samples from historical records to match the counts estimated in the entry module for each subgroup. This cross-sectional dataset of sampled individuals is the base-file for the model.

We then turn to estimating the amount of HELP debt each student has when they leave university. Because the base-file of simulated students is sampled from past records of real students, we can use the study records of each student to estimate their debt. First, we link higher education records to our base-file, which provides data on every subject each student attempted at university. Next, we assign each subject a cost according to the subject's field of study and equivalent full-time load, based on a set of policy rules. We then sum the cost across all subjects for each student to provide a total HELP debt. This assumes students make no upfront payments or repayments before leaving university, and that past study patterns (within a field of study) are broadly representative of current patterns.[6]

This approach provides a richer method to simulate student debts under policy changes than previous literature. It captures the real study patterns of recent students, including any electives they study, whether they fail and repeat units, or if they choose to switch courses part-way through their degree. This gives a more realistic estimate of the heterogeneous mix and number of subjects different students study during a bachelor's degree. Importantly, it is also flexible and can be used to model different policy scenarios, where the costs of study differ from current or historical costs.

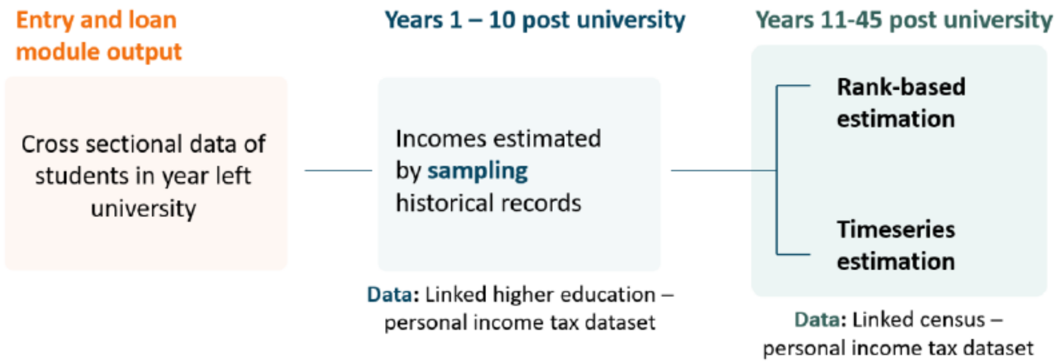
The income module

At this stage, the base-file is a simulated cross-sectional dataset of ex-students in the year they left university. The dataset includes demographic information and the estimated HELP debt for each individual. Next, we simulate their lifetime incomes using a combination of sampling and estimation, to create a longitudinal dataset.

We rely on empirical distributions and stochastic methods to estimate incomes, described in detail below. For the first 10 years after leaving university, we generate students' incomes by sampling historical records. Beyond the sampling period, we construct two different methods for estimating incomes, both of which rely on Monte Carlo simulations. The first method uses empirical distributions of income growth patterns to estimate growth in income over time. The second uses Copula models to estimate transitions between income ranks over time.

6 A small share of domestic bachelor students make upfront payments toward their degree; further iterations of the model could account for upfront payments.

Figure 1: Income module schematic



Income sampling

To sample incomes, we use the linked dataset of higher-education and personal income tax records which includes the income history of bachelor's students who left university between 2005 and 2021. First, we restrict this data to individuals who left university before 2012, and we truncate the data so there are 10 years of history for each student.

We then draw a random 10-year income profile from the linked higher-education – tax dataset described above for each simulated person in the base-file. This random draw is matched to the simulated individual by their sex, age, field of study, graduation status, and whether they studied a double degree. This gives each simulated individual the 10-year income history of a past bachelor student who shares similar characteristics, creating a 10-period panel of simulated students and incomes.

Income estimation

Beyond the 10-year sampling period, we estimate student incomes through Monte Carlo simulations. Two methods are available, described in turn below. In general, we prefer the timeseries simulation approach, because it captures some structural change in female incomes over time and provides slightly more realistic estimates of income mobility over time. However, both models give very similar estimates of repayment times and costs to government under most policy scenarios.

For these estimates, the linked higher-education – personal income tax dataset used for sampling is less useful because it only includes records of students left university more recently and tend to be younger. Instead, we rely on the linked census – personal income tax dataset to inform our assumptions, which contains individuals with a bachelor's degree across all age groups, as well as their incomes between 2002 and 2022, and self-reported field of study.

Timeseries simulation

The first method used to estimate incomes relies on Monte Carlo simulations to stochastically model how incomes grow over time. The model has two key components. We first estimate and apply a simple labour force model to capture large income movements and provide structure to the model. We then pull random draws from empirical distributions to capture smaller changes to incomes over time. In many ways, this form is similar to early Markov-chain income modelling approaches, such as those pioneered by Champernowne (1953), but with an autoregressive component.

Our simple labour force model has three states: 'no income', 'lower income', and 'regular income'. Individuals have 'no income' if they have $0 income in a year (imputed or actual); 'lower income' if their income is below $50,000 or has dropped by 50% from the previous year; and 'regular income' if they earn above $50,000 and their income is above 50% of their income in the previous year. To calculate the likelihood of transitioning between each state, we estimate a probability table from the linked census – personal income tax dataset. Each of the parameters used to define the earning state can be modified.

We first assign individuals in the dataset an income state in each year, according to the rules. We then estimate the probability of moving between each state, by age, sex, field of study, and previous earning state history. From these probability tables, Monte Carlo simulations are used to determine if an individual will move between states in each given year. An individual's income is imputed as $0 if they are assigned the 'no income' state.

Once individuals are assigned an earning state for the latest year (denoted year *t + 1*), their income is estimated conditional on this state. To inform our estimates of incomes, we generate a series of empirical cumulative distribution functions, using income histories in the linked census – personal income tax dataset.

We estimate two types of ECDF's: distributions of incomes in (real) dollar amounts in year *t + 1*, denoted $\hat{F}_{income}(x)$; and distributions of the percentage change in income between year *t* and *t + 1*, denoted $\hat{F}_{growth}(x)$. For each, we estimate an individual ECDF for different 'cohorts' individuals, which we consider as a unique combination of sex, age, field of study, earning state in year *t* and *t + 1*, and income quintile in the previous year (year *t*).

To estimate all non-zero incomes in the year *t + 1*, a random draw is made one of the two types of distributions, for the appropriate cohort. Whether the 'income' or 'growth' distribution is used depends on the earning states in the current and previous year (*t*). In general, an individual's income is estimated using the $\hat{F}_{growth}(x)$ distribution if they remain in the same earning state between years. We tend to use the income distribution ($\hat{F}_{income}(x)$) where there are changes in earning states between years, particularly where income is increasing off a low, or zero, base (see appendix for a decision matrix on which distribution is used).

If the income distribution ($\hat{F}_{income}(x)$) is used, a dollar is randomly drawn and applied directly as the individual's income in the new period (year *t + 1*). If the growth distribution ($\hat{F}_{growth}(x)$) is used, the percentage change value drawn from the distribution is multiplied by the individuals previous income in year *t*. I.e:

$$Income_{t+1} = \begin{cases} N, \quad where\ N \sim \hat{F}_{income}(x), \quad if\ income\ distribution\ used \\ Income_t * (1 + N), \quad where\ N \sim \hat{F}_{growth}(x), \quad if\ growth\ distribution\ used \end{cases}$$

This estimation is repeated yearly until there is 45 years of simulated income history for everyone (including the 10-year sampling period). Once all incomes are estimated, headline (real) wage growth assumptions are applied to account for structural growth over time.

Rank-based simulation

The rank-based method estimates incomes over the life course by simulating how an individual's position on the income distribution (their income rank) changes over time. Following Dearden et al. (2006), we use copula functions to capture the dependence structure between income ranks over time.

This estimation has three steps. First, we use a simple labour force model to estimate the likelihood that an individual earns 'significant' income in a year. We use the linked census – personal income tax dataset of individuals with a bachelor's degree to estimate a probability table of the likelihood of earning above $20,000, by age, sex, field of study, and the previous history of earning above $20,000. Beyond the sampling period, we run Monte Carlo simulations on these probabilities to determine if an individual will earn above $20,000 in the year $t + 1$. We impute the incomes of individuals not earning above $20,000 as $0, as they would not make repayments toward their HELP debt under most policy settings. Conditional on earning above $20,000, we then estimate income ranks in year $t + 1$.

To estimate income rank transitions, we fit copula models for each combination of sex, age, and field of study. To fit these models, we calculate income ranks from the linked census – personal income tax dataset for individuals earning above $20,000, for each combination of age, sex, field of study and year. Consistent with previous literature (for example, Higgins and Khemka 2024), we fit t-copula models for each adjacent 'pair' of income ranks in each subgroup, for example individuals' income ranks in years n and $n + 1$.

We then use the fitted models to estimate the income rank transitions of the simulated individuals in the model. First, we calculate the income rank for each individual compared to their peers, in year $t = 10$ (the final year of the sampling period). We use the fitted copula models to run Monte Carlo simulations, which estimate an income rank for each simulated person in year $t + 1$, based on their initial income rank in year t and the joint probability distribution of the fitted copula model.

Finally, we must map each person's income rank to an income figure in dollar terms. To do this, we calculate cross-sectional income distributions from the linked census – tax dataset of individuals with a bachelor's degree, for each combination of age, sex and field of study. We then map each individual's income rank to the equivalent position on the cross-sectional distributions to assign a dollar figure in each year. This process is repeated until each person has 45 years of income history. Once all incomes have been estimated, we apply headline (real) wage growth assumptions to all incomes.

The repayments module

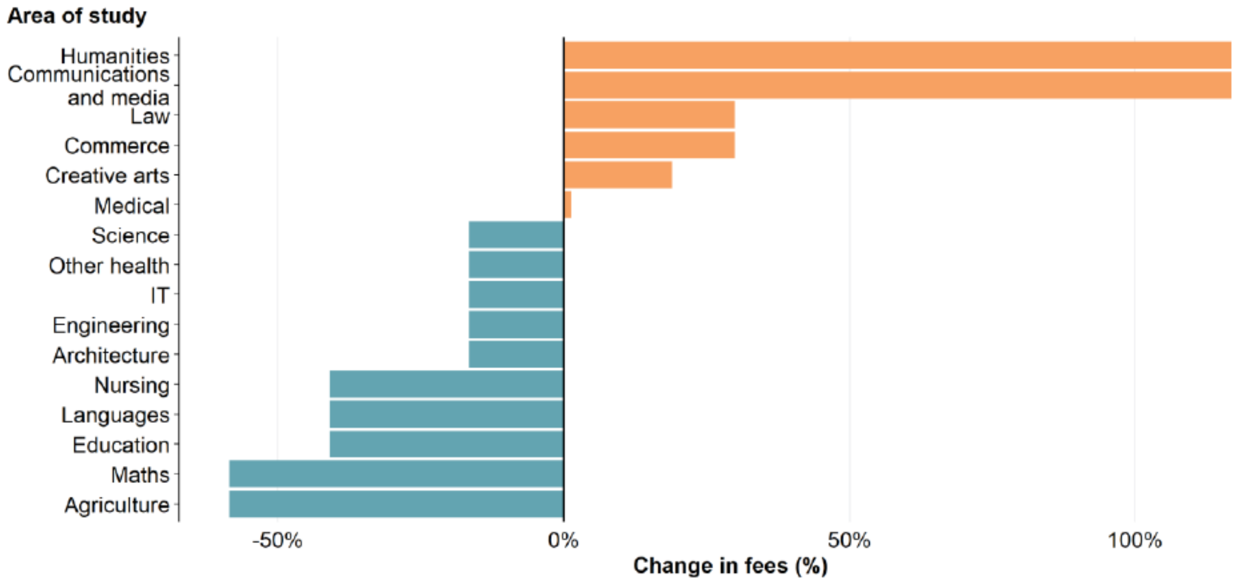
After estimating incomes, we have a longitudinal dataset of simulated individuals, their demographics, initial HELP debts upon leaving university, and 45 years of estimated annual earnings. We then use this file to estimate the expected HELP repayments iteratively in each year. If an individual has a remaining HELP balance, we calculated their expected repayments according to policy rules, assuming students do not make voluntary repayments. This expected repayment is subtracted from the outstanding balance of HELP debt; the calculation then moves to the next year until there is no HELP debt remaining, or the 45-year limit is reached. Because the model operates in real terms, indexation is only applied if the assumed indexation rate of HELP debt differs from CPI.

4. Using HELPEM to examine a recent policy change: The Job-Ready Graduates package

Using the introduction of the 2021 Job-Ready Graduate policy changes as an example case, the remainder of this paper demonstrates how HELPEM can be used to simulate the effects of policy changes. We explore how the policy change affected the amount of debt students hold, expected repayment times, and costs to government of providing HELP loans, under the recent change and previous policy settings.

In 2021, the Job-Ready Graduates (JRG) package changed the amount that government and students would contribute to the cost of study. Chart 1 shows the change in student fees under the Job-Ready Graduates package.[7] Study in some fields became cheaper, while other areas of study increased in price. Previously, the highest fee rate was twice the amount of the lowest fee rate; under JRG, the difference between the lowest and highest rate increases to four times.

Chart 1: Changes to student contributions under the JRG package



Note: There are different numbers of students in each field. Field of study definitions in this chart differ slightly from those used in HELPEM, due to the structure of the fee changes.

Model scenarios

To explore the policy change, we simulate two scenarios using the model:

1. Pre-JRG: A scenario where students pay for all their study under 2020 fee rates (inflated to 2024 dollars)
2. JRG: A scenario where students pay for all their study under current (Job-Ready Graduate) fee rates

In both scenarios, policies are assumed to be at 'full effect', with students undertaking all of their studies under a single policy and set of fees. To test the differences between the two scenarios holding 'all else equal', we simulate the lifetime incomes of a single cohort of bachelor students who leave university at the end of 2024. We include students who do not complete their degrees in the analysis, and it is assume that all eligible students defer the cost of their studies with HELP loans and do not make voluntary debt repayments. Aside from fees, we assume that other policies follow current legislation as of December 2024.

In both scenarios, we assume that CPI and WPI forecasts follow internal forecasts until 2030, at which point they reach rates of 2.5 and 3.7 (nominal) per cent annually and remain stable in the long-term. We assume that the cost of borrowing is 5 per cent, in line with other models (O'Niell and Antcliff 2009, Higgins, 2023).

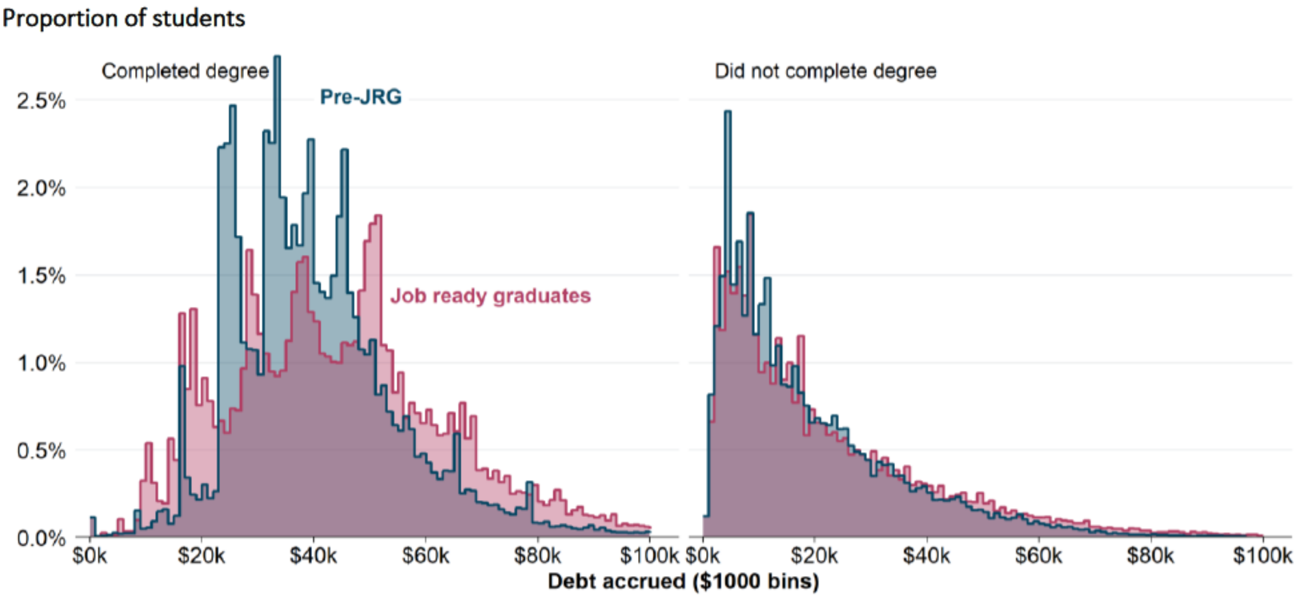
7 For this analysis, we only focus on changes to student contributions.

Effect of JRG on the distribution of student debts

Under the JRG scenario, average graduate debts are 12-13 per cent higher in real terms than in the Pre-JRG scenario. However, there are more significant changes across the distribution.

Many more graduates would attract much lower or higher debts under JRG rates (Chart 2). More than twice as many graduates would leave university with debts under $20,000 under JRG fees, and the share who leave university with debts of $50,000 would increase by 70 per cent. About 30 per cent fewer students would graduate with debts towards the centre of the distribution (between $20,000 - 50,000), compared to under pre-JRG settings. These trends remain if debts are standardised per full-time year of study.

Chart 2: Distribution of bachelor debts under JRG and pre-JRG fee scenarios



Note: The distribution of debts included in this chart differs slightly from the distribution used in the final model run for which results are included, as it is based on earlier research conducted before changes to indexation where legislated.

Under Job Ready Graduates, student fees changed according to the field of each unit of study. Compared to pre-JRG fee rates, the difference in cost between the cheapest and most expensive units increased from about two-fold to four-fold. JRG also changed which units were more and less expensive relative to one another. Changes to the distribution of debts not only reflect a 'widening' of the distribution, but also a 'rearrangement' (Chart 3).

Some fields such as Engineering and Nursing became notably cheaper, while others such as Commerce, Law and Humanities became more expensive. Under pre-JRG fee rates, the majority (65 per cent) of simulated humanities graduates have debts below $35,000; under JRG fees, a similar share (62 per cent) have debts over $50,000.[8]

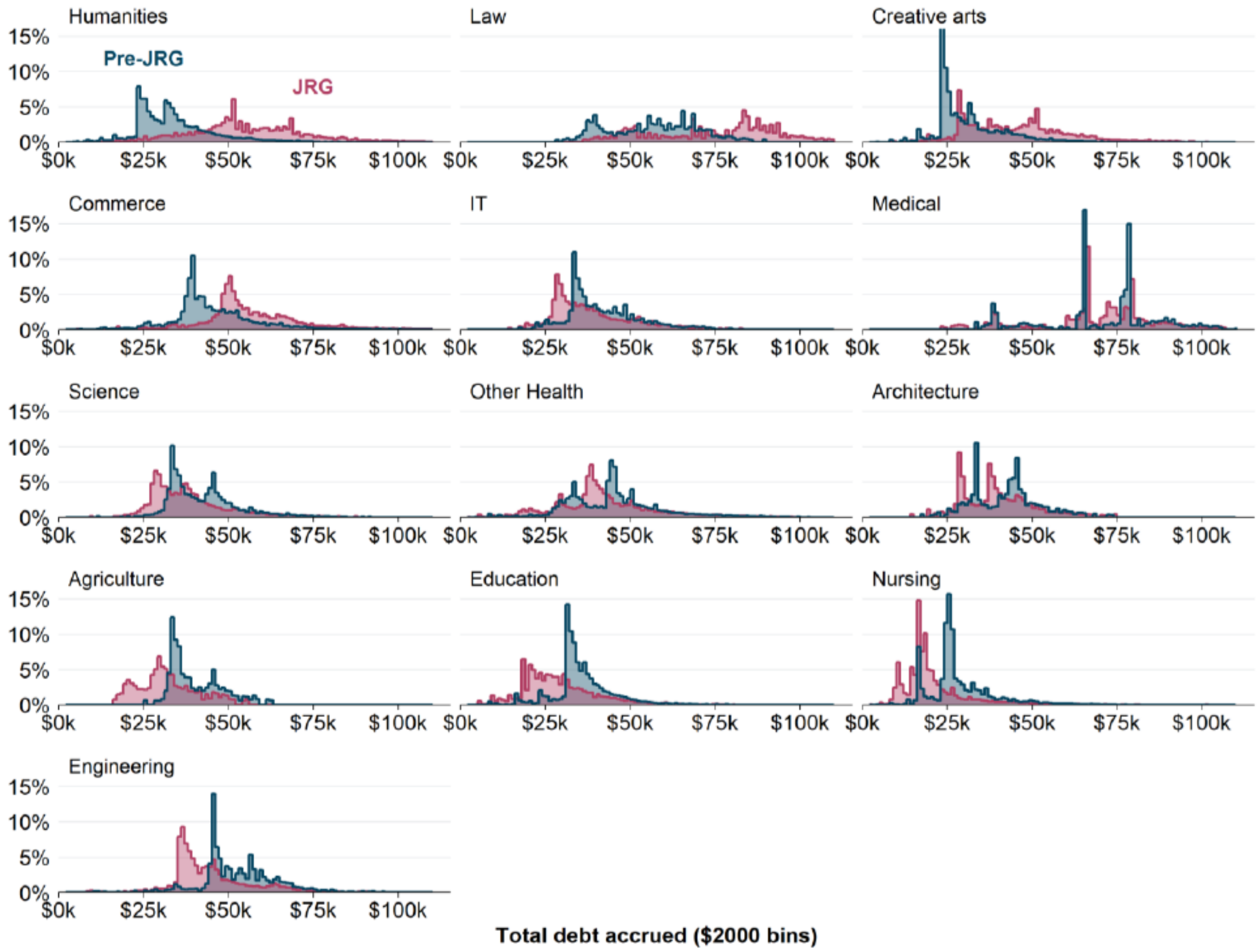
Under the JRG scenario, law graduates have the highest levels of debt, owing to high student contribution fees, the long length of law degrees, and because many students who study law take the course as part of a double degree. Often, humanities or commerce is the second field in these double

[8] These figures only include 'graduates'; those who complete their degree.

degrees, both of which also attract the highest student contribution rate. Under the JRG scenario, a substantial fraction of law graduates accrue over $75,000 worth of HECS/HELP.

Chart 3: Distribution of debts under pre-JRG and JRG fee scenarios, by course field of study

Proportion of students



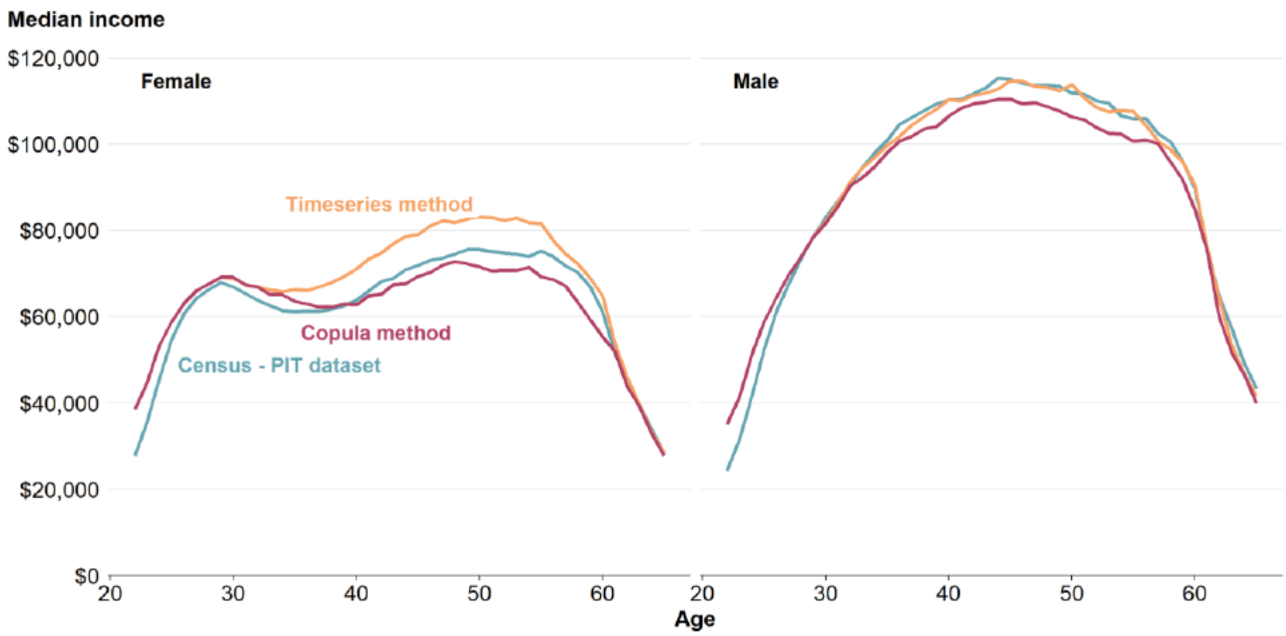
Note: The distribution of debts included in this chart differs slightly from the distribution used in the final model run for which results are included, as it is based on earlier research conducted before changes to indexation where legislated.

Income estimates

Because we assume students' future incomes are independent of their HELP debts, incomes are identical under the pre-JRG and JRG scenarios.

Chart 4 shows the estimated median incomes, by age and sex, under the two income modelling approaches in HELPEM. For comparison, we include an additional series which contains the cross-sectional income profile of incomes for individuals with a bachelor's degree calculated from the linked Census – personal income tax dataset. To provide a fair comparison, incomes in the chart below are all inflated and deflated using WPI. We include all future years of incomes estimated by the model to give results over the entire age range.

Chart 4: Estimated median income by age and sex, WPI adjusted



Until ages of about 30-35, both model results are identical because incomes in this period are sampled from historical records. Beyond the sampling period, the timeseries approach has a higher estimate of median incomes than the rank-based (copula) method, particularly for women.

The higher estimates in the timeseries model stem from the model's autoregressive approach. The income values in the copula method are taken directly from cross-sectional estimates of historical income distribution, which does not account for any future structural change over time. By comparison, the timeseries approach 'carries forward' recent structural changes in the incomes in younger female cohorts. It therefore accounts for some level of future structural change, giving higher estimates of female incomes over the lifecycle.

There are also some small differences between the historical median incomes and those estimated by the rank-based model. These differences are because the students in the model studied a different mix of fields than individuals in historical data, and because our model includes students who did not complete their degree who are not included in the census - personal income tax data.

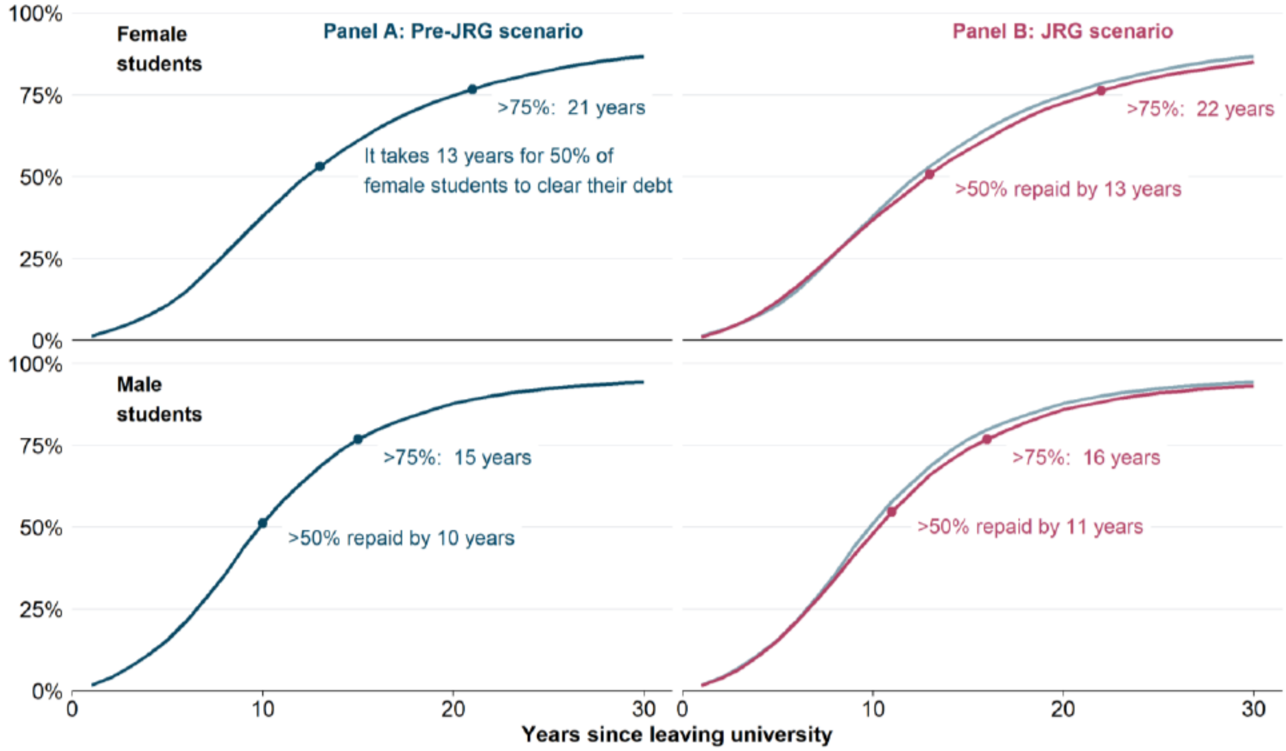
In general, both models provide a similar level of year-to-year income mobility to historical data (not shown). The rank-based approach generally shows slightly higher levels of income mobility, particularly over time periods longer than one year. This finding is consistent with previous work (Dearden et al 2006), and probably stems from the assumption that income rank transitions can be modelled through a single-period Markov process.

For brevity, the remainder of this paper only includes results where incomes were estimated using the timeseries. We generally prefer this method as it captures some structural change in female incomes over time. However, both models give very similar estimates of repayment times and costs to government and our conclusions are not sensitive to the choice of model.

Repayment times increase under JRG fees

Chart 2 shows the modelled length of time it takes for students to fully repay their bachelor debts under the pre-JRG and JRG scenarios.

Chart 5: Proportion of students who have repaid their bachelor debt



Note: The faded blue line on the right panels indicates the pre-JRG scenario. Medians and 75th percentile estimate are not precisely at 50 or 75 per cent, because repayments are made yearly. The curves smooth the discrete yearly values.

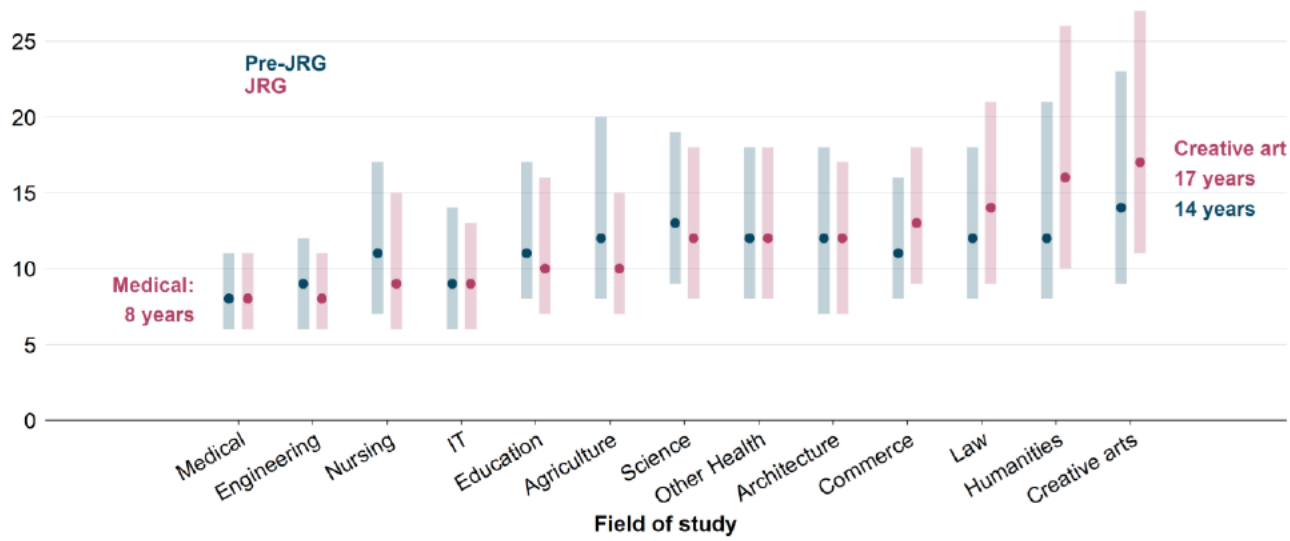
Under the pre-JRG scenario, we estimate it would take about 13 years for half of female students, and 10 years for half of male students to repay their debts (left panels). The introduction of the JRG package slightly increases repayment times by less than one year for women, and about one year for men (right panels). The time taken for 75 per cent of students to repay their debts also increases by about one year for both male and female students.

Repayment times increase under JRG because average student fees are higher, and because fee increases are largest in fields with lower future incomes. However, the aggregate effect of JRG on average repayment times is relatively small, because students in many fields benefit from lower fees than previously (see Chart 3).

Increases to repayment times are not felt equally across disciplines

For any individual student, the impact of JRG on repayment times depends on the field they studied (Chart 6). The field with the shortest median repayment time, medicine, remains the same with half of students expected to repay their debts in 8 years. For students who studied creative arts, the field with the longest repayment times, median repayment times increase from 14 to 17 years. Previously, the median creative arts student paid off their debt 2 years later than the overall student median, but this increases to 5 years under JRG fee rates.

Chart 6: Time taken to repay debt, for 25th percentile, median, and 75th percentile, by field of study



Note: Points are median repayment times. Shaded areas indicate 25th and 75th percentiles.

Under pre-JRG fees, most fields (8 of 13) have a median repayment time of 11 or 12 years. Three fields have median repayments times of less than 11 years, and the remaining 2 fields had median repayment times above 12 years. The shift to JRG fees increases the spread of median repayment times. Under JRG, only 3 fields have median repayment times of 11 to 12 years. Four fields have median repayment times longer than 12 years, and 6 fields have median times of less than 11 years.

The time taken for 75 per cent of students to repay their bachelor debt depend more on model specifications and assumptions. However, we generally find that differences in repayment times between the two scenarios are larger at the 75th percentile compared to the median. For example, repayment times for the 75th percentile of creative arts and humanities students are 4 and 5 years higher respectively under JRG fee rates than pre-JRG fees. In both creative arts and humanities, we expect it to take more than 25 years for three-quarters of students in each field to repay their bachelor debt.

The increase to the range and spread of repayment times mean there are large differences between fields of study in when students are likely to clear their debts. For a humanities student who enters university directly after high school and completes their degree in three years, a median repayment time of 16 years under JRG fee rates suggests they will be 37 years old by the time they have repaid their debts. Comparatively, a similar nursing student will clear their debt at age 30.

How does JRG affect the fair value of HELP loans?

HELP allows students to delay paying university fees until their income reaches a repayment threshold. Although this smooths the cost for students, it does carry a cost to government, which has two main components:

1. The amount of debt not expected to be repaid, written off upon the death of the debtor.

2. An 'interest subsidy': the difference between the indexation rate of the loan and the government borrowing rate.[9] This is equivalent to a discounting of future repayments - repayments further into the future are worth less to government.

At any given time, the total amount of debt in the HELP system is referred to as the 'face value' of debt. The face value of debt overstates how much that debt is worth to government because it does not account for debt that will never be repaid, or for interest subsidies.

Accounting for the debt which is not expected to be repaid and for interest subsidies gives the 'fair value' of debt, which is a better indication of what it is worth to government. The difference between the 'face value' and 'fair value' of debt can be expressed as a percentage, often called the total subsidy rate, which indicates the fraction of debt government loses by providing HELP loans. The expected fair value of aggregate HELP debt is formally modelled by the Government Actuary.

Chart 7: Fair and face value of debt for students leaving university in 2024, by scenario

Chart 7 shows the estimated face and fair value of debt for a single cohort of bachelor students, under each of the policy scenarios. Our estimates are broadly in line with estimates from other Australian models. For example, the Government Actuary estimates the subsidy rate of new HELP debt accrued in 2022-23 is 28.3 per cent (Department of Education, 2023). Modelling repayments of current JRG fee-levels, Higgins and Khemka (2024) estimate a subsidy rate of 28.4 per cent for graduates studying a 3.5-year degree.

Under the pre-JRG scenario, the face value of debt for the bachelor cohort is about $6.4b, with a fair value of $4.7b. This implies a total subsidy rate of about 27 per cent. Under the JRG scenario, the total pool of debt would be 12-13 per cent, or $820m higher, but the fair value of debt increases by just over half this amount ($440m). The lower increase to the fair value occurs for two reasons.

First, a marginal increase to a student HELP debt will necessarily lead to a smaller increase in fair value, because the additional debt is repaid 'last', once a student has repaid their initial balance. The

[9] Outstanding loan balances are currently indexed at the lower of WPI or CPI. The long-term government borrowing rate varies, but we assume a rate of 5 per cent per year, in line with the Government Actuary. In a year where CPI is 2.5 per cent, WPI is 4 per cent, and the borrowing rate is 5 per cent, the interested subsidy will be the difference between CPI and the borrowing rate (2.5 per cent).

marginal dollar therefore attracts a higher interest subsidy than the average, and there is a higher chance that is never repaid.[10]

Second, the JRG package changed how debts were distributed between students. Before the Job-Ready Graduate reforms, fees were more closely aligned to students' future incomes: fields of study with higher earnings incurred higher fees. Because JRG increases fees in fields where students tend to have lower earnings, a greater share of debt is held by students who are slower to repay and less likely to fully repay their debts (see Higgins and Khemka 2024, for a more detailed discussion of subsidy rates across fields of study).

The combination of these two factors increases the subsidy rate by about 2.3 percentage points; about 1.1 ppt increase due to higher overall fees, and about 1.2 ppt increase due to the reallocation of debts toward students with worse repayment outcomes.

Conclusion

Australia was the first country to introduce an income contingent loan system to support higher education funding, and the program is often lauded as a policy success. Given the size, widespread use, and evolving nature of the HELP system, it is important for policymakers to understand the effects of new or proposed policy changes.

This paper provides an overview of a new model to estimate HELP debt repayments. HELPEM aims to give policymakers a new tool to explore how changes to higher education policies could affect outcomes for students and governments.

Our model builds on existing literature and applies both previously used and new methods to richer administrative data. In particular, we build on past methods to estimate student debts by making use of detailed information about student study patterns. This allows us to estimate how the distribution of student debts would change under different policies settings in the absence of behavioural responses.

We use the model to explore a recent policy change, the 2021 introduction of the Job-Ready Graduates package. Although our findings are not projections or forecasts, because they only model policy scenarios and do not take account for upfront or voluntary repayments, the results give broad insights about the expected longer-term effects of the Job-Ready Graduates program.

Under the Job-Ready Graduate package, we estimate that the average student will take longer to repay their bachelor debt compared to previous policies. We estimate that JRG will contribute to larger differences in repayment times between students who studied different fields.[11] For some fields, the result of the policy change is long expected repayment times. In the absence of upfront or voluntary payments, we expect it to take 16 and 17 years for half of humanities and creative arts students to pay off their debts. We estimate that a quarter of students will take more than 25 years to completely repay. Many students in these disciplines may continue to make repayments toward their debt in their late 30's, 40's, and beyond.

[10] Assuming that the allocation of debts between students remains unchanged; i.e. that all student debts increase in a similar way regardless of field of study or other characteristics.
[11] Compared to under previous fee rates.

Although we find that JRG policies increase the total amount of debt bachelor students accrue, this is caused by large fee increases to a small number of fields of study. In many cases, graduates from these fields also have relatively lower average earnings.

Because students in lower earning fields have worse repayment prospects, we find that government subsidies increase under the JRG scenario, and that more HELP debt is written off by government. Under the JRG package, bachelor debts increase by 12-13 per cent for new cohorts. However, we estimate that increases to the fair value of debt are about half as large as they appear to be when taken at face value, because much of the increase in debt is held by students with worse repayment prospects. Overall, the effective subsidy rate on a HELP loan increases by about 2 percentage points in the JRG scenario, compared to the pre-JRG scenario.

This is the first version of HELPEM, which could be further developed to create more realistic estimates of HELP debt repayments by future users. Initial suggestions for further work include adding estimates of voluntary and upfront repayments, improving the entry module to give more realistic estimates of student numbers, and more detailed treatment of how macroeconomic and structural changes may affect incomes and repayments. Further work could also be undertaken to compare how methods for estimating incomes compare to other approaches commonly used in the literature, such as methods used by Britton, van der Erve and Higgins (2019), with a particular focus on the level of income mobility over the life cycle.

References

Altonji, Hynso and Vidangos (2023), "Individual earnings and family income: Dynamics and distribution", Review of Economic Dynamics, https://www.sciencedirect.com/science/article/abs/pii/S1094202522000461

Andrews (1997), "The Effect of HECS on Interest in Undertaking Higher Education", Department of Employment, Education, Training and Youth Affairs, https://books.google.com.au/books/about/The_Effect_of_HECS_on_Interest_in_Undert.html?id=gscKtwAACAAJ&redi r_esc=y

Antcliff and O'Neill (2009b), "The Higher Education Loan Programme (HECS/HELP) – Microsimulation Modelling of Individual Repayment Prospects ", Conference proceedings, Biennial Convention 2009, Institute of Actuaries Australia, https://www.actuaries.asn.au/Library/1430%207E%20ONeill.pdf

Aungles (2002), "HECS and opportunities in higher education: A paper investigating the impact of the Higher Education Contributions Scheme (HECS) on the higher education system", https://web.archive.org.au/awa/20070219035517mp_/http:/pandora.nla.gov.au/pan/68064/20070219 0000/www.dest.gov.au/sectors/higher_education/publications_resources/profiles/hecs.pdf

Beavis and Elsworth (1998), Individual demand for tertiary education: interests and fields of study., Department of Employment, Education, Training and Youth Affairs.

Britton, van der Erve and Higgins (2019), "Income contingent student loan design: Lessons from around the world", Economics of Education Review, https://www.sciencedirect.com/science/article/pii/S0272775717301668

Champerdowne (1953), "A model of income distribution", The Economic Journal, http://www.casa.ucl.ac.uk/new-zipf/papers/champernowne.pdf

Cherastidtham and Norton (2018), "What helps and what hinders university completion?", Grattan Institute, https://grattan.edu.au/wp-content/uploads/2018/04/University-attrition-background.pdf

Dearden, Fitzsimons, Goodman and Kaplan (2006), "Estimating Lifetime Earnings Distributions Using Copulas", https://www.researchgate.net/publication/264872149_Estimating_Lifetime_Earnings_Distributions_Using

Dearen (2019), "Evaluating and designing student loan systems: An overview of empirical approaches", Economics of Education Review, https://www.sciencedirect.com/science/article/pii/S0272775717304995

Department of Education (2023), "Reporting of HELP Receivable at 30 June 2023", https://www.education.gov.au/higher-education-publications/resources/reporting-help-receivable-30-june-2023

DEWR (2009), "Year 12 Student Choices: A survey on factors influencing year 12 decision-making on post-school destination, choice of university and preferred subject", https://web.archive.org.au/awa/20110629214716mp_/http:/pandora.nla.gov.au/pan/127723/20110629 1522/www.deewr.gov.au/HigherEducation/Publications/Documents/Year12StudentChoicesReport.pdf

Erjnaes and Browning (2014), "The persistent-transitory representation for earnings processes", Quantitative Economics, https://onlinelibrary.wiley.com/doi/epdf/10.3982/QE239

Harding (2007), "Challenges and Opportunities of Dynamic Microsimulation Modelling", International Microsimulation Association, https://microsimulation.pub/articles/00280

Higgins (2011), "Implications of earnings models complexity to costing Income Contingent Loans", Implications of Earnings Model Complexity to Costing Income Contingent Loans by Tim Higgins :: SSRN

Higgins and Khemka (2024), "HECS-HELP Prices and Distributional Equity: Price setting Based on Expected Future Income", https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4790069

Higgins and Sinning (2013), "Modelling income dynamics for public policy design: An application to income contingent student loans", Economics of Education Review, https://www.sciencedirect.com/science/article/abs/pii/S0272775713001118

Hu, Moffitt and Sasaki (2019), "Semiparametric estimation of the canonical permanent-transitory model of earnings dynamics", Quantitative economics, https://onlinelibrary.wiley.com/doi/full/10.3982/QE1117

Norton (2022), "From Private to public benefit: The shifting rationales for setting student contributions", Melbourne Centre for the Study of Higher Education, Occasional Paper, https://melbourne-cshe.unimelb.edu.au/__data/assets/pdf_file/0014/4320401/From-public-to-private-benefits.pdf

O'Niell and Antcliff (2009), "The Higher Education Loan Programme (HECS/HELP) – Microsimulation Modelling of Individual Repayment Prospects", Institute of Actuaries of Australia, https://www.actuaries.asn.au/library/con09_paper_oneill%20and%20antcliff.pdf

Universities Accord review panel (2024), "Australian Universities Accord", Department of Education, https://www.education.gov.au/australian-universities-accord/resources/final-report

Yong, Coelli and Kabatek, (2023), "University fees, subsidies and field of study", Melbourne Institute, https://melbourneinstitute.unimelb.edu.au/__data/assets/pdf_file/0011/4751741/wp2023n11.pdf

Appendix

Table A1: Field of education definitions and student contribution rates

Field	Field of education codes
Science	01 (Natural and physical sciences)
IT	02 (Information technology)
Engineering	03 (Engineering and related technologies)
Architecture	04 (Architecture and building)
Agriculture	05 (Agriculture, Environment and related studies)
Medical	0601 (Medical studies), 0607 (Dental studies), 0611 (Veterinary studies)
Nursing	0603 (Nursing)
Other health	06 (Health), excluding those included in medical and nursing
Education	07 (Education)
Commerce	08 (Management and commerce), and 0919 (Economics and econometrics)

Humanities	09 (Society and culture), excluding 0919 (Economics and Econometrics) and 0909 (Law)
Law	0909 (Law)
Creative arts	10 (Creative arts).

Source: Cherastidtham, Norton & Mackey (2018). *Student contribution rates are as of 2024, and are for a single full-time study load equivalent year (*Department of Education*). Note that we refer to 'bands', based on student contribution rates, not 'funding clusters', which are based on the total revenue received for the unit (student and Commonwealth contributions).*

Timeseries income simulation decision matrix

Income state in current year *t*	Income state in year *t - 1*	Ever previously been in 'normal income' state?	Method for estimating incomes
No income			Imputed \$0
Lower income	No income		Draw from income distribution: $\hat{F}_{income}(x)$
Lower income	Lower income		Draw from income growth distribution: $\hat{F}_{growth}(x)$
Lower income	Regular income		Draw from income growth distribution: $\hat{F}_{growth}(x)$
Regular income	No income	Yes	Last regular income value carried forward, with adjustment
Regular income	No income	No	Draw from income distribution: $\hat{F}_{income}(x)$
Regular income	Lower income	Yes	Last regular income value carried forward, with adjustment
Regular income	Lower income	No	Draw from income distribution: $\hat{F}_{income}(x)$
Regular income	Regular income		Draw from income growth distribution: $\hat{F}_{growth}(x)$