For Online Publication

Supplementary Appendix for:
Liquidity Constraints, Fiscal Externalities and Optimal Tuition Subsidies

Nicholas Lawson

Université du Québec à Montréal

lawson.nicholas@uqam.ca

A   Introduction

In this supplementary online appendix, I present several results used in the calculations in the paper, as well as a series of sensitivity analyses. I begin in appendix B by showing that the welfare derivative presented in section 2 applies much more generally, by comparing it with results from Lawson (2015). Appendix C then examines how the liquidity ratio in section 2.2 can be approximated with heterogenous liquidity constraints. Appendix D presents the calculations leading to the elasticity of average earnings with respect to student grants, $\varepsilon_{\bar{Y}b}$. Appendix E considers a number of alterations to the model, and finds that the main conclusions and numerical results in

the paper are strongly robust. Finally, Appendix F presents a simple model in which income taxes impose distortions on multiple margins, including both education and job search, and shows that corrective policies on each margin are required; an alteration of UI benefits to correct tax distortions on job search does not eliminate the effect of fiscal externalities on optimal tuition subsidies, as fiscal externalities affect optimal policies independently on both margins.

References to numbered equations, tables, etc. not found in this appendix refer to the numbers from the main paper.

B Comparison of Welfare Derivative with General Formula from Lawson (2015)

In section 2, I solved a specific model of college education for the derivative of social welfare with respect to the tuition subsidy denoted as (3). While the model was simplified and stylized, the results remain very general, as demonstrated by the analysis of Lawson (2015), where a very similar equation is derived from a general model that applies to any government transfer program. Consider the following equation from page 21 of Lawson (2015):

$$\frac{dW}{db_j} = D_j \left[\frac{E_j[U'(c)] - E_{\bar{y}}[U'(c)]}{E_{\bar{y}}[U'(c)]} - \sum_{l=1}^{M} \frac{D_l b_l}{D_j b_j} \left(\varepsilon_{b_j}^{D_l} - \varepsilon_{b_j}^{\bar{y}} \right) \right]. \tag{B.1}$$

In the setting of that paper, there are M different transfer programs, with (B.1) providing the welfare derivative for program j. D_j is discounted average percentage of time spent on program j, while $E_j[U'(c)]$ represents the average marginal utility of all individuals on program j, and $E_{\bar{y}}[U'(c)]$ is average income-weighted marginal utility for all individuals. Finally, $\varepsilon_{b_j}^{D_l}$ represents the elasticity of time spent on program l with respect to benefit b_j, while $\varepsilon_{b_j}^{\bar{y}}$

is the elasticity of the tax base (average income) with respect to b_j.

In the setting of the current paper, there are three "programs": the tuition subsidy b, the additional education-related expenditure p, and the exogenous spending G, and the latter is unresponsive to b by assumption. When considering the welfare derivative with respect to b, D_j is equal to $\frac{S}{R_1}$ (the general model assumes a single continuous period, whereas the current model has 12 periods), and it is clear that $\sum_{l=1}^{M} \frac{D_l b_l}{D_j b_j} \left(\varepsilon_{b_j}^{D_l} - \varepsilon_{b_j}^{\bar{y}} \right) = \left(\frac{b+p}{b} \right) \varepsilon_{Sb} - \left(1 + \frac{G+Sp}{Sb} \right) \varepsilon_{\bar{Y}b}$.

This only leaves the marginal utility terms. First, $E_j[U'(c)] = E_1[u'(c_{ui})]$. $E_{\bar{y}}[U'(c)]$, meanwhile, is the average marginal utility of all individuals weighted by labour income, and this is identical to v_c^*. Therefore, equation (B.1) is identical in the notation of the current paper to:

$$\frac{dW}{db} = \frac{S}{R_1} \left[\frac{E_1[u'(c_{ui})] - v_c^*}{v_c^*} - \left(\frac{b+p}{b} \right) \varepsilon_{Sb} + \left(1 + \frac{G+Sp}{Sb} \right) \varepsilon_{\bar{Y}b} \right]. \quad \text{(B.2)}$$

There are only two differences between (B.2) and (3): an alternative normalization by S rather than $\frac{S}{R_1}$, which simply rescales the welfare derivative, and the substitution of $v_c^0(c_v^0, l_0)$ for v_c^*. The latter is a conservative assumption allowing empirical evaluation using estimates of the liquidity effect, which depends on $v_c^0(c_v^0, l_0)$. Therefore, aside from this slight modification, the results from the general model in Lawson (2015) apply directly to the tuition subsidy case.

C   Liquidity Ratio with Heterogeneous Constraints

To be as general as possible, let me allow for the possibility that η_i and A_i are jointly distributed according to some bivariate distribution function $F(\eta, A)$. Let me define $S_A(A)$ to be the probability of college attendance for an indi-

vidual with debt limit A; this can be written as:

$$S_A(A) = 1 - F_{\eta|A}[R_1 v(c_v^0, l_0) - u(c_u(A)) - R_2 v(c_v^1(A), l_1(A))|A]$$

where $F_{\eta|A}$ represents the conditional cdf. Then the overall probability of college attendance is simply $S = \int_A S_A(A) f_A(A) dA$, where f_A is the marginal density of A.

Next, observe that:

$$\frac{\partial S}{\partial b} = \int_A \frac{\partial S_A(A)}{\partial b} f_A(A) dA = \int_A f_{\eta|A}(\eta_A^*|A) f_A(A) u'(c_u(A)) dA$$

$$\frac{\partial S}{\partial a_1} = \int_A \frac{\partial S_A(A)}{\partial a_1} f_A(A) dA = \int_A f_{\eta|A}(\eta_A^*|A) f_A(A)[u'(c_u(A)) - v_c(c_v^0, l_0)] dA$$

where η_A^* is the critical value for $A_i = A$. Therefore, using the definition of L from the text:

$$L = \frac{\int_A f_{\eta|A}(\eta_A^*|A) f_A(A)[u'(c_u(A)) - v_c(c_v^0, l_0)] dA}{\int_A f_{\eta|A}(\eta_A^*|A) f_A(A) v_c(c_v^0, l_0) dA}.$$

Meanwhile, the term I wish to replace is $\frac{E_1[u'(c_{ui})] - v_c(c_v^0, l_0)}{v_c(c_v^0, l_0)}$; this is greater or less than L as:

$$E_1[u'(c_{ui})] \gtrless \frac{\int_A f_{\eta|A}(\eta_A^*|A) f_A(A) u'(c_u(A)) dA}{\int_A f_{\eta|A}(\eta_A^*|A) f_A(A) dA}$$

$$\frac{\int_A [1 - F_{\eta|A}(\eta_A^*|A)] f_A(A) u'(c_u(A)) dA}{\int_A [1 - F_{\eta|A}(\eta_A^*|A)] f_A(A) dA} \gtrless \frac{\int_A f_{\eta|A}(\eta_A^*|A) f_A(A) u'(c_u(A)) dA}{\int_A f_{\eta|A}(\eta_A^*|A) f_A(A) dA}.$$

If the conditional hazard rate $\frac{f_{\eta|A}(\eta_A^*|A)}{1 - F_{\eta|A}(\eta_A^*|A)}$ is constant, these two terms will be equal, and I can safely replace $\frac{E_1[u'(c_{ui})] - v_c(c_v^0, l_0)}{v_c(c_v^0, l_0)}$ with L in (3). More generally, let me continue by substituting $h(A)$ for the conditional hazard rate, and let me also write $k(A) = [1 - F_{\eta|A}(\eta_A^*|A)] f_A(A)$ to represent the measure of enrollees at a particular value of A_i; then the comparison becomes:

$$\frac{\int_A k(A) u'(c_u(A)) dA}{\int_A k(A) dA} \gtrless \frac{\int_A k(A) h(A) u'(c_u(A)) dA}{\int_A k(A) h(A) dA}$$

4

$$\frac{\int_A k(A)h(A)dA}{\int_A k(A)dA} \frac{\int_A k(A)u'(c_u(A))dA}{\int_A k(A)dA} \underset{<}{\overset{\geq}{}} \frac{\int_A k(A)h(A)u'(c_u(A))dA}{\int_A k(A)dA}$$

$$E_1[h(A)]E_1[u'(c_u(A))] \underset{<}{\overset{\geq}{}} E_1[h(A)u'(c_u(A))]$$

$$0 \underset{<}{\overset{\geq}{}} Cov_1[h(A), u'(c_u(A))].$$

Therefore, if the covariance of the hazard and the marginal utility among students is close to zero, it will be a reasonable approximation to insert L into (3). Meanwhile, I will tend to overestimate the liquidity effect if the covariance is positive, which would follow, for instance, if $h(A)$ is decreasing in A (given that $u'(c_u(A))$ should be non-increasing in A).

Given that η_A^* is decreasing in A, I would want the hazard to be increasing in η in order for the liquidity effect to be overestimated, thus making my results on the importance of liquidity constraints an upper bound. Certain distributions, such as the Pareto and the χ^2 for degrees of freedom less than 2, feature decreasing hazards, but many other distributions, including the logistic that I use in my calibration, feature an increasing hazard, in which case I will tend to overestimate the importance of L.

D   Calculation of $\varepsilon_{\bar{Y}b}$

First, assuming that the only effects of b on $\bar{Y}$ are from b's effect on schooling and from the effect of the tax change on earnings Y_{01} and Y_{11}, I can write:

$$\varepsilon_{\bar{Y}b} = \frac{b}{\bar{Y}}\frac{d\bar{Y}}{db} = \frac{b}{\bar{Y}}\left[\frac{\partial \bar{Y}}{\partial S}\frac{dS}{db} + \frac{\partial \bar{Y}}{\partial \tau}\frac{d\tau}{db}\right].$$

It is clear that $\frac{\partial \bar{Y}}{\partial S} = \gamma_2 Y_{11} - \gamma_1 Y_{01} = [\gamma_2(1.08)^4 - \gamma_1]Y_{01}$, and given that I assume that the elasticity of taxable income is 0.4, I have $\frac{\partial \bar{Y}}{\partial \tau} = -0.4\frac{\bar{Y}}{1-\tau}$. Using (2) for $\frac{d\tau}{db}$ and assuming $p = 0$, the equation for $\varepsilon_{\bar{Y}b}$ becomes:

$$\varepsilon_{\bar{Y}b} = \left[\gamma_2(1.08)^4 - \gamma_1\right]Y_{01}\frac{S}{\bar{Y}}\varepsilon_{Sb} - 0.4\frac{Sb}{(1-\tau)\bar{Y}}\left[1 + \varepsilon_{Sb} - \left(1 + \frac{G}{Sb}\right)\varepsilon_{\bar{Y}b}\right]$$

5

and rearranging, I arrive at:

$$\varepsilon_{\bar{Y}b} = \left[\frac{[\gamma_2(1.08)^4 - \gamma_1]S}{\gamma_2(1.08)^4 S + \gamma_1(1 - S)} - \frac{0.4\tau}{1 - \tau}\left(1 + \frac{G}{Sb}\right)^{-1} \right] \frac{1 - \tau}{1 - 1.4\tau}\varepsilon_{Sb}$$

$$-\frac{0.4\tau}{1 - 1.4\tau}\left(1 + \frac{G}{Sb}\right)^{-1}.$$

Inputting $\frac{G}{Sb} = 88.41$ into the equation above, I find that $\varepsilon_{\bar{Y}b}$ takes a value of 0.0063 when $\varepsilon_{Sb} = 0.1$, and 0.0142 when $\varepsilon_{Sb} = 0.2$, as presented in Table 1.

E   Robustness Analyses

This section will be devoted to an examination of the robustness of my results. I begin with an analysis of optimal policy when G is a public good rather than an exogenous quantity of required spending, and then I examine the sensitivity of my results to the coefficient of relative risk-aversion. Next, I use the estimates of fiscal costs and benefits from Trostel (2010) to assess the impact on my conclusions of how these fiscal effects are modelled, and I also examine the optimal policy when part of the government's tax revenue pays for a lump-sum transfer, so that the marginal tax rate is not equal to the average tax rate. I also extend the model to consider uncertainty about future incomes, as well as heterogeneity in liquidity constraints and returns to education. The quantitative results are only slightly altered in each case, and the qualitative conclusions remain very similar: in the baseline case, it is always optimal to raise tuition subsidies to at least offset the value of tuition, and the effects of liquidity constraints on optimal subsidies are small and often negative.

E.1   Endogenous G

In the main model of the paper, I assume that G is an exogenous amount of required government spending above and beyond tuition subsidies. However,

the results are nearly identical if G is instead a public good chosen optimally by the government. To demonstrate this, I assume that each individual receives utility $a \ln(G)$ from the total amount G of public good provided.

First, consider the results of the sufficient statistics analysis: if both τ and G are allowed to change in response to b, the welfare derivative in (1) becomes:

$$\frac{dV}{db} = E\left(\frac{\partial V_i}{\partial b}\right) + E\left(\frac{\partial V_i}{\partial \tau}\right)\frac{d\tau}{db} + \frac{a}{G}\frac{dG}{db}.$$

Next, consider that, if G has been chosen optimally, the following first-order condition must be satisfied:

$$\frac{dV}{dG} = \frac{a}{G} + E\left(\frac{\partial V_i}{\partial \tau}\right)\frac{d\tau}{dG}|_{db=0} = 0.$$

This means I can rewrite the welfare derivative with respect to b as:

$$\frac{dV}{db} = E\left(\frac{\partial V_i}{\partial b}\right) + E\left(\frac{\partial V_i}{\partial \tau}\right)\left[\frac{d\tau}{db} - \frac{d\tau}{dG}|_{db=0}\frac{dG}{db}\right]$$

and the term in square brackets is equal to $\frac{d\tau}{db}|_{dG=0}$, the adjustment in taxes required to pay for an increase in b holding G constant, and so:

$$\frac{dV}{db} = E\left(\frac{\partial V_i}{\partial b}\right) + E\left(\frac{\partial V_i}{\partial \tau}\right)\frac{d\tau}{db}|_{dG=0}.$$

The adjustment in taxes to pay for b holding G constant is exactly what I used in my calculations in section 2.2. The essential point is that, if G is chosen optimally, the marginal welfare impact of changing it must be equal to the marginal welfare impact of changing τ, which means that it doesn't matter for welfare purposes which is adjusted, and the welfare derivative is identical to that derived earlier.

As b changes, however, the government has the option of changing either τ or G, or both, to balance the budget, so the welfare gain from raising b would be at least as large if G is endogenous. However, there are no credible empirical

7

estimates of how G should change with b; this depends on the structure of the model, so I can only evaluate it using my calibrated structural model. It is easy to incorporate endogenous G into the model: I simply calibrate the individual parameters as before, and then find the value of a that makes $G = 68.606$ optimal; the resulting values range from 3.2364 to 3.2427 depending on the case. Then I can numerically evaluate welfare as a function of both b and G and find the optimum, which is displayed in Table E.1 below.

Table E.1: Results from Calibration and Simulation with Endogenous G

ε_{Sb}	$\dfrac{\partial \hat{S}}{\partial a_1}$	
	0	0.0021
	B. Optimal Combined b and G	
0.1	$7047 & 69.194	$6537 & 68.822
0.2	$9441 & 70.789	$9311 & 70.565
	C. Welfare Gains from Moving to Optimum	
0.1	$1240 (39.9%)	$1409 (45.4%)
0.2	$4727 (152.3%)	$4600 (148.2%)

As expected, the optimal values of b and welfare gains are slightly higher, though only by a few hundred dollars at most. However, it is slightly more surprising that when b increases, the optimal G actually increases slightly, rather than decreasing to help pay for the increased tuition subsidies. This is the result of an income effect: with increased college enrollment and loosened liquidity constraints, average consumption has increased, and marginal utility of income is lower on average. Therefore, the average individual is more willing to substitute towards the public good, and the welfare-maximizing value of G increases slightly. In any case, the changes in optimal policy are minimal relative to the baseline estimates.

E.2 Increased Risk-Aversion

My second sensitivity analysis considers how the results change when I specify a coefficient of relative risk-aversion of $\rho = 2$ in employment. Since I only need to specify this parameter when using the structural method, it will only affect my simulation results. Calibration proceeds as before, assuming $p = 0$ as usual, and simulation yields the results displayed in Table E.2. The optimal values of b and welfare effects are smaller in most cases, but the conclusion of more than offsetting median public tuition continues to hold in the baseline case, and the effects of liquidity constraints remain small and negative.

Table E.2: Results from Calibration and Simulation for $\rho = 2$

	$\frac{\partial \hat{S}}{\partial a_1}$	
ε_{Sb}	0	0.0021
	A. Numerical Estimate of $\frac{dW}{db}$ at $b = 2$	
0.1	0.2097	0.2564
0.2	0.4085	0.4307
	B. Optimal Student Grants	
0.1	$4127	$3707
0.2	$7806	$7611
	C. Welfare Gains from Moving to Optimum	
0.1	$572 (18.4%)	$808 (26.0%)
0.2	$2748 (88.5%)	$2702 (87.1%)

E.3 Evidence from Trostel (2010) on Fiscal Effects of Education

In this subsection, I will test the robustness of my results to a different value of p, which has been set to zero in all other numerical results in the paper. I perform my analysis again using the most pessimistic estimates from Trostel (2010), in which he concludes that each year of college costs the government

$17845 and saves expenditures amounting to $13955 in present value. Accounting for the fact that $2000 per year of the costs are already included in b, I therefore set $p = 2$, and evaluating (4) and using the same statistical extrapolations as before leads to the results displayed in Table E.3. The values of $\frac{dW}{db}$ are smaller now, as are the optimal grants and the welfare gains; however, the optimal grants all still represent significant increases from $b = 2$, and the baseline finding is over $1000 above and beyond the value of tuition. The effects of liquidity constraints on optimal policy are again small in the baseline case.

Table E.3: Results from Sufficient Statistics and Extrapolation with $p = 2$

	$\frac{\partial \hat{S}}{\partial a_1}$	
ε_{Sb}	0	0.0021
	A. Estimate of $\frac{dW}{db}$ at $b = 2$	
0.1	0.1346	0.1817
0.2	0.3213	0.3435
	B. Optimal Student Grants	
0.1	$4982	$7339
0.2	$6920	$7253
	C. Welfare Gains from Moving to Optimum	
0.1	$607 (19.6%)	$1275 (41.1%)
0.2	$2168 (69.9%)	$2468 (79.5%)

Calibration and simulation follows the same procedure as before, and the results are found in Table E.4. In every case, the welfare derivative at baseline is smaller, as are the optimal grants and the welfare gains from moving to the optimum; the optimal grants drop by about $1700, but the baseline result still involves a subsidy greater than tuition, and once again stronger liquidity constraints have a small negative effect.

Table E.4: Results from Calibration and Simulation with $p = 2$

	$\frac{\hat{\partial S}}{\partial a_1}$	
ε_{Sb}	0	0.0021
A. Numerical Estimate of $\frac{dW}{db}$ at $b = 2$		
0.9	0.1546	0.2016
1.8	0.3299	0.3521
B. Optimal Student Grants		
0.9	$5225	$4734
1.8	$7497	$7380
C. Welfare Gains from Moving to Optimum		
0.9	$635 (20.5%)	$909 (29.3%)
1.8	$2686 (86.5%)	$2693 (86.8%)

E.4 Marginal Tax Rate $\neq$ Average Tax Rate

For simplicity, the main model in the paper features a single flat tax rate applied to all earned income. In reality, most developed countries' tax systems features multiple tax brackets, including an exemption of some initial amount of earnings from taxation. In this appendix, I consider a setting in which there is still a single marginal tax rate, but in which part of the spending financed by this tax is a lump-sum transfer to all individuals, making the average tax rate lower than the marginal rate.

To be precise, since my calibration corresponds to 2007, I use the average tax rates for that year. In a publication entitled "Average Federal Tax Rates in 2007," the Congressional Budget Office estimates that the average household paid 20.4% of their income in federal taxes in 2007; however, this includes excise and corporate taxes, which I exclude from my analysis, and a larger portion of social insurance taxes than considered in my analysis. On average, federal income tax was 9.3% of income, to which I add 3% for the Medicare tax, as well as 1.9% in state income taxes, which is the average reported by

the Institute on Taxation & Economic Policy for 2007, leading to an overall average tax rate of 14.2%.

Thus, I assume that all individuals receive a lump-sum transfer equal to m in each period, which I hold exogenously fixed. I find the m that makes the average tax rate equal to 14.2% for employed individuals: $0.142 = 0.23 - \frac{\hat{S}\gamma_2 + (1-\hat{S})\gamma_1}{\bar{Y}}m$, which means $m = \frac{0.088\bar{Y}}{\hat{S}\gamma_2 + (1-\hat{S})\gamma_1} = \$5451$. I can then back out G as the remainder of tax revenues not spent on m or on tuition subsidies.

The results with the sufficient statistics method are actually identical to baseline, because there what matters are the elasticities and the size of government, which haven't changed (if the size of government is reinterpreted to include the lump-sum transfers). The equation for $\frac{dW}{db}$ takes exactly the same form as in (4), except that G is replaced by $G + \gamma_1 m$, which remains an exogenous quantity of required government spending.

In the calibration case, the changes are modest since the parameters of the model are calibrated to a given consumption gap, but introducing m changes the calibrated value of the risk-aversion coefficient θ. Table E.5 presents the numerical results, and the welfare derivatives at $b = 2$ are smaller than at baseline, because the calibration implies a considerably larger θ, so that the marginal value of an additional dollar to a student is smaller; the welfare gains at the optimum are unsurprisingly lower as a result. However, the optimal subsidies are nearly identical to the baseline case, lower by no more than \$194.

Thus, the main results of the paper are unaffected by this alteration. And it should be noted that an alternative, and perhaps more realistic, specification of the tax system generates even larger optimal subsidies and welfare gains: if the lump-sum transfer is available only to employed individuals, and not to students (who the model assumes have no labour earnings), then there is an added fiscal bonus of tuition subsidies in that they prevent students from

12

Table E.5: Results from Calibration and Simulation with Lump-Sum Transfers

	$\frac{\partial \hat{S}}{\partial a_1}$	
ε_{Sb}	0	0.0021
	A. Numerical Estimate of $\frac{dW}{db}$ at $b = 2$	
0.9	0.1191	0.1739
1.8	0.3386	0.3656
	B. Optimal Student Grants	
0.9	$6848	$6281
1.8	$9150	$9036
	C. Welfare Gains from Moving to Optimum	
0.9	$724 (23.3%)	$566 (18.2%)
1.8	$3451 (111.2%)	$3181 (102.5%)

entering the workforce quickly and gaining the right to a transfer. In that case, the optimal subsidies increase by at least $2500 over baseline in the sufficient statistics analysis, and by at least $4000 in the calibration analysis; full results are available upon request. Therefore, the main results of the paper are robust to the exact specification of the tax system: tuition subsidies should increase significantly, and liquidity constraints have little impact on optimal policy.

E.5 Income Uncertainty

Next, I consider a case with uncertainty about future incomes. To keep the problem simple, I assume that all uncertainty is resolved after the first period. Thereafter, educated individuals receive either $Y_{1tH} = (1 + g)^{t-1} Y_{11H}$ in each period or $Y_{1tL} = (1 + g)^{t-1} Y_{11L}$, each with probability 0.5, where $Y_{11H} > Y_{11L}$ and $\frac{Y_{11H} + Y_{11L}}{2} = Y_{11}$. Meanwhile, uneducated workers begin with Y_{01} in the first period, and thereafter receive $Y_{0tH} = (1 + g)^{t-1} Y_{01H}$ or $Y_{0tL} = (1 + g)^{t-1} Y_{01L}$, each with probability 0.5, where $\frac{Y_{01H} + Y_{01L}}{2} = Y_{01}$. The corresponding consumption values will be denoted as c^1_{vH} and c^1_{vL} for educated workers and c^0_{vH} and c^0_{vL} for uneducated workers, with c^0_{v1} representing the consumption of

13

first-period workers.

In deriving $\frac{dW}{db}$, the only meaningful change will come from the fact that $\frac{\partial V}{\partial \tau}$ takes a different form, specifically:

$$\frac{\partial V}{\partial \tau} = -\frac{\gamma_2}{2} S \left(v_c^1(c_{vL}^1, l_{1L}) Y_{11L} + v_c^1(c_{vH}^1, l_{1H}) Y_{11H} \right)$$

$$-\frac{1-S}{2} \left(v_c^0(c_{vL}^0, l_{0L})(Y_{01} + \gamma_2 Y_{01L}) + v_c^0(c_{vH}^0, l_{0H})(Y_{01} + \gamma_2 Y_{01H}) \right).$$

However, this equation cannot be used in its current form, and the most reasonable simplification is still $\bar{Y} v_c^*$, where $\bar{Y}$ remains equal to $S\gamma_2 Y_{11} + (1-S)\gamma_1 Y_{01}$, so that (4) holds in this case as well, and the results are unchanged.

I will therefore focus on the structural analysis. I assume that α_s and δ take the same values as the baseline case, but I also assume that first-period workers cannot adjust their labour supply from $l_{01} = 1$ (for example, assume that they are on an apprenticeship program of fixed labour intensity). Then, after the first period and the resolution of uncertainty, I allow workers of all types to solve for optimal labour supply, and I solve for the wages that generate $\{Y_{01L}, Y_{01H}, Y_{11L}, Y_{11H}\}$ in equilibrium.

The calibration then proceeds largely as before, with $p = 0$ as usual, except that A and θ must be chosen simultaneously to generate consumption choices which match $E(c_v^1) = 1.26 E(c_v^0)$ and $u'(c_u) = (\hat{L}+1)v'(c_{v1}^0)$. For the variability of income, I use the median and interquartile range of income for high school and college graduates from the CPS in the 4th quarter of 2012. Then I consider three cases: one case in which I choose the values of $\{Y_{0L}, Y_{0H}, Y_{1H}, Y_{1L}\}$ that produce the same interquartile range, specifically 74.3% of the median for high school graduates and 81.5% for college graduates, one case in which I cut the high school IQR in half, and one in which I cut the college IQR in half. The results are displayed in Table E.6, and while the welfare derivatives are

14

similar to baseline, the optimal grants and welfare gains are larger in every case, especially when the college wage variance is large; in those cases, raising grants allows graduates who end up in low-wage jobs to increase consumption from a low level, which has a significant positive impact on utility. Effects of liquidity constraints are small and negative in every case.

Table E.6: Results from Calibration and Simulation with Uncertain Income

	CPS Variance		Low HS Variance		Low College Variance	
			$\frac{\partial \hat{S}}{\partial a_1}$			
ε_{Sb}	0	0.0021	0	0.0021	0	0.0021
	A. Numerical Estimate of $\frac{dW}{db}$ at $b=2$					
0.1	0.2227	0.2695	0.2143	0.2523	0.2237	0.2706
0.2	0.4073	0.4296	0.4185	0.4317	0.4073	0.4296
	B. Optimal Student Grants					
0.1	$12230	$12140	$12492	$12424	$7622	$7179
0.2	$13520	$13438	$13827	$13758	$9680	$9576
	C. Welfare Gains from Moving to Optimum					
0.1	$3713	$3550	$4598	$4280	$1409	$1650
0.2	$7854	$7589	$9524	$9251	$4191	$4183

E.6 Heterogeneity in Liquidity Constraints and Two-Tier Grants

In appendix C, I examined how robust the sufficient statistics condition in (4) is to a distribution of debt limits; an alternative examination of the robustness of the results to heterogeneous liquidity constraints can be performed using a structural approach. I allow for two groups, each representing half of the population,[1] one of which is unconstrained while the other faces a debt limit A. I calibrate the model for $\{A, \theta, \mu, \sigma\}$ using the sufficient statistics as aver-

---

[1]Brown, Scholz and Seshadri (2012) find that approximately half of the children in their sample did not receive post-schooling cash transfers from their parents, which they claim as an indicator for student liquidity constraints.

ages, and then solve for the optimal lump-sum student grant, with the results displayed in Table E.7. The values of $\frac{dW}{db}$ are slightly smaller than in Table 3, which is to be expected because the logistic distribution for η has an increasing hazard (see appendix C). The optimal levels of b are also smaller, but only by $50 at most, and the welfare gains are slightly higher, while the strength of the liquidity constraints has a negative impact on subsidies.

Table E.7: Results from Calibration and Simulation with Heterogeneous Liquidity Constraints

ε_{Sb}	$\frac{\partial \hat{S}}{\partial a_1}$	
	0	0.0021
	A. Numerical Estimate of $\frac{dW}{db}$ at $b = 2$	
0.1	0.1951	0.2406
0.2	0.4107	0.4317
	B. Optimal Student Grants	
0.1	$6954	$6425
0.2	$9158	$9041
	C. Welfare Gains from Moving to Optimum	
0.1	$1375 (44.3%)	$1699 (54.7%)
0.2	$4696 (151.3%)	$4709 (151.7%)

With this calibrated model in hand, I could go one step further and consider what policy the government would want to set if they could observe individuals' debt limits; with two types of individuals, the government could introduce a two-tier grant system, with one grant amount b_1 for the constrained group and another amount b_2 for unconstrained students. However, when I numerically maximize welfare (still measured as equally-weighted utilitarian social welfare) over the pair (b_1, b_2), I find the same uniform grant scheme described in Table E.7; that is, $b_1 = b_2$, with both equal to the values listed there. The reason for this is that the parameters used suggest that a grant of $6000 or more is sufficient to remove any liquidity constraints that may exist, so for the

purpose of optimal policy, whether a particular group of people is constrained or not is irrelevant, because they will no longer be constrained as the subsidy approaches the optimum. This is another way of making the point that liquidity constraints are of second-order importance for welfare analysis of policy: in the current model, even a perfectly-informed government could not improve the outcome with policies specifically targetting constrained individuals.

E.7 Heterogeneous Returns to Education

I now investigate how sensitive the results are to allowing for heterogeneous returns to education. I assume that the college wage premium P (where $w_{11} = Pw_{01}$) follows some distribution $F_P(P)$, and to be precise I use a quadratic approximation to the marginal treatment effect distribution presented in Figure 4 of Carneiro, Heckman and Vytlacil (2011), as displayed in my Figure E.1.

Figure E.1: College Wage Premium Distribution
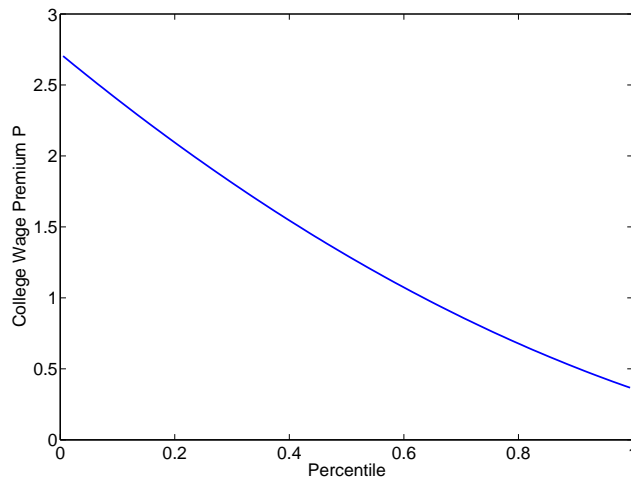
I divide the population into 100 equal masses denoted by $j = \{1, 2, ..., 100\}$, with wage premia equal to $\{F_P^{-1}(0.005), F_P^{-1}(0.015), ..., F_P^{-1}(0.995)\}$, and nor-

malize l_0 and l_{1j} to one for $j = 50$. Then I allow for a distribution of η for each mini-population, where η is allowed to be correlated with P. In particular, I let $\eta_{ij} = \bar{\eta}_j + \eta_i$, where $\bar{\eta}_j$ is deterministic for each j and η_i comes from a logistic distribution with mean 0 and scale parameter σ; therefore, college attendance is positively correlated with ability, but not perfectly, as idiosyncratic tastes for college still play a role.

I specify $\bar{\eta}_j$ to match the pattern of responsiveness to b found in Carneiro, Heckman and Vytlacil (2011); specifically, I assume $\bar{\eta}_j = U_0 - U_{1j} + z - \mu_s \left(\left(\frac{j-40.5}{100} \right)^3 + 0.395^3 \right)$, where $U_{1j} = u(c_{uj}) + R_2 v^1(c^1_{vj}, l_{1j})$, and where z is some constant. This generates the baseline pattern of responsiveness to b found in my Figure E.2, which compares closely to the dashed line with triangles in Figure 5 in Carneiro, Heckman and Vytlacil (2011), although the latter corresponds to a more abstract concept of increasing the entire set of instruments that determine the value of college attendance. As the tuition subsidy increases, the curve in Figure E.2 will tend to shift to the right, as individuals with lower returns to college are increasingly affected, and thus the marginal return to the subsidy will decrease.

Allowing for a distribution of wage premia makes it important to model the tax system more realistically: I assume that the state and Medicare tax rates do not vary with income, but I use an approximation to the US federal system in 2008, with a 15% marginal rate up to $41500 and a 25% rate beyond. To account for the personal exemption of $3500 and the standard deduction of $5450, as well as the fact that the first $8025 of taxable income is only taxed at a 10% rate, I assume a universal tax refund of $1743.75. To avoid discontinuities in the marginal tax schedule, I use a smoothed approximation to the tax rate between $39000 and $44000, specifically a sine connecting $\tau = 0.23$ at $39000 to $\tau = 0.33$ at $44000. I assume that the tax rate threshold moves up

Figure E.2: Responsiveness of Enrollment to b Across Distribution
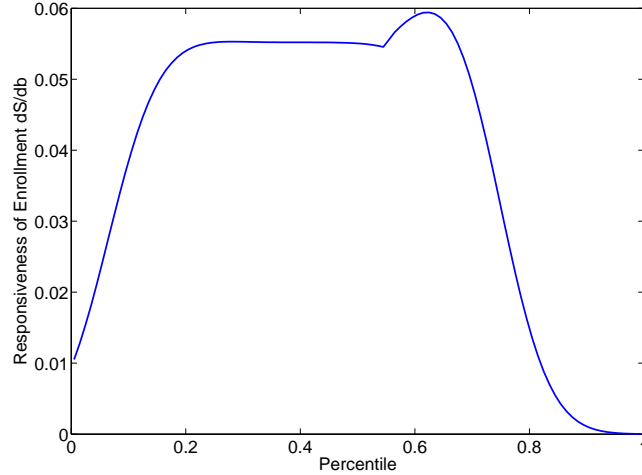
with wage growth, and that when taxes need to adjust to balance the budget, the base (state and Medicare) tax rate is the one that moves.

When calibrating, I select values for $\{A, \theta, \mu_s, \sigma, z\}$ in order to match five quantities, three of which are familiar: $E_1[u'(c_{uj})] = (\hat{L} + 1)v_c^0(c_v^0, l_0)$, $\hat{S} = 0.388$, and $\varepsilon_{Sb} = \{0.1, 0.2\}$, although in this case ε_{Sb} is interpreted as a partial derivative. I also choose z to generate a probability of attendance of 95% for the highest-return group, and I use the fact that college graduates consume 73.9% of their pre-tax income and high school graduates consume 83.4% to motivate setting $\frac{\frac{E_1(c_v^1)}{E_1(Y_1)}}{\frac{c_v^0}{Y_0}} = \frac{0.739}{0.834}$.

This leads to the results presented in Table E.8. The striking finding is that the welfare derivative at baseline is significantly larger, because the average return to education among those induced to go to school is higher using the estimates from Carneiro, Heckman and Vytlacil (2011), further suggesting that my assumption of an 8% return to a year of education was a conservative estimate. However, there are diminishing returns to inducing college

19

attendance, because increasingly generous grants induce students with lower monetary returns to go to school; therefore, optimal grants are lower when $\varepsilon_{Sb} = 0.2$, though they are larger when $\varepsilon_{Sb} = 0.1$ because the returns to inducing college attendance do not decline as quickly in that case. However, the qualitative conclusions are essentially identical to the baseline case: the optimal subsidy is greater than median public tuition in each case, and effects of liquidity constraints are small. Meanwhile, the welfare gains are significantly larger than before, amounting to $80.5 billion per year in the baseline case.

Table E.8: Results from Calibration and Simulation with Heterogeneous Returns to Education

ε_{Sb}	$\hat{L}$	
	0	$\frac{1}{3}$
A. Numerical Estimate of $\frac{dW}{db}$ at $b = 2$		
0.1	0.6102	0.6865
0.2	1.1968	1.2501
B. Optimal Student Grants		
0.1	$8912	$9455
0.2	$7236	$7478
C. Welfare Gains from Moving to Optimum		
0.1	$6946 (223.8%)	$8567 (276.0%)
0.2	$11504 (370.6%)	$12563 (404.7%)

This analysis provides us with a sense of how heterogeneity in returns can affect the results, but naturally has more of a "black box" character than the baseline analysis. The current method is not as well suited to answering questions about how financial aid could be better targetted at students on the margin of attending college or from groups with high returns; research with multiple dimensions of heterogeneity such as Abbott et al. (2013) provides a complementary analysis which provides answers to some such questions, and additional future work along these lines could also be useful.

F   A Simple Model of Multiple Fiscal Externalities

The analysis of optimal tuition subsidies in this paper, as well as the analysis of optimal unemployment insurance in Lawson (2016), starts from the assumption of a distortionary marginal income tax used to raise revenues for government spending. In each case, this tax also imposes distortions on another margin, such as the decision about how much education to obtain, or how hard to search for a new job: the tax reduces the return to both actions, leading to inefficiently low education and job search. As a result, when considering a policy that is targetted at one of these margins, we must account for the pre-existing distortion of taxation, and the optimal policy involves increased subsidies to education, and decreased unemployment benefits, to offset the distortion from income taxes.

In this appendix, I demonstrate that these results hold in a model with multiple margins in combination; specifically, I present a simple model in which income taxation affects a total of three margins: labour supply while employed, education, and job search while unemployed. With a distribution of income among employed individuals, there is a role for distortionary income taxes to provide for redistribution, and such taxation requires changes to education subsidies and unemployment insurance: as the government revenue requirement increases, the jointly optimal policy features rising marginal income tax rates and education subsidies, and decreasing unemployment insurance, exactly as found when each policy was considered on its own.

This analysis demonstrates that the main results in the present paper, as well as those in Lawson (2016), are not dependent on any assumption that the policy in question is the only way to offset pre-existing tax distortions, or even

necessarily the best way. Rather, the point is that taxes can distort a wide variety of margins, from the intensive margin of labour supply to job search to education, and beyond,[2] and if we are considering policy on any of these dimensions, the distortionary impact of income taxation must be taken into account. Thus, solving the distortion of taxation on job search by reducing unemployment insurance does not solve the distortion on education, as they are two very different margins of choice; reducing unemployment benefits generally does not restore the education choice to the efficient level. The essential assumption is simply that the distortionary marginal tax is necessary, that lump-sum taxes are either not possible or not desirable; the main model in the current paper simply imposes this assumption, but this appendix demonstrates that the same general results hold when it arises endogenously from inequality.

F.1   Model Setup

The model features a population of ex-ante identical individuals who live for three periods. In the first, the representative individual chooses the level of education they wish to obtain, denoted by e, and they also choose a level of borrowing d_1 to provide for consumption, as I assume that individuals begin and end life with zero wealth. They receive consumption utility $U(c_e)$, where $c_e = d_1 + b_e e$ is total consumption, and where b_e is the government subsidy to education attainment; for simplicity, I assume that the only cost of education is a convex loss of utility due to effort, $h_e(e)$.

---

[2]An anonymous referee rightly points out that "there are countless ways in which to change the distortions created by public finance needs", and that fiscal externalities could thus be an argument for changing other policies, such as investment subsidies for physical capital. If other margins are impacted by taxation, corrective policies on those margins could well be optimal.

In the second period, the representative individual starts out unemployed and searches for a job; they choose their search intensity s subject to a convex effort cost $h_s(s)$, and spend the first $(1-s)$ percent of the period unemployed and the remaining s employed at a wage of 1. Individuals also choose a level of borrowing d_2 when re-employed; I assume that no borrowing is allowed during unemployment, so that consumption during that time is $c_u = m + b_u$, where m is exogenous home production and b_u is an unemployment benefit. Consumption when re-employed is $c_j = 1 - \tau + b_w + \frac{d_2}{s}$, where b_w is a lump-sum transfer to employed individuals and τ is the marginal tax rate on employment income; borrowing d_2 is divided by s because the borrowed amount is only needed for a fraction s of the period, generating an instantaneous consumption stream of $\frac{d_2}{s}$. Education does not affect productivity in this period; this is a simplifying assumption to prevent strong interactions between education and job search, and corresponds to the idea that the first job is a training job at a lower productivity. Therefore, consumption utility in period 2 is $(1-s)U(c_u) + sU(c_j)$.

Finally, in period 3, the representative individual chooses a labour supply l subject to a convex effort cost $h_l(l)$, and works at a stochastic wage that is realized after the labour supply choice: each with probability 0.5, the wage can be $w_L = (1-\theta)(1+e)$ or $w_H = (1+\theta)(1+e)$. Expected utility from consumption is $0.5U(c_L) + 0.5U(c_H)$, where $c_L = (1-\tau)w_L l + b_w - d_1 - d_2$ and $c_H = (1-\tau)w_H l + b_w - d_1 - d_2$.

Putting everything together, and assuming zero interest and discount rates, overall expected utility given by:

$$V = U(c_e) - h_e(e) + (1-s)U(c_u) + sU(c_j) - h_s(s) + 0.5U(c_L) + 0.5U(c_H) - h_l(l).$$

F.2 First-Order Conditions & the Effect of Taxation

To understand the workings of this model, consider the first-order conditions for the representative individual given a policy vector $\{b_e, b_u, b_w, \tau\}$:

$$\frac{\partial V}{\partial d_1} = U'(c_e) - 0.5u'(c_1) - 0.5u'(c_2) = 0$$

$$\frac{\partial V}{\partial d_2} = U'(c_j) - 0.5u'(c_1) - 0.5u'(c_2) = 0$$

$$\frac{\partial V}{\partial e} = U'(c_e)b_e - h'_e(e) + 0.5(1-\tau)l\left[u'(c_1)(1-\theta) + u'(c_2)(1+\theta)\right] = 0$$

$$\frac{\partial V}{\partial s} = -U(c_u) + U(c_j) - U'(c_j)\frac{d_2}{s} - h'_s(s) = 0$$

$$\frac{\partial V}{\partial l} = 0.5(1-\tau)(1+e)\left[u'(c_1)(1-\theta) + 0.5u'(c_2)(1+\theta)\right] - h'_l(l) = 0$$

The first two first-order conditions, for d_1 and d_2, ensure consumption smoothing between the first period, the time spent working in the second period, and the third period (in an average marginal utility sense). Thus, the only remaining opportunities for consumption smoothing are in unemployment, and between high- and low-income third-period workers. The rest of the first-order conditions are a bit more complicated, but we can easily evaluate the effect of an increased tax rate τ on e, s and l:

- The substitution effect of a higher τ on education is clearly negative, as $\frac{\partial V}{\partial e}$ is decreasing in τ; of course, there is also an income effect if c_1 and c_2 are affected, so it is not clear that a higher tax rate will necessarily decrease e, but we know that an income effect is not distortionary. Therefore, we can conclude that taxation causes e to be inefficiently low.

- While τ does not appear directly in the first-order condition for s, it does appear in the definition of c_j, and from the perspective of job search, a reduction in $U(c_j)$ from an increase in τ is a marginal substitution effect,

changing the relative returns attached to unemployment and employment. Therefore, higher taxes reduce $\frac{\partial V}{\partial s}$, leading to an inefficiently low s (unless d_2 is very large, so that $\frac{\partial V}{\partial s}$ is actually increasing in s, which cannot be true in equilibrium unless s is at the upper bound of 1).

- Finally, the substitution effect of a higher τ reduces $\frac{\partial V}{\partial l}$, making l inefficiently low.

From this analysis, it is clear that higher marginal tax rates not only distort labour supply l downwards, they also lead to inefficiently low values for both e and s. Furthermore, those distortions are not dependent on each other - the tax rate impacts both margins independently.[3]

To consider the welfare implications of policy, I also need to present the government budget constraint:

$$\tau(s + (1 + e)l) = b_e e + (1 - s)b_u + (1 + s)b_w + G$$

where G is a revenue requirement beyond UI and education. However, a sufficient statistic welfare analysis is not as effective in this case, with three different margins and four policy instruments. I need to simulate the structural model, and the next subsection presents the results.

F.3   Optimal Policy Simulations

This model is meant to be illustrative, rather than a precise model of the real world; in particular, the assumption of three periods of equal length is clearly

---

[3]I could, of course, alter the model to introduce more interactions between choices on the various margins, but it would still be true that a higher tax rate makes both s and e too low because the returns to each have declined. For example, if the income from work in the second period depended positively on education, then there would be a positive feedback in which higher education raises the return to job search, and more intense job search raises the time spent working and thus the return to education. This would worsen the overall impact of income tax distortions, but solving one margin still wouldn't solve the other.

not realistic. Therefore, I do not attempt a calibration of the model to real-world empirical quantities. Instead, I use a simple parameterization described below.

To begin with, I assume that utility from consumption is logarithmic. I assume home production of $m = 0.2$, and an earnings distribution parameter of $\theta = 0.1$. Meanwhile, the effort disutility functions all take the form $h_x(x) = \sigma_x \left(\frac{1-(1-x)^{1-\kappa_x}}{1-\kappa_x} - x \right)$; this function takes a value of zero at $x = 0$, and has a slope that rises from zero and goes to infinity as x approaches 1 (as long as $\kappa_x > 0$), thus constraining e, s and l to be between zero and one. The parameters are: $\{\sigma_e, \kappa_e\} = \{6, 0.2\}$; $\{\sigma_u, \kappa_u\} = \{1, 0.5\}$; and $\{\sigma_l, \kappa_l\} = \{0.5, 0.5\}$.

Given a value of G, I can then solve for the optimal combined policy $\{b_e, b_u, b_w, \tau\}$ subject to the government budget constraint. I calculate the optimum for values of G at intervals of 0.05 from 0 to 0.8, and the results can be found in Figures F.1 and F.2; Figure F.1 presents the optimal transfers, while the optimal marginal tax rate τ is displayed in Figure F.2.
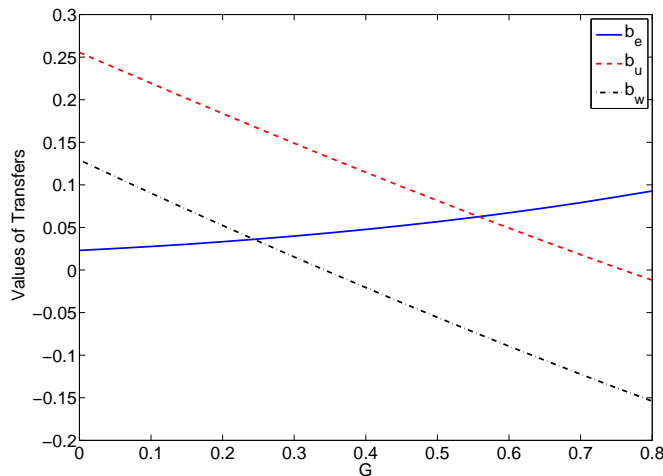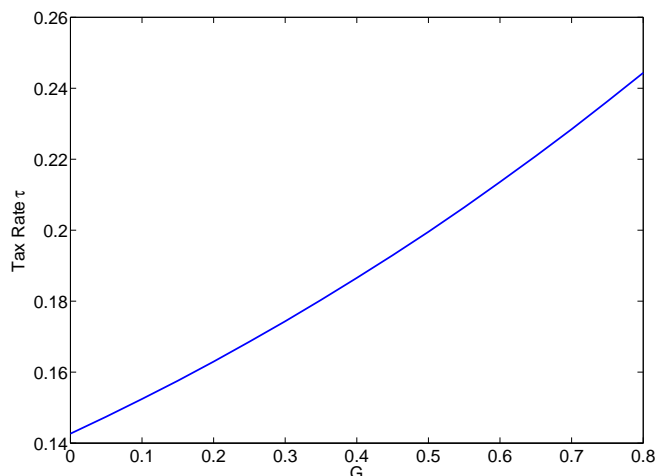
Figure F.1: Optimal Transfers

Figure F.2: Optimal Budget-Balancing Tax Rates

First of all, we can see that the marginal tax τ increases and the lump-sum transfer b_w decreases with G; whatever the mix between lump-sum taxes (or transfers) and marginal taxes that is optimal when $G = 0$, both are used to raise more revenue when G increases. While unsurprising, this supports my assumption that a positive G implies a larger marginal tax rate.

More interestingly, the optimal unemployment benefit declines rapidly with G, from over 0.25 at $G = 0$ to slightly below zero when $G = 0.8$, while the education subsidy increases substantially, from about 0.023 to 0.093. While the latter numbers may seem relatively small, the average employed income in the current model is about 1.28 at baseline, or about 1.34 at $G = 0.8$; in the main model of the paper, baseline tuition is about 15% of average income, and thus my results roughly correspond to a subsidy on tuition going from about 12% up to 46%. Given that tuition is not included in the current model, and thus does not need to be financed by the agent, this is quantitatively a large subsidy, particularly at the upper end.

Thus, while I again note that the model is not calibrated to the real world, the results do confirm the intuition discussed at the beginning of this appendix: we need to subsidize education more and unemployment less when G is higher. Solving the distortion on one margin doesn't solve the distortion on the other margin, because the cause of the distortion in the first place - the marginal income tax - is still in place. The implications for policy extend well beyond just unemployment insurance and education subsidies, but the analysis of the current paper and Lawson (2016) indicate that the impact is first-order on both of those margins.

References

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni L. Violante. 2013. "Education Policy and Intergenerational Transfers in Equilibrium." Working Paper.

Brown, Meta, John Karl Scholz, and Ananth Seshadri. 2012. "A New Test of Borrowing Constraints for Education." *Review of Economic Studies*, 79(2): 511–538.

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

Lawson, Nicholas. 2015. "Social Program Substitution and Optimal Policy." *Labour Economics*, 37: 13–27.

Lawson, Nicholas. 2016. "Fiscal Externalities and Optimal Unemployment Insurance." Unpublished Paper.

Trostel, Philip A. 2010. "The Fiscal Impacts of College Attainment." *Research in Higher Education*, 51(3): 220–247.