Evaluating and designing student loan systems: An overview of empirical approaches

Lorraine Dearden

Department of Social Science, University College London, Institute for Fiscal Studies and IZA, 20 Bedford Way, London WC1H 0AL, UK

ABSTRACT

To understand and design student loan systems, realistic earnings and/or income projections for current and future graduates are crucial. In this paper, Current Population Survey (CPS) data from the US is used to demonstrate empirical approaches that can be exploited to simulate lifetime income and earnings profiles for graduates which are needed to understand and design effective and sustainable student loan systems. The crucial element in getting this analysis correct is having reliable simulations of the whole distribution of future graduate earnings and income. Typically, in this literature, the repayment burdens (RBs) of student loans are calculated at different quantiles of the graduate income or earnings distribution. Often, unconditional quantile regression (UQR) is used to calculate age–earnings profiles for different quantiles of the income or earnings distribution. The paper shows that this approach has limitations when evaluating student loans and that simple raw quantile estimation by age with some age smoothing is preferable. This approach can also be used when income is censored and recorded in income bands as occurs with relevant data in some countries. The paper shows a simple way of incorporating dynamics utilizing these age–earnings profiles by quantile even when only very short panel data is available. This involves using copula functions. Having reliable dynamic estimates turns out to be important in assessing not only the taxpayer costs of designing an income-contingent loan (ICL) but also for correctly assessing the extent of loan repayment hardship for individuals.

1. Introduction

One of the most important, and generally undiscussed, issues in applied labor and education economics relates to the use of cross-sectional data to infer the likely future earnings or income[1] of individuals. The research in this area usually assumes that point-in-time estimates can be used as accurate projections of lifetime outcomes. Nowhere is this inference more critical than in research motivated by the need to understand the empirical implications of student loan policy design.

Current research into student loans focusses on two related and separate questions. The first concerns the impact that so-called time-based repayment loans (TBRLs) might have on the hardship of debtors because these types of student loans require repayments of debt irrespective of a debtor's capacity to repay debt. This issue has motivated a substantial literature based on calculations of so-called "repayment burdens" (RBs), the proportion of a debtor's disposable income that must be used to service a TBRL debt. This burden is seen, quite validly, to be a critical aspect of TBRLs because when an individual has a high RB it would seem to follow that repayment is a challenging experience,

leading to anxiety, hardship, the need to require family and/or friends to assist in debt repayment and, in an extreme situation, having to default on the loan. As discussed in the Introduction of this Special Issue, these costs for a debtor – particularly those associated with default – are significant.

The second research and policy-related issue in the student loan research area concerns the costs for government associated with the adoption of an alternative student loans system, an issue also addressed in the Introduction to this Special Issue. What matters here are the subsidies involved when TBRLs are replaced with their only policy alternative, what are known as "income-contingent loans" (ICLs), in which repayment of a loan depends on the future income of debtors. While the conceptual advantages of an ICL are clear and considered in the Introduction, what matters for policy are the design parameters of an ICL related to finding the associated taxpayer subsidies.

For both issues, it is critical to understand the strengths and weaknesses of using cross-sectional data concerning lifetime earnings projections. To address these issues, the paper uses the Current Population Survey (CPS) data from the US to illustrate the importance

---

E-mail address: l.dearden@ucl.ac.uk.

[1] In this paper, graduate income *and* earnings by age are simulated. Income is generally used for repayment burden analysis whereas income-contingent loan systems apply typically to labor earnings.

of estimation approaches that go well beyond the cross-sectional methods typically used.

There are several crucial elements to getting this analysis right, and these and the related conclusions are as follows:

(i) While current research recognizes the need to have reliable simulations of the *whole* distribution of current and future graduate earnings or income in a cross-section, it is demonstrated that the methods currently used with this approach, involving unconditional quantile regression (UQR) analysis, are generally not correct and in some cases can lead to inaccurate - even sometimes misleading - conclusions; and

(ii) Inferences based on cross-sectional analysis that necessarily restricts individuals to remain in the same part of the income or earnings distribution for life (allowing for no income or earnings dynamics) are not credible. The paper illustrates a relatively simple way of incorporating income and earnings dynamics using copula functions and basic panel data that is available in almost all countries. It demonstrates that in most cases, the estimated subsidies for ICL are too high and the estimated RB problems with TBRL are too low when earnings and income dynamics are ignored.

Section 2 discusses the best way to estimate the cross-sectional distribution of age–earnings profiles. Unconditional quantile regression (UQR) methods are compared with smoothed raw percentiles by age, and the paper shows that UQR is not always appropriate and probably should not be used in this context. The section also demonstrates how researchers can effectively deal with earnings or income survey data that has been banded (or partially banded) by using either interval regression techniques or midpoints, coupled with age smoothing, to obtain age–earnings profiles across the distribution that match the actual age–earnings profiles well. Finally, it shows how static lifetime income and earnings projections can be estimated exploiting these cross-sectional distributions of age–earnings profiles. Section 3 shows a simple but sophisticated approach to estimating dynamic lifetime earnings and income profiles when good longitudinal data is not available, which involves a simple extension of the approach developed in Section 2 and copula functions. Section 4 shows the implications for RB analysis of using dynamic rather than static income simulation. Section 5 shows the implications of incorporating earnings dynamics for understanding the consequences of ICL design for graduates and taxpayers. Section 6 concludes.

2. Cross-sectional distribution of age–earnings profiles

2.1. Introduction

In this section, the strengths and weaknesses of the methods used for estimating cross-sectional age–earnings profiles by quantile of the income and earnings distribution are discussed. Getting these profiles right is essential to understanding RB problems with student loans as well as designing student loan systems and estimating the subsidies involved. Methods used in previous work on student loans, such as Chapman and Lounkaew (2015), are shown to be problematic, and this has important implications for policy work; in particular, it is highly likely to underestimate RB problems with TBRL loans.

2.2. Data description

All the analysis in this paper uses data from the March income supplement of the US Current Population Survey (CPS) from 2014, 2015, 2016 and 2017. CPS data was also used in Chapman and Lounkaew (2015). From the CPS, a sample of individuals who have completed a four-year bachelor's degree or higher degree who are aged 23 to 65 is selected. Data on individual income (from all sources) and labor earnings is used in the analysis in the paper.

The total CPS sample across these four years consists of 142,385 observations of which 64,376 (45%) are for males and 78,009 (55%) are for females. A panel is constructed for the subset of bachelor's degree graduates who are observed in two consecutive years - i.e. March 2014 and 2015, March 2015 and 2016 or March 2016 and 2017 - removing all clearly anomalous cases.[2] This results in a panel of 30,917 individuals (61,834 observations) of whom 13,979 are males (45%) and 16,938 are females (55%). Summary statistics for the income and earnings variables for the whole CPS sample and the CPS panel are given in Table 1.

The sample sizes mean that for the panel, there are an average of about 330 individuals per age transition for men and 400 per age transition for women. For the CPS cross-sectional data, there are around 1,500 observations per age for men and 1,800 for women. There is of course variation by age in both datasets, with the lowest numbers concentrated among young male bachelor's degree graduates aged 23 to 25 and female graduates aged 60 and above.

2.3. Estimating the cross-sectional distribution of age–earnings profiles

As highlighted in the Introduction to this Special Issue, estimating RBs across the distribution of income or earnings is essential to understanding how student loan repayments impact on graduates. To do this, one can simply estimate the percentiles of the income and earnings distribution at each age (the marginal distribution at each age) and plot age–earnings profiles for different percentiles of the earnings or income distribution. This allows researchers to look at the associated RBs at different parts of the distribution by age.

Typically, in the repayment burden / student loan literature, these estimates are smoothed by age; in the student loan literature, unconditional quantile regressions have been used to do this at different quantiles of the income or earnings distribution (see Chapman and Lounkaew, 2015). UQRs are important for many questions dealing with causal impact across the distribution of outcomes, but they are not applicable for RB analysis or student loan design. Estimation of RBs at each age across the distribution of earnings needs information on how the qth quantile of actual or "raw" earnings or income *conditional on age,* $Q_q(y|A)$, changes by age. This is because the repayment burden is measured as the loan repayment at age t as a proportion of *actual* income at age t. A UQR instead identifies the impact of the population aging by one year on the qth quantile of unconditional earnings, $Q_q(y)$, *across all ages.*

When regression (mean) techniques are used then $E[y|A]$ averages up to the unconditional mean $E[y]$ over the range of A because of the law of iterated expectations, i.e. $E[E[y|A]] = E[y]$. The estimated coefficient on age (and any polynomials) gives the impact of a change in age on both $E[y]$ and $E[y|A]$. However, for quantile regression, this does not hold. Firpo, Fortin, and Lemieux (2009) show that to obtain the unconditional effect of the variable of interest (A) on the outcome of interest (y), one needs to perform conditional quantile regression (CQR) and then *integrate out* over all the conditioning variables to get the unconditional effect. They show how this can be done using the re-centered influence function (RIF). This is not needed for RB analysis: instead, percentiles conditional on age are required.

To illustrate potential problems, the cross-sectional CPS data for 2014, 2015, 2016 and 2017 from the age of 23 until 65[3] is used to

---

[2] These include a change in sex, age going up by more than two years, change in ethnicity or change in where individual and/or parents were born.

[3] Chapman and Lounkaew (2015) started their earnings profiles at the age of 22. The number of 22-year-old graduates in the CPS data is relatively small, so in this paper the base category is 23-year-olds, and 22-year-old graduates are recoded as being 23. Having a sufficiently large sample size is important when calculating percentiles of the income distribution. As in Chapman and Lounkaew (2015), military personnel are excluded but the self-employed, and BA graduates who go on to do further postgraduate study, are retained. This

Table 1
Summary statistics for bachelor's degree graduates: CPS 2014–17.

	Income ($US, 2017 prices)				Earnings ($US, 2017 prices)			
	Whole sample		Panel		Whole sample		Panel	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Females	54,496	62,315	54,527	50,445	48,651	60,297	48,849	48,330
Males	96,585	112,710	97,553	95,012	87,831	109,654	88,858	91,996
Observations	142,385	61,834	142,385	61,834				

calculate the raw percentiles of income and earnings by age, sex and year.[4] Smoothing techniques are then compared. Because zero earnings and income are included in these quantile estimates, a regression model with a high-order polynomial in age is generally necessary to capture the drop-off in earnings during child-rearing ages for women but also for earnings and income at the bottom of the distribution where fluctuations are more likely. A quintic in age fits the CPS data best for both men and women, although a locally weighted scatterplot smoothing (lowess) approach works just as well (see Cleveland, 1979).[5] The preferred approach is compared with the UQR method advocated by Chapman and Lounkaew (2015). The UQR method turns out to be very sensitive to the functional form used (whether log income and earnings are used or levels and what polynomial in age is used) as well as the age range over which the model is estimated. For low and high quantiles, it proves to be very unstable.[6]

In Fig. 1, estimates from UQR and the preferred approach are compared for the 10th percentile of the male income distribution using 2014–17 US CPS data in 2017 prices. "Raw percentile data" is the 10th percentile of income at each age. "Exponential UQR quadratic" is the model used by Chapman and Lounkaew (2015). "Exponential UQR quintic" is the same model but includes a much more flexible polynomial in age (quintic). "Quintic raw percentile data" is the predictions from a linear regression of the raw percentile level data on a quintic in age.

Fig. 1 shows that both the UQR approaches approximate the raw 10th percentile data very poorly over the full range of ages. The best fit is given by running a regression with the raw conditional quantile data as the dependent variable and a quintic polynomial in age as the independent variable and obtaining the prediction from this regression. A quadratic performs equally as well in this case, as does a lowess smoothing procedure (not shown).

In Fig. 2, the same exercise is repeated for female bachelor's degree graduates in the 25th percentile of the income distribution.[7] Again the UQR with either a quadratic or quintic does not replicate the raw percentile data well, and the preferred model does much better. The
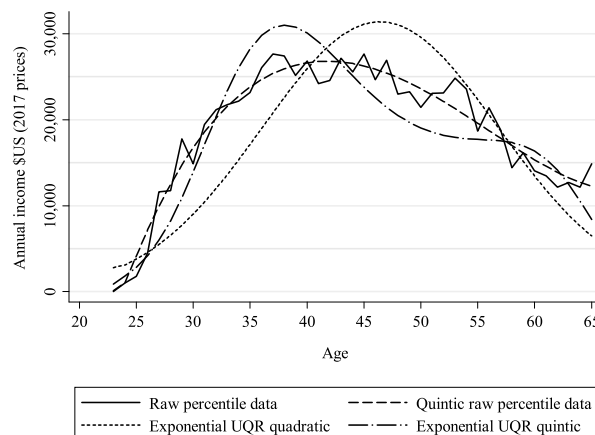

Fig. 1. Male bachelor's degree graduate 10th percentile of income distribution: comparing methods.
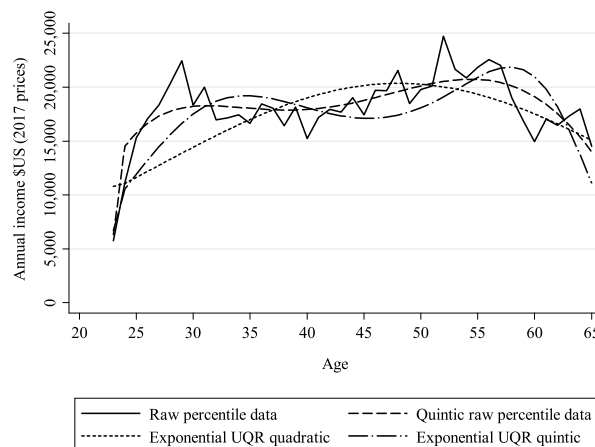
Fig. 2. Female BA graduate 25th percentile of income distribution: comparing methods.

importance of having a quintic specification is also evident here.

In Fig. 3, the estimates of median income (50th percentile) for male bachelor's degree graduates in the sample are shown. While UQR with either quadratic or quintic specification performs quite well for most ages, it overestimates income at young ages which is crucial for RB work. For example, at age 23 the overestimate with the quadratic UQR approach is just over $10,000 or 45%.

In Fig. 4, the implications for high-earning graduates are considered by looking at estimates for women in the 95th percentile of graduate *earnings*. Both specifications of the UQR model overestimate earnings up until about the age of 35, and this overestimation is substantial at low ages (by around 60% or just over $32,000 at the age of 23 with the quadratic UQR specification).

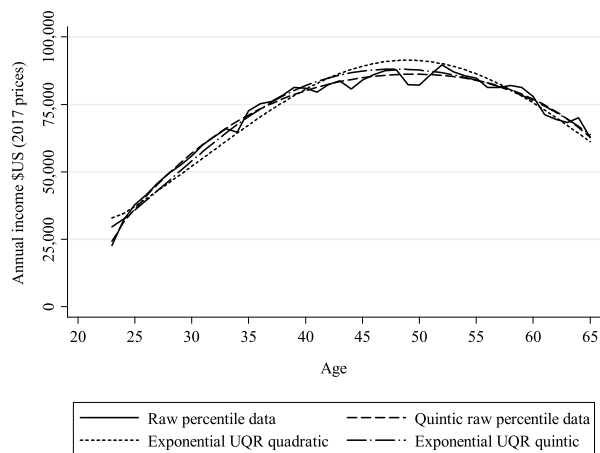This process has been repeated for every percentile[8] of the US

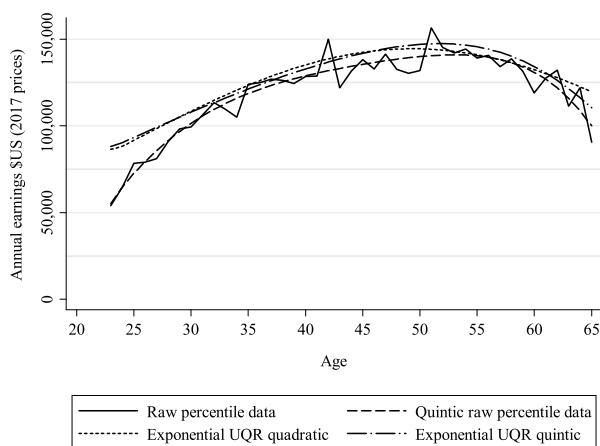**Fig. 3.** Male bachelor's degree graduate 50th percentile of income distribution: comparing methods.

Fig. 4. Female bachelor's degree graduate 95th percentile of earnings distribution: comparing methods.

income and earnings distributions for bachelor's degree graduate men and women. These profiles will be used for simulating graduate income and earnings in Sections 2.5 and 3.4. The approach that best approximates the raw percentile data, essential for RB analysis, involves smoothing the raw percentile estimates by age using a flexible polynomial in age or using "lowess" smoothing of the raw percentile data with an appropriate bandwidth. UQR methods should not be used for this exercise.

Does this mean that all the studies using UQR methods to calculate RBs are wrong? The analysis undertaken using CPS data for this paper shows that a UQR approach with a sufficiently flexible functional form in age generally gives reasonable estimates of the quantiles of the income or earnings distribution at each age (which are needed for RB analysis and student loan design) - except at low and high percentiles of the income and earnings distributions and at young ages.[9] The

(footnote continued)

each age, which are needed for simulation, estimates of the 0.5th, 1.5th, 2.5th, ..., 99.5th percentile of the earnings distribution at each age have been calculated and then smoothed using a quintic polynomial in age. This is equivalent to dividing the earnings at each age into percentiles and taking the median earnings or income at each percentile by age. Hence what is illustrated as the median is the 49.5th percentile rather than the median.

[9] Where earnings and incomes are zero, this is not the case, but it is true for low values of positive income or earnings and relatively high values of income or earnings. For instance, the UQR estimates for the 20th percentile of female income were not sensible (way too high) in all years except 2016 and just not

sensitivity of the UQR to model specification means that without careful exploration of the data, estimated earnings profiles used for RB analysis and/or student loan design is likely to be incorrect.

2.4. Dealing with banded data

The US CPS has all its income and earnings data measured without banding, although to preserve confidentiality there is an income swapping procedure applied to prevent the identification of individuals with extremely low or high incomes. However, this is not true for all countries and/or all data used for estimating age–earnings profiles across the distribution. In Japan, for example, all Labor Force Survey (LFS) earnings and income data is banded into around 10 to 12 bands, and in many countries - for example, Australia - census data has banded income (Armstrong, Dearden, Kobayashi & Nagase, 2019).

Banded data, particularly when the number of bands is small, limits the ability to estimate accurate percentiles of the distribution at a particular age. If there are only 10 bands, there will only be 10 unique percentile estimates, which is problematic. The estimates at each age will be heavily influenced by the distribution of respondents within each band by age and RB calculations at most percentiles will be inaccurate.

With appropriate age smoothing of raw quantile data at each age, this problem may be ameliorated, but it is an open question. In the Japanese LFS data and Australian census data, like the CPS data, there are lots of rich covariates which should be able to reliably position individuals within their known (log) income band, and this can easily be done using interval regression (see Stewart, 1983) and then predicting income *conditional on the band that the individual is in.* Interval regression can be compared with simple age smoothing of midpoint estimates of income for the case where data does not have rich background characteristics.

The approach is tested by banding the full CPS's income data into 14 income groups.[10] The income groups and the proportion of male and female bachelor's degree graduates falling into each category are shown in Table A1 in Appendix A. For nonzero incomes, logs of the lower and upper bounds of each band are taken and interval regression[11] is performed and then the predicted log income *conditional on being within the observed band* is calculated. These predictions are then converted into income levels.[12] The covariates include a cubic in age, year dummies, dummies for grouped total family income, a quadratic in hours of work, detailed industry and occupational dummy variables, ethnicity dummy variables, whether the individual was US-born, whether their father and mother were US-born, regional dummy variables and a metropolitan dummy variable. These variables are highly endogenous, but the sole purpose of this exercise is to get good predictions of earnings and income *within the known earnings and income bands,* so endogeneity is highly desirable for this exercise (unlike most applications).

From these predictions and the simple midpoint estimates, raw

(footnote continued)

credible. Conditional quantile regression could be used, but occasionally there are convergence problems when earnings or income is zero at some ages and positive at other ages.

[10] In an earlier version of this paper, earnings and income were placed in 20 groups with the maximum income group being $150,000 per year. In this version, 14 groups are used, but the high-income group is now those earning over $175,000 per year. This improves the approximation at the top of the male and female earnings distributions considerably and does not change estimates at the lower end of the distributions (where there are now half the number of groups). These changes better replicate the situation in the Japanese LFS data and demonstrate the robustness of the approach.

[11] Stewart (1983) calls this type of estimation "grouped dependent variable estimation".

[12] These predictions include an estimated residual; therefore one can simply exponentiate the within-band prediction.

percentiles by age are calculated. A quintic polynomial in age is used to smooth these raw percentile estimates by age, gender and year using the methodology of Section 2.3.[13] Fig. 5 shows the observed uncensored raw data alongside the smoothed quantile earnings profiles based on the interval regression and midpoint age-smoothed predictions for men in the 5th, 15th, 25th, 50th, 75th, 85th and 95th percentiles. Fig. 6 shows the corresponding diagram for women.

Fig. 5 shows that the age-smoothed profiles perform exceptionally well for men at all quintiles up to the 75th percentile and reasonably well at higher percentiles. With both the interval regression and mid-point approach, earnings are slightly too high for the 85th percentile from the age of 40. For the 95th percentile, they are slightly too high below the age of 35 and slightly too low after the age of 38. The midpoint age-smoothed estimates assume income is $250,000 if men earned above $175,000, the top income group. The US Stafford Loan should generally be repaid within 10 years, so for all but high earning graduates, the estimated income profiles are accurate for the terms of these loans and RB analysis.

Further, as Barr, Chapman, Dearden, and Dynarski (2019) show, high-earning graduates pay off loans more quickly with an ICL than with a Stafford Loan, so again these potential inaccuracies in profiles for high earning graduates, will have minimal implications for ICL design work. The problem at high percentiles arises because there is a large proportion of men earning above $110,000 (see Table A1) and there are only three income bands for this group covering just over 25% of male graduates in the CPS sample. This, however, is less of a problem at younger ages.

For women, the interval regression and midpoint procedures work right across the distribution of female graduates. This reflects the fact that women are more equally distributed in the constructed income bands, as can be seen in Table A1, and the 95th centile is below the top band, unlike the case with men.

The success or otherwise of using age-smoothed interval regression predictions or midpoint predictions by quantile will depend crucially on the distribution of individuals within each band and, in the case of interval regression, the richness of the background data used in the regressions. However, in most countries with well-distributed banded income and earnings data, the censoring is unlikely to cause any significant problems for RB analysis and designing student loan systems.

2.5. Static lifetime income projections

The simplest way to generate lifetime earnings projections for graduates is to assume they remain in the same part of the graduate earnings distribution their whole life. For example, if a graduate has median earnings at age 23, they are assumed to stay in the 50th percentile of the earnings distribution their entire life. This is termed a "static" lifetime earnings projection or simulated lifetime earnings with no dynamics in this paper.

The first important question when doing static lifetime income or earnings projections is how many quantiles are used to summarize the marginal distribution at any one age. In this paper, 100 percentiles are used as this seems to capture the distribution of earnings and income in the CPS well. To test the reliability of this decision, the distribution of the actual (continuous) income data is compared with the distribution of the percentile approximation, as shown in Fig. 7. The figure shows that the distribution of observed income is replicated closely with this simple approximation which uses the smoothed age–earnings profiles by percentile from Section 2.3. It performs equally well for earnings (not illustrated).

To simulate static lifetime earnings, 10,000 males and 10,000 females are assigned into a unique percentile from 1 to 100 (100 in each
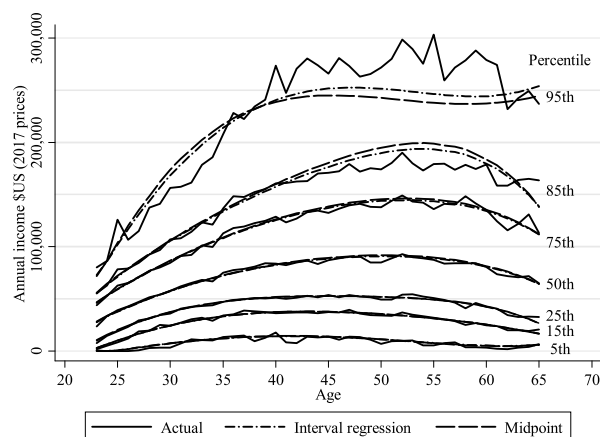

Fig. 5. Male bachelor's degree graduates' earnings by quantile using age-smoothed interval regression and midpoint predictions.
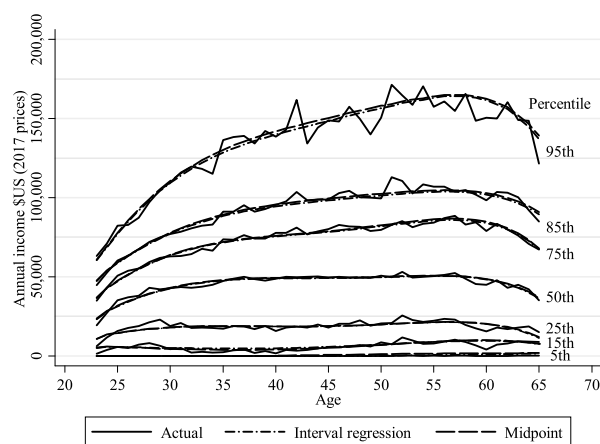
Fig. 6. Female bachelor's degree graduates' earnings by quantile using age-smoothed interval regression and midpoint predictions.

percentile) at age 23.[14] They are assumed to stay in the same percentile from the age of 23 until 65. The smoothed percentile earnings and income by age for men and women are then mapped into the data to provide the baseline earnings and income projections by gender and age. By construction, the earnings paths of the 10,000 men and 10,000 women cannot cross. The observations are appropriately weighted to represent the current size and gender composition of the most recent bachelor's degree graduate population using figures from the Digest of Education Statistics for 2015.[15]

3. Estimating and simulating income dynamics

3.1. Introduction

To accurately understand the implications of student loan systems and estimate the likely taxpayer costs, it is essential to have realistic simulations of future earnings and income over a graduate's working life. The assumption that a person stays in the same part of the income

---

[13] For the top band, females are assigned $220,000 and males $250,000.

[14] In fact, just 100 males and 100 females could have been used as each graduate in a set percentile at age 23 will have identical lifetime earnings and income paths with static simulation. However, this is not true with dynamic simulation discussed in the next section, where much larger sample sizes are needed.

[15] In 2015, 812,669 BA degrees were conferred on men and 1,082,265 on women (see https://nces.ed.gov/programs/digest/d16/tables/dt16_301.10.asp).

Fig. 7. Kernel density estimates of actual vs percentile approximated income (ages 23 to 65).

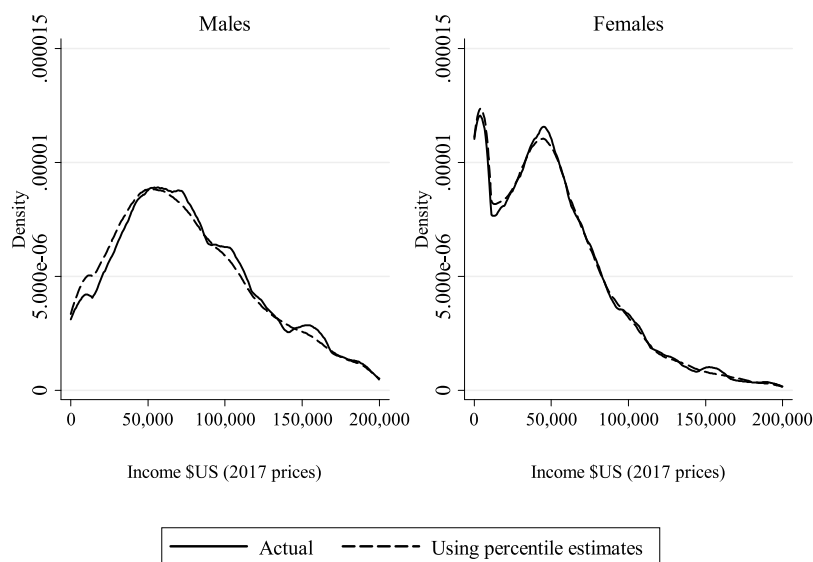or earnings distribution is not credible. When one follows graduates across the lifecycle, their position in the earnings distribution at any age is likely to change as they age, by choice (for instance taking time out to have a child), due to luck (becoming unemployed or moving to a higher-paying job) or for other reasons. In some countries there is lots of mobility (e.g. the US), whereas in other countries there is not much mobility (e.g. Japan).

In this section of the paper, a relatively simple way of simulating dynamic income and earnings paths for graduates is developed, that accurately captures the different income and earnings paths that graduates are likely to face. Importantly, the method proposed for undertaking dynamic simulation can be implemented in virtually all countries with micro labor force data as these generally involve rotating panels of individuals. Further, it also exploits the smoothed age–earnings profiles by quantile discussed in Section 2.3 and used in the static projections discussed in Section 2.5.

In general, estimating dynamics in a reliable way is very important, and this is shown empirically in Section 4 for RB analysis and Section 5 for ICL design. The corollary of this is that if simulated earnings or income dynamics have too much mobility, they are likely to exaggerate RB problems and underestimate the cost of designing an ICL. This also needs to be borne in mind.

3.2. Estimating labor market dynamics with limited data

With long panels, sophisticated methods can be used to get dynamics correct - for example, the approach outlined in Britton, van der Erve, and Higgins (2019, Section 3.1). In many countries, good panel data is not available. The sophisticated methods cannot be implemented with short panels, and simple regression models are not reliable as they assume linear dependence across the income or earnings distribution, which does not accurately reflect observed transitions.

With short panels, it is better to use methods that rely on estimating rank dependence allowing dependence to vary across the income or earnings distribution. Modeling rank dependence can also better overcome issues with measurement error in income or earnings.

A simple way of estimating rank dependence uses copula functions. This involves modeling the joint cumulative distribution function (CDF) of the two marginal CDFs of income or earnings (including zeros) at adjacent ages. This is a simplified version of the approach used by Dearden, Fitzsimons, Goodman, and Kaplan (2008) and Bonhomme and Robin (2009) and provides a simple parametric way of estimating

income or earnings transition matrices (of any dimension). Hence it is related to the dynamic simulation approach used by Higgins and Sinning (2013) with rich Australian longitudinal data.

Crucially, the approach involves the assumption that an individual's rank in the income or earnings distribution next period only depends on their current rank (i.e. is first-order Markov). Bonhomme and Robin (2009) show that for French LFS data with three income observations, this assumption is reasonable and matches the observed transitions over one and two years well, despite the first-order Markov assumption. Armstrong et al. (2019) show this first-order Markov rank dependence assumption replicates correlations over one, two and three years well using Japanese panel data.

The copula function approach is so named as it defines the way two (or indeed many) continuous univariate marginal distributions can be "coupled together" to form their joint bivariate (or multivariate) distribution F. For the approach used in this paper, it is assumed that earnings and income are continuous and observed for every individual at age t (y_{it}) and age $t + 1$ (y_{it+1}). These earnings and incomes are then turned into their CDFs at each age, u_{it} and u_{it+1}. These, by definition, are standard uniform. From Sklar's theorem (Sklar, 1959), if these CDFs are continuous and have joint distribution $F(u_t, u_{t+1})$ and marginal distributions $F(u_t)$ and $F(u_{t+1})$, there is a unique copula function C_t such that:

$$F(u_t, u_{t+1}) = C_t\left(F(u_t), F(u_{t+1})\right) = C_t(u_t, u_{t+1}), \quad t = 23, 24, \ldots, 64.$$

Note that in the setting of this paper, $F(u_t) = u_t$ and $F(u_{t+1}) = u_{t+1}$ since u_t and u_{t+1} are the CDFs of the income or earnings variable at each age t and hence the marginal distributions are also standard uniform. More generally, the marginal distributions of income can be modeled at each age using any distribution or mixture, but the approach used in this paper is to use the empirical marginal distribution by age estimated in Section 2.[16] In this example, C_t is a two-dimensional copula, but the method extends to higher dimensions.[17] Another

---

[16] In Section 2.3, 100 percentiles of the income and earnings distribution at each age have been estimated by gender for BA graduates and can be mapped onto the CDF by multiplying by 100 and rounding up to the nearest percentile. This approach has already been shown to fit the continuous data well (Figure 7) and is explained in more detail in Section 3.4.

[17] Moreover, this joint distribution can be decomposed as a function of the copula function and the marginal densities, i.e. $f(u_{it}, u_{it+1}) = c_t(F(u_{it}), F(u_{it+1})) f_t(u_{it}) f_{t+1}(u_{it+1}) = c_t(u_{it}, u_{it+1})$ where c_t is the copula density and f_t and f_{t+1} are the marginal densities of the copula, which are equal to 1 as the marginal

Table 2
Measures of income dependence in the CPS panel.

Age group in first year	Kendall's tau		Income correlation including zero incomes		Income correlation for nonzero incomes		Log income correlation for nonzero incomes	
	Males	Females	Males	Females	Males	Females	Males	Females
All ages	0.414	0.489	0.435	0.441	0.430	0.420	0.526	0.634
< 25	0.286	0.273	0.148	0.144	0.096	0.125	0.200	0.358
25–29	0.379	0.446	0.256	0.321	0.237	0.262	0.491	0.522
30–34	0.411	0.511	0.391	0.405	0.384	0.371	0.457	0.639
35–39	0.443	0.544	0.355	0.480	0.351	0.446	0.501	0.706
40–44	0.448	0.526	0.411	0.507	0.407	0.486	0.516	0.680
45–49	0.415	0.508	0.475	0.450	0.479	0.428	0.547	0.669
50–54	0.413	0.492	0.414	0.361	0.408	0.351	0.517	0.603
55–59	0.387	0.463	0.435	0.466	0.425	0.454	0.492	0.599
60–65	0.433	0.443	0.447	0.483	0.463	0.472	0.505	0.572

attraction of the copula function is that it makes simulation very easy once parameter estimates by age are obtained. This is discussed further in Section 3.4.

Typically, parametric copula functions are used, and different copula functions allow for different types of dependence (including symmetric and nonsymmetric tail dependence - cf. regression). A goodness-of-fit criterion, such as the Akaike information criterion (AIC), can be used to choose the model that best fits the data.

3.3. Copula model estimates and reliability

The CPS panel of BA graduates from 2014 to 2017 is used to operationalize the copula estimation. This panel was described in Section 2.2. The basic dependence characteristics of the panel are presented in Table 2. It shows that the rank correlation of the CDFs at adjacent ages, measured by Kendall's tau, varies by age group in the sample. The table also shows the correlation of income and log income. Kendall's tau correlation is used to measure rank dependence as this can be easily estimated from the estimated parameters of the copula model for comparative purposes, which is not true of correlation parameters, and is less prone to bias due to earnings or income measurement error (see Appendix B for full details of Kendall's tau).[18] It is important to emphasize that if a person has zero earnings or income, then they are randomly distributed at the bottom of the CDF at each age. For comparison, income and log income correlations for those with nonzero income in both periods are also shown in Table 2.

What is evident from the table is that rank dependence varies by age and there is a lot more mobility at younger ages. This will need to be captured in the estimation and simulations. The lifecycle patterns of correlation exhibited for men and women are also different. It is evident that the (linear) income and log income correlations are quite different too, though they show similar patterns by age to those of the rank correlation. The difference between the income and log income correlations strongly suggests nonlinear dependence. Hence observed dependence is better captured by a rank correlation measure such as Kendall's tau, which does not impose linearity and instead evaluates the monotonic relationship between the ranks of two adjoining income or earnings variables (see Appendix B for details). The copula approach does not require linear dependence, and this flexibility is crucial with short panels.

The estimation strategy involves finding a copula function that best captures the dynamics between the CDFs of income or earnings at adjacent ages from 23 to 65.[19] For almost all ages, the t-copula provides

the best fit for the CPS data, and this is true whether modeling earnings or income dynamics.[20] Dearden et al. (2008) find that the t-copula also works best with earnings data from the UK Labour Force Survey. The t-copula has the dependence structure implicit in a bivariate t-distribution.[21] It has two parameters – the correlation parameter, ρ, and the degrees of freedom parameter, ν. These can be broadly interpreted as describing the overall level of immobility in the distribution (higher ρ) and the excess immobility in the tails of the distribution (lower ν). Kendall's tau (τ) can be directly estimated from the model estimates and, for the t-copula, is given by $\tau = 2\nu^{-1}(\arcsin(\rho))$.

To take account of the observed change in dependence, the t-copula model is estimated separately by gender as well as for every age transition from 23 to 65.[22] The estimates of the two t-copula parameters rho (ρ) and degrees of freedom (ν) and the associated confidence intervals by age, gender and for both income and earnings are shown in Figs. 8, 9, 10 and 11. Smoothed estimates by age are also shown, and it is these smoothed estimates that are used in the simulations.

Fig. 8 shows that for males, there is much more mobility at early ages and then, from the age of 40, mobility settles and remains relatively stable. There is slightly more income mobility than earnings mobility. Fig. 9 shows that there is much less tail dependence at both early and late ages (higher degrees of freedom) and that tail dependence is slightly stronger for earnings than for income.

Figs. 10 and 11 show the corresponding estimates for females. Fig. 10 shows that there is high mobility again at early ages but it stops decreasing from the age of 30. Earnings mobility then remains relatively flat whereas income mobility increases slowly until the age of 65. At all ages, there is more income mobility than earnings mobility, which was also true for men. Fig. 11 shows that at both young and old ages there is less tail dependence than at other ages, where tail dependence is flat. These results from the t-copula model show that the nature of the first-order rank dependence for males and females in the US is very nuanced and help explain why the simple correlation parameters reported in Table 2 could not fully capture the first-order dependence found in the data.

The model's performance is tested for the CPS panel by comparing the predictions of income and earnings at age $t + 1$ from the t-copula model with the actual earnings and income outcomes at age $t + 1$.

(*footnote continued*)
distributions are standard uniform.
[18] The calculations use Kendall's tau-b where ties are counted as concordant rather than discordant; see Appendix B for more details. This makes no difference with continuous marginal CDFs but does if these are made discrete, e.g. turned into 100 percentiles.
[19] The R "Copula" and "VineCopula" packages are used to do this. Transitions

(*footnote continued*)
are modeled at every age, and then the goodness-of-fit tests are used to see which copula best fits the data using "fitCopula" from the "Copula" package.
[20] For example, for male income dynamics, the t-copula is best for 33 age transitions, the Frank copula for six age transitions and BB1, BB7 and survival BB8 for one each.
[21] For detailed information on the t-copula, including a formal definition, see Demarta and McNeil (2005).
[22] This was explicitly built into the Maximum Likelihood Estimation procedure in Dearden, Fitzsimons, Goodman, and Kaplan (2008), but this is not available for the R copula packages used. Instead, separate estimates are obtained for each age transition, and then these are smoothed before simulation.

Fig. 8. Estimates of rho from t-copula: males.

Fig. 9. Estimates of degrees of freedom (df) from t-copula: males.

Quintile transition matrices are also compared. The simulation method proposed involves mapping the age–earnings profiles for income and earnings by percentile, age, year and gender to the estimated percentile from the t-copula model. The way dependence is modeled matters crucially for the distribution of the *difference* in income and earnings between ages t and $t + 1$. Figs. 12 and 13 show that the t-copula model performs reasonably well for both income and earnings changes over one year for men and women.[23] The simulations do not entirely capture very small changes in earnings or income (lower peaks) because of the

percentile approximation of income and earnings at age $t + 1$. However, the overall rank correlation for the one-year-ahead prediction from the t-copula model is always slightly higher than the actual rank correlation for men, although slightly lower for women.[24] The US has more income and earnings mobility than most countries, and Armstrong et al. (2019) show that a copula model assuming first-order Markov rank dependence performs considerably better in countries with less mobility, such as Japan. This means that mobility may be overestimated in this model.

As a final check, quintile transition matrices from the t-copula model and the observed CPS panel data are compared for income in Table 3 for men and Table 4 for women. The model replicates the observed income transitions reasonably well, although it does not get the

---

[23] If the panel member was observed in March 2014 and March 2015, March 2014 income or earnings profiles by percentile are mapped to their actual percentile observed in 2014, and March 2015 income or earnings profiles are mapped to both their actual and predicted percentiles (from the t-copula estimate) in March 2015. This is repeated for panel members observed in March 2015 and March 2016 and for those observed in March 2016 and March 2017. Figure 12 shows the mapping of their actual income and mapped percentile income in the base year and Figure 13 shows the mapping of their actual earnings and mapped percentile earnings in the base year.

[24] For males (females), the actual rank correlation for income is 0.419 (0.494) and the rank correlation with the one-year-ahead prediction is 0.425 (0.490). For earnings, the actual rank correlation for males (females) is 0.461 (0.535) and the rank correlation with the one-year-ahead prediction is 0.468 (0.527).

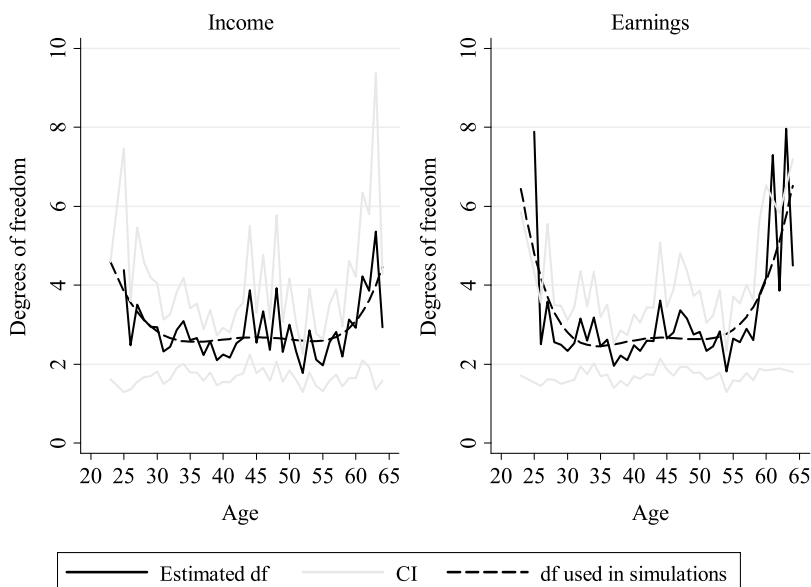Fig. 10. Estimates of rho from t-copula: females.

Fig. 11. Estimates of degrees of freedom (df) from t-copula: females.

slight asymmetry observed in the female income transition matrix, which shows higher tail dependence at lower incomes than higher incomes. Further, it is clear that the percentages in the diagonal of the transition matrix are too low in the simulations, which again suggests that the simulations may have slightly too much mobility. This will need to be borne in mind in the analysis of RBs and ICL design below.

3.4. Dynamic lifetime income and earnings projections

With the t-copula estimates, it is easy to recursively simulate lifetime income and earnings percentiles. This is done for 10,000 males and 10,000 females. This involves the following steps:

1 Place each man and woman in a percentile at the age of 23 (100 men and 100 women in each percentile) using a draw from the random uniform distribution (u_{23}).[25]

2 Estimate the conditional distribution function of u_{24} given u_{23}, which is given by:

$$c_{u_{23}}(u_{24}) = F(u_{24}|u_{23}) = \frac{\partial}{\partial u_{23}} C_{23}(u_{23}, u_{24})$$

where C_{23} is the t-copula with the age 23 rho (ρ_{23}) and degrees of freedom (v_{23}) model estimates.

3 Generate a random standard uniform variable r with the same dimension as u_{23}.

4 Generate $u_{24} = c_{u_{23}}^{-1}(r)$ to get the uniformly distributed predicted rank at age 24, which has a stochastic element due to the rank prediction being determined by not only the rank at age 23 but also the draw from the random uniform. This means that individuals from the same percentile will end up in different percentiles at age

_____

[25] The pseudo-observations command in R is used to ensure that the random

(footnote continued)

draw is precisely uniformly distributed. This is also required for the draw from the random uniform in step 3.
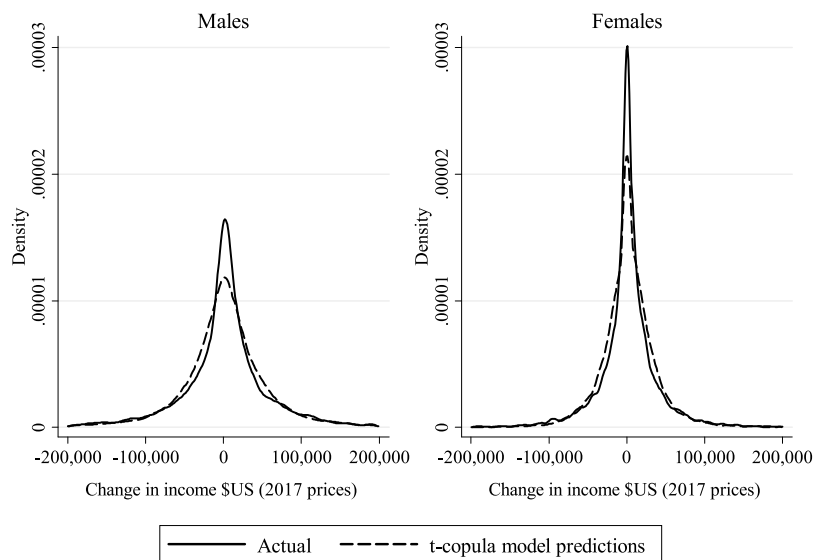
Fig. 12. Predicted difference in incomes at adjacent ages.
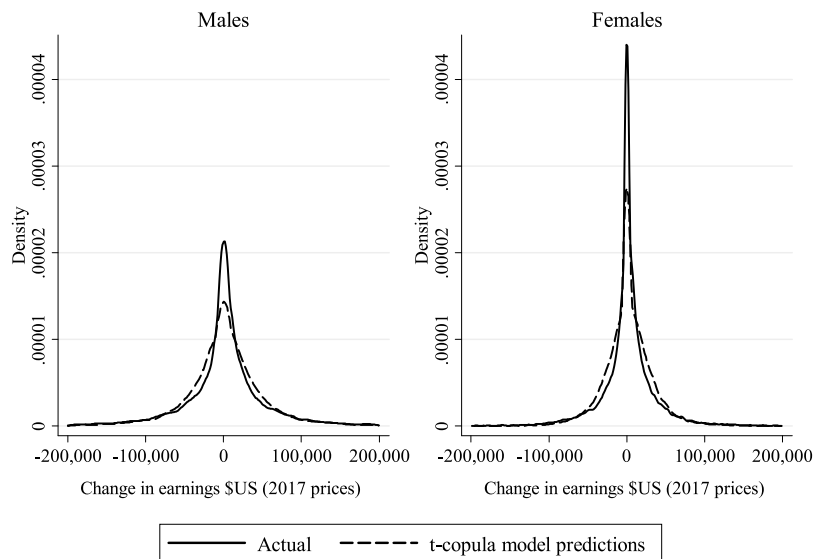
Fig. 13. Predicted difference in earnings at adjacent ages.

Table 3
Male income transition matrices: actual vs predicted.

Quintile at age t	Actual Quintile at age $t + 1$					Predicted from t-copula model Quintile at age $t + 1$				
	1	2	3	4	5	1	2	3	4	5
1	54.79	21.22	10.13	7.49	6.11	53.21	22.22	12.34	6.49	5.46
2	20.61	39.59	21.11	10.20	8.70	22.84	32.60	23.15	14.59	7.05
3	10.20	19.47	37.45	21.75	11.29	10.74	24.32	29.39	25.46	10.24
4	7.72	11.70	18.58	40.41	21.82	7.18	14.41	24.75	30.92	22.97
5	6.68	8.02	12.73	20.15	52.08	6.03	6.46	10.38	22.54	54.28

Table 4
Female income transition matrices: actual vs predicted.

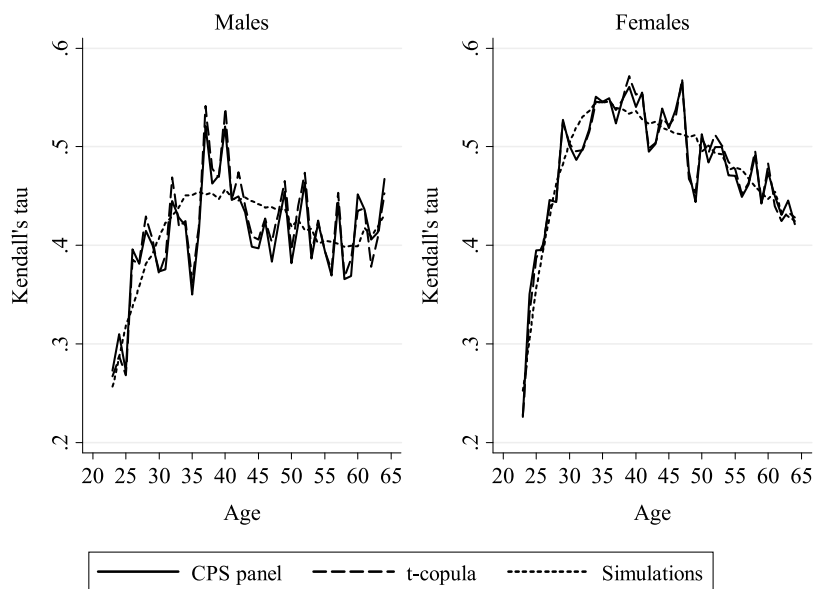| Quintile at age t	Actual Quintile at age $t+1$					Predicted from t-copula model Quintile at age $t+1$				
	1	2	3	4	5	1	2	3	4	5
1	65.76	18.55	6.77	4.77	4.00	58.89	22.35	9.34	5.21	4.03
2	18.14	43.33	19.85	10.81	7.92	22.55	34.39	25.33	12.69	5.19
3	7.01	19.96	40.56	22.06	10.50	8.91	25.41	30.96	25.50	10.32
4	4.76	10.22	22.71	41.90	20.58	5.62	12.31	24.45	36.04	21.77
5	4.32	7.83	10.10	20.47	57.00	4.02	5.54	9.93	21.55	58.69

Fig. 14. Comparison of Kendall's tau for income dependence from CPS sample, t-copula estimates and simulated sample.

24 (the extent to which this happens depends on model parameters).[26]

5 Repeat steps 2 to 4 recursively for every age.

The final simulations of lifetime earnings and income are obtained by multiplying the predicted ranks (uniformly distributed between 0 and 1) by 100 and rounding up to get the simulated percentile by age. The smoothed percentile earnings and income by age for men and women are then mapped into the smoothed income and earnings age–earnings profiles from Section 2, as was the case with the static simulations. This provides the dynamic earnings and income projections by gender and age. It is assumed these graduates started college in 2017 and will graduate in 2021, and the simulated projections are re-weighted by gender to reflect the latest US BA completions. This is important when working out the budgetary implications of different ICL systems.

Fig. 14 compares the estimates of Kendall's tau (τ) from the CPS panel, from the actual model estimates (where $\tau = 2v^{-1}(\arcsin(\rho))$ and from the simulated income sample (where smoothed model estimates of ρ and v were used). It illustrates that the model estimates replicate the raw CPS first-order rank dependence well. As a result, the dependence structures over adjacent ages of the simulated sample, which use smoothed parameter estimates of ρ and v, also mirror the CPS panel rank dependence. For both males and females, rank dependence in income increases until they are aged around 35 and then slowly decreases. However, the shape and extent of this dependence are very different by gender. Females have much more immobility in income than males from the age of 30 onwards, with the gender difference largest at around age 35. This gender difference in rank dependence, however, has almost disappeared by the age of 65.

4. Implications for estimating RBs using dynamic simulations

4.1. Introduction and conceptual issues

In this section of the paper, the implications of using dynamic rather than static simulations of graduate income for RB analysis are assessed. Including dynamics is essential to fully understand RBs over the loan term of a TBRL such as the US Stafford Loan. In most cases, the more mobility there is among student loan debtors, the more likely it is that RB hardship with TBRLs will be underestimated in static simulations. US data is used to demonstrate this, and it is shown that the RB problem in the US is much larger than previous studies such as Chapman and Lounkaew (2015) and Chapman and Dearden (2017) have suggested.

Why might this be generally true? Assume that all graduates take out a loan and the loan amount is the same for all graduates. In this case, if there is no mobility, then the quantile of the income and earnings distribution a graduate finds themselves in will not change

---

[26] For example, suppose that, based on the estimates, a person in percentile 1 had a 50% chance of being in percentile 1, a 35% chance of being in percentile 2 and a 15% chance of being in percentile 3 (and 0% probability of being in any other percentile). Those with a draw from the uniform distribution above 0.85 would be assigned percentile 3, those between 0.50 and 0.85 percentile 2 and the remaining individuals percentile 1.

over their lifetime from the one they were in at the age of 23. Hence a person who remains in the 5th percentile of the income or earnings distribution will face RBs of over 100% for the term of a typical Stafford Loan. Therefore, the *number of individuals* facing at least one high RB will be determined entirely by those facing a high RB at the age of 23 when earnings are lowest. Suppose there is a small amount of income mobility, say a person in the 10th percentile of the earnings distribution moves to the 50th percentile at age 24 and the person whom they replace moves to the 10th percentile at age 24. Both these individuals will now have one period of facing a high RB but also have one period with a relatively low RB. However, the number of individuals facing at least one period of a high RB by the age of 24 will increase by one. Hence introducing *any* dynamics, in this case, cannot decrease, but can possibly increase, the number of individuals facing a high RB for at least one period over the term of a TBRL.

In reality, not all students take out loans, and loan sizes vary, so the above might not hold. However, in most plausible cases, the more mobility there is in income or earnings dynamics during the term of a TBRL, then the higher the probability that a *different individual* at a point of time will experience a bad labor market outcome (and hence a high RB at one point in time if they have a loan). This means that generally, assuming no mobility shows a *lower bound* of the number of individuals potentially facing a high RB during the term of a TBRL. This is shown to be the case for the US in Section 4.3 using plausible assumptions.

4.2. Hypothetical loan characteristics

In the illustrations, only graduates who take out loans are analyzed, and it is assumed the future lifetime earnings paths are the same for debtors as for those who do not take out loans.[27] It is also assumed that the average US student debtor will take out a loan of $35,000 over a four-year BA degree, which is just under the current average of all US student loans for 2017 graduates of $39,400.[28] These loans are assumed to be log-normally distributed and have a standard deviation of $20,000. It is also assumed that the loan amount a BA student takes out is positively correlated with the first 10 years of their total graduate earnings. In these simulations, a correlation of 0.3 is assumed, but the sensitivity of results to this assumption is tested. Papers that have access to US administrative student loan data and tax records confirm that there is a positive correlation between debt levels and later earnings (see Looney and Yannelis, 2019, tables 3 and 4).

Fig. 15 shows the average repayment schedule for a Stafford student loan of $35,000 in 2017 US prices (the average in the simulated sample) assuming 2% inflation. Stafford Loans are TBRLs that in 2017/18 have a nominal interest rate of 4.45% and the majority must be paid back within 10 years. In the first year of the loan, an average student must pay back just over $4,700 regardless of income, as shown in Fig. 15.

4.3. Empirical results

Typically in RB analysis, the repayment burden is shown by age and percentile. However, this implicitly assumes that all graduates stay in the same percentile of the income distribution at each age, i.e. it assumes no income dynamics. This is the approach taken in Chapman and Lounkaew (2015) and has been followed in some of the papers in this issue. This approach demonstrates that with most TBRLs, there is generally a significant problem at low incomes particularly at young ages. Barr et al. (2019) show that even women in the 35th and 50th
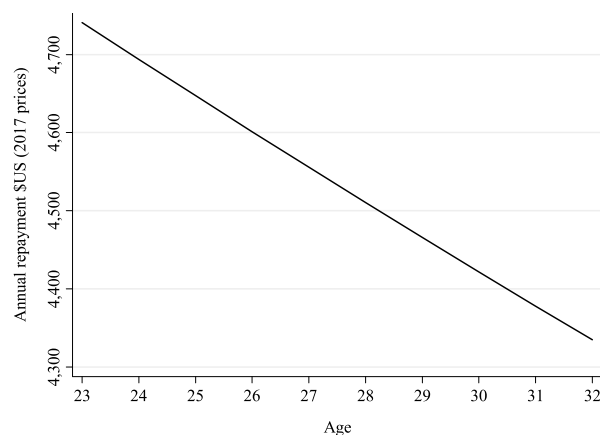

Fig. 15. Loan repayment schedule for Stafford Loan of $35,000.

percentiles of the income distribution in the US face high RBs at young ages.

There is another way of looking at the RB problem which highlights the importance of using dynamics.[29] The simulations from Section 2.5 and 3.4 are used, and it is assumed that there will be 1% real income growth per year over a graduate's lifetime.[30] The loan amount is assigned to individuals using the baseline correlation of 0.3 with the first 10 years of the graduate's total income. Using these simulations, the number of times over the 10-year term of the Stafford Loan an individual will face RBs of more than 18% and more than 40% is calculated (so this could be 0, 1, 2, …, 10 times).

Defining what constitutes an excessive RB is subjective and will crucially depend on other circumstances such as whether the graduate is married and/or has children. Salmi (2003) suggests that RBs at or below 18% of disposable personal income should be the limit to classify RBs as manageable, and this is the first RB examined. A relatively high RB of 40% is also chosen to exemplify the importance of modeling dynamics.

Males and females are pooled together using the weights constructed earlier, which reflect current BA graduation rates. In Table 5, the percentage of the cohort of borrowers falling into each category is shown. A comparison is made between the number of years of excessive RBs when using the simulations with and without dynamics. A comparison is also made with the simulations based on the no dynamics case using the quadratic UQR approach of Chapman and Lounkaew (2015).

As pointed out at the beginning of this section, assuming no mobility is likely to underestimate the RB problem. Table 5 shows that if no mobility is assumed, around 48% of graduates would never face RBs greater than 18% and 70% would never face RBs greater than 40%. If the UQR approach were used, then these estimates would be even higher – 60% and 78% respectively. The estimates from the dynamic model suggest that the correct figures could be as low as 15% and 32% respectively. Further, just under 50% of graduates may face three or more years of having RBs greater than 18%, and just under 25% could face three or more years of having RBs greater than 40%.

This nuanced picture is not captured if income dynamics are not included and helps explain the current default and delinquency problems with student loans in the US highlighted in Barr et al. (2019, Table 1).

The RB hardship figures reported in Table 5 will overestimate hardship if the simulations have too much mobility, which is likely

---

[27] If those who do not take out loans have better lifetime income projections, then the RBs of graduates with debt will be underestimated. If the opposite is true, then RBs will be overestimated.

[28] https://studentloanhero.com/student-loan-debt-statistics/.

[29] For RB work, (gross) income is used as the denominator.

[30] Chapman and Lounkaew (2015) assume 1.5%.

Table 5
Measures of years of excessive RBs for average $35,000 US Stafford Loan.

Number of years of excessive RBs	RB > 18% (percentage)			RB > 40% (percentage)		
	No dynamics, UQR quadratic	No dynamics, raw smoothed quintic	Dynamics, raw smoothed quintic	No dynamics, UQR quadratic	No dynamics, raw smoothed quintic	Dynamics, raw smoothed quintic
0	60.15	47.93	14.71	78.24	69.77	31.90
1	1.70	13.08	17.93	0.85	10.14	25.72
2	2.68	5.40	17.71	1.19	2.65	17.69
3	2.33	4.12	14.68	1.14	1.96	10.30
4	2.30	3.22	11.35	1.06	1.21	6.76
5	2.03	2.39	8.55	0.79	0.90	3.62
6	1.79	1.71	6.12	0.69	0.60	1.94
7	1.62	1.36	3.87	0.65	0.49	1.25
8	1.55	1.15	2.54	0.59	0.41	0.50
9	1.33	0.77	1.48	0.51	0.32	0.22
10	22.52	18.87	1.08	14.30	11.55	0.10

given the limitations of the first-order rank dependence assumption in a country such as the US. However, assuming no dynamics gives a misleading picture of RB problems. Of course, all approaches fail to account for other factors that will determine whether a person faces financial hardship in repaying their loan. A more sophisticated RB analysis would look at RBs by household and consider other factors that affect the ability to pay, such as the number of children, and household taxes and benefits. This should be addressed in future work looking at student finance.

5. Implications for analyzing the design and cost of an ICL using dynamic simulations

5.1. Introduction and conceptual issues

In this section, the implications of including dynamics for graduates and taxpayers with an ICL are considered. It is assumed that all students taking out an ICL have the same distribution of lifetime earnings as those who do not take up the ICL. Further, it is assumed, as in the previous section, that student loans are log-normally distributed with an average loan of $35,000 and a standard deviation of $20,000. It is also assumed that loan size is correlated with the first 10 years of graduate earnings with a correlation of 0.3. The empirical work shows that allowing for labor market dynamics significantly reduces the estimated costs of introducing an ICL in the US.

Why is this the case conceptually? To explain it, two extreme cases are considered.

The first is where there is complete mobility, i.e. earnings or income outcomes are completely random each year. This would mean that approximately half of every graduate's time in the labor market would involve receiving income or earnings above the median and half below the median. With most ICLs, this would mean that every graduate would pay off their loan in full.

Conversely, with no mobility, the same people would be in the top half of the income or earnings distribution every year and, with most ICLs, pay off their loan. There is no advantage to the taxpayer of them staying above the median after the loan is repaid. Conversely, graduates in the bottom of the income or earnings distribution will contribute nothing towards the loan as they will earn below the repayment threshold every year. Just one graduate moving from just below the repayment threshold to just above the repayment threshold and being replaced in the earnings or income distribution by a graduate who has paid off their loan will increase government revenue. In reality, not all students would take out ICLs, and loan amounts vary by student, so there are cases where this may not necessarily be true, but these would be unusual. The illustration in the next section allows for variation in loan amounts but assumes the lifetime earnings of debtors and non-debtors are the same.

5.2. Hypothetical ICL parameters

To illustrate the impact of including dynamics, the Stafford Loan interest and government cost of borrowing parameters are used in conjunction with other ICL parameters. It is assumed that there is:

(i) A first income repayment threshold of $17,000 per year and a second threshold of $35,000 (in a policy reality, these would both be uprated annually with inflation). The $17,000 threshold is similar to that used with the income-based repayment (IBR) scheme currently operating in the US.
(ii) A marginal 3% repayment rate on earnings above the first threshold and 10% marginal for earnings above the second threshold. Again, the 10% marginal rate is similar to that used with the current IBR scheme.
(iii) A zero real interest rate while a student is at college and while a graduate is below the first income threshold (i.e. debt increases with inflation only); and then a real interest rate equal to the current Stafford Loan rate, which is 4.45% nominal or 2.45% real. With the means-tested component of the Stafford Loan, a zero real interest rate applies while students are at college. No means testing is used for this simulation.
(iv) An inflation rate of 2% and a government cost of borrowing of 2.4% nominal or 0.4% real.
(v) A loan write-off after 25 years.

To compare the full distributional implications of this ICL as well as the size of the taxpayer subsidy, the *earnings* simulations from Sections 2.5 and 3.4 are used. Taxpayer costs are estimated both under the assumption of no dynamics and when dynamics are incorporated. The taxpayer subsidy is calculated by pooling the male and female results using current BA enrolment proportions, as highlighted earlier in the paper. All costs and repayments are discounted back to when the student takes out the loan at age 18 and are in $US 2017 prices. The taxpayer subsidy is calculated by comparing the net present value (NPV) of repayments (which depends on future earnings simulations and ICL parameters) with the NPV of providing the loans (which depends on the number of loans taken out). Real earnings growth of 1% per year is assumed for all graduates throughout their working life. Inflation is assumed to be 2%, and the government cost of borrowing is set to the current 10-year US bond rate. This is currently used to determine the Stafford Loan interest rate, which is set at the government cost of borrowing (currently 2.4% nominal or 0.4% real) plus 2.05 percentage points.

5.3. Empirical results

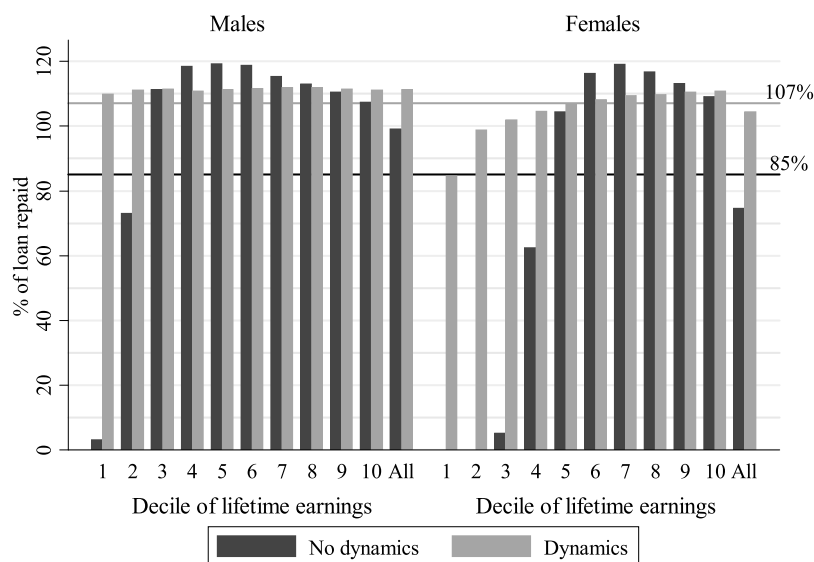Fig. 16 shows the distributional impact of ICL repayments (by

Fig. 16. Percentage of ICL repaid by decile of lifetime earnings: dynamics vs no dynamics
Note: Overall graduate contribution shown by horizontal lines. The subsidy is calculated by taking this amount away from 100%.

deciles of the male and female college *lifetime earnings* distributions[31]) under the assumptions of no dynamics and of dynamics. The analysis shows that when earnings dynamics are ignored, the estimated taxpayer subsidy for this ICL is around 15%; when dynamics are included, the estimated taxpayer subsidy is −7%. All graduates receive a taxpayer subsidy in this scheme while they are at college and while they earn below the first threshold. However, once they are above the first threshold, they receive no subsidy as the interest rate is 2.05 percentage points above the government cost of borrowing, so they are net contributors. Those who do not repay their loan within 25 years may also receive a subsidy due to the loan write-off. The difference in estimates of the taxpayer subsidy is large as there is high earnings mobility in the US for BA graduates and, with the ICL operating over 25 years rather than the 10 years of the Stafford Loan, there is a much higher chance of individuals making some repayments. Of course, if the earnings simulations involve too much mobility (as is likely), then the taxpayer subsidy will be underestimated. However, in the example shown above, the mobility would have to have been massively overestimated before the ICL would involve any taxpayer subsidy.

The extent of the difference in estimated taxpayer subsidy across the distribution also depends on other factors such as the size of the ICL loan and the ICL loan parameters (see Barr et al., 2019 and Britton et al., 2019) as well as the correlation between the total student loan taken out and future earnings. For example, if there is no correlation between loan size and earnings in the first 10 years, the estimate of the taxpayer subsidy increases to 16% (no mobility) and −6% (mobility). If there is a perfect negative correlation between loan size and gross earnings in the first 10 years, the taxpayer subsidy estimates increase to 19% (no mobility) and −4% (mobility).

As shown in Barr et al. (2019) and Britton et al. (2019), this subsidy can be reduced or increased very easily by varying the ICL parameters, including introducing a surcharge on the loan, changing the interest rate and/or changing other ICL parameters such as repayment thresholds and repayment rates. However, given the income mobility of BA graduates in the US, it is clear that a well-designed ICL could work (see Barr et al. (2019) for more details) and would have considerable advantages over the current Stafford Loan system. Additional work simulating the earnings dynamics of two-year college graduates and

dropouts suggests a similarly designed ICL could work beyond BA graduates. Of course, tight regulation of loans would need to be implemented, particularly with the for-profit sector, but this is also true with the current US loan system.

6. Conclusions

This paper reviews the empirical approaches that are needed to both evaluate and design student loan systems. An innovation of the paper is that it has suggested relatively straightforward methods for improving income and earnings simulation when data is poor (e.g. the data has banded income or good panel data is not available in the country). Another innovation is that the method proposed extends work that is routinely, and arguably inaccurately, done in countries evaluating student loan systems.

The paper shows that for RB analysis, it is generally better to use raw percentile estimates of income or earnings by age and gender and age smoothing rather than UQR methods. Having banded income data (as is the case in countries such as Japan) does not appear to be a significant problem for all but the highest earners, and RB analysis (or indeed student loan design) is not affected by the grouping of income or earnings data.

The paper shows how income and earnings dynamics can be easily introduced even with short panels that have a minimum of two observations for the same individual. This involves using copula functions, which better capture the complex dependence between income or earnings over one year. With traditional dynamic panel data methods, this is only possible to do reliably with longer panels.

Critically, the paper highlights the importance of including dynamics in both assessing the RBs associated with current loan systems and designing ICLs. Ignoring dynamics will firstly underestimate the proportion of individuals facing repayment hardship with a TBRL and secondly result in overestimating the taxpayer costs of an ICL.

In the case of the US, the modeling shows that the typical cross-sectional approach significantly understates the incidence of "excessive" RBs. The cross-sectional approach suggests that just under 50% of graduates with loans will not face RBs greater than 18%, whereas the dynamic simulations suggest that this proportion is closer to 15%. Further, the static cross-sectional simulation approach exaggerates the extent of the subsidies associated with ICLs by around 20 percentage points.

From the methodology employed, both illustrations imply strongly

---

[31] The lifetime earnings of all individuals from the ages of 23 to 65 are summed without discounting and deciles are constructed from this measure.

that the typically used empirical approaches to student loan policy assessment and design have the potential to significantly understate the benefits for both graduates and governments of the use of ICL compared with TBRL. This is particularly true in countries such as the US with high income and earnings mobility, but has pertinence for all countries considering the introduction of ICLs.

Acknowledgments

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.econedurev.2018.11.003.

Appendix A

Table A1
Distribution of grouped income variable: BA graduates, 2014–17 CPS.

Income lower bound ($US per year)	Income upper bound ($US per year)	Percentage of males	Percentage of females
0	0	2.95	6.65
0	5,000	2.44	8.11
5,000	10,000	1.73	3.67
10,000	20,000	4.63	8.19
20,000	30,000	5.92	8.39
30,000	40,000	7.16	10.12
40,000	50,000	8.25	10.84
50,000	60,000	8.00	9.36
60,000	70,000	7.67	7.75
70,000	85,000	10.47	8.49
85,000	110,000	13.20	8.66
110,000	135,000	8.20	3.96
135,000	175,000	8.48	2.97
175,000		10.90	2.84
Sample size		64,376	78,009

Appendix B. Kendall's tau (τ)

In the paper, all measures of rank correlation/dependence use Kendall's tau (τ). This is a measure of association based on the number of concordant, discordant and tied pairs there are in the cumulative distribution of income at age t (u_t) and $t + 1$ (u_{t+1}) in the CPS panel. A pair of CDFs $\{(u_{ti}, u_{t+1i}), (u_{tj}, u_{t+1j})\}$ are:

- concordant if $u_{ti} < u_{tj}$ and $u_{t+1i} < u_{t+1j}$ or $u_{ti} > u_{tj}$ and $u_{t+1i} > u_{t+1j}$; the number of concordant pairs is denoted by n_c;
- discordant if $u_{ti} > u_{tj}$ and $u_{t+1i} < u_{t+1j}$ or $u_{ti} < u_{tj}$ and $u_{t+1i} > u_{t+1j}$; the number of discordant pairs is denoted by n_d;
- tied if $u_{ti} = u_{tj}$ or $u_{t+1i} = u_{t+1j}$; the numbers of tied pairs are denoted by n_t and n_{t+1} respectively.

Kendall's tau-b rank correlation (τ), which is used in the paper, is given by:

$$\tau = \frac{n_c - n_d}{\sqrt{(n_c + n_d + n_t)(n_c + n_d + n_{t+1})}}$$

and the total number of pairs that can be constructed and compared for a sample of size T is:

$$n = {}^{1}/_{2}T(T - -1).$$

References

Armstrong, S., Dearden, L., Kobayashi, M., & Nagase, N. (2019). Student loans in Japan: Current problems and possible solutions. *Economics of Education Review, 71*, 120–134. https://doi.org/10.1016/j.econedurev.2018.10.012.

Barr, N., Chapman, B., Dearden, L., & Dynarski, S. (2019). The US college loans system: Lessons from Australia and England. *Economics of Education Review, 71*, 32–48. https://doi.org/10.1016/j.econedurev.2018.07.007.

Bonhomme, S., & Robin, J-M. (2009). Assessing the equalizing force of mobility using short panels: France, 1990-2000. *Review of Economic Studies, 76*(1), 63–92.

Britton, J., van der Erve, L., & Higgins, T. (2019). Income contingent student loan design: Lessons from around the world. *Economics of Education Review, 71*, 65–82. https://doi.org/10.1016/j.econedurev.2018.06.001.

Chapman, B., & Dearden, L. (2017). Conceptual and empirical issues for alternative student loan designs: The significance of loan repayment burdens for the US. *The ANNALS of the American Academy of Political and Social Science, 671*(1), 249–268.

Chapman, B., & Lounkaew, K. (2015). An analysis of Stafford Loans repayment burdens. *Economics of Education Review, 45*(3), 89–102.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatter-plots.

Journal of the American Statistical Association, 74(368), 829–836.

Dearden, L., Fitzsimons, E., Goodman, A., & Kaplan, G. (2008). Higher education funding policy. *The Economics Journal, 118*(526), F100–F125.

Demarta, S., & McNeil, A. J. (2005). The t copula and related copulas. *International Statistical Review, 73*(1), 111–129.

Firpo, S., Fortin, N., & Lemieux, T. (2009). Unconditional quantile regressions. *Econometrica, 77*(3), 953–973.

Higgins, T., & Sinning, M. (2013). Modeling income dynamics for public policy design: An application to income contingent student loans. *Economics of Education Review, 37*, 273–285.

Looney, A., & Yannelis, C. (2019). How Useful Are Default Rates? Borrowers with Large Balances and Student Loan Repayment. *Economics of Education Review, 71*, 135–145. https://doi.org/10.1017/j.econedurev.2018.10.004.

Salmi, J. (2003). Student loans in an international perspective: The World Bank experience. World Bank, Working Paper 27295.

Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut Statistique de l'Université de Paris, 8*, 229–231.

Stewart, M. B. (1983). On least squares estimation when the dependent variable is grouped. *Review of Economic Studies, 50*(4), 737–753.