This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.
Abstract

This thesis explores inequality and its effects on the macroeconomy. Each chapter is concerned with heterogeneity between households or firms in some dimension: the first simply explores whether a standard model can capture the degree of inequality observed in the data; the second and third tackle the implications of such inequality for some aspect of policy. In all three papers I adopt the approach now common in the literature of viewing inequality as arising from idiosyncratic shocks and incomplete markets. In this lay summary I briefly discuss each chapter.

Matching the Wealth Distribution with Income Inequality & Risk. I show that a standard incomplete markets model with labour income risk alone, when modelled accurately using a new method, can broadly match the distribution of wealth observed in the US and UK, whereas adding capital income risk does little to concentrate wealth further; both findings are in contrast to the prior literature. Moreover, I show that increased labour income inequality and risk can account for around a third of the rise in the concentration of wealth among the top 1% in the US since the 1970s, and about 0.5pp of the decline in real interest rates.

Should I Stay (in School) or Should I Go (to Work). How should governments calibrate the desire for redistribution via progressive taxation against the need to incentivise the
accumulation of human capital through schooling? I explore this question in a heterogeneous agent model featuring stochastic human capital accumulation, endogeneous labour supply and an education choice modelled as a stopping time problem, where agents choose an optimal number of years to study before starting work. The latter feature addresses a shortcoming of much of the literature that typically models education decisions as a time allocation problem – which overstates the ability of older workers to insure themselves against obsolescence of their human capital and hence understates the welfare benefits of public insurance, for example through progressive taxation – or as a stylised problem with no resource or opportunity costs, which understates the bite of financial frictions, and hence the potential welfare gains from subsidising higher education. The social welfare-maximising policy features generous subsidies for education and highly progressive labour taxes. This result is robust to myriad extensions, including weakening the extent of financial frictions by allowing students to borrow.

**EM Forever Blowing Bubbles: Global Imbalances & the Limits of Fiscal Space.**

What are the limits on how much a government can borrow when the real interest rate on public debt is below the growth rate of the economy? I explore this question in a framework where household entrepreneurs make risky investments under incomplete markets. Such risks induce them to engage in precautionary saving, which lowers equilibrium interest rates and hence expands the fiscal space available to governments. I expand on the prior literature by introducing emerging market (EM) economies – where entrepreneurs face even greater risk and hence engage in more aggressive precautionary saving – as well as limits to the private supply of safe assets. Both elements reduce equilibrium real interest rates in the developed world and further inflate the bubble in public debt, affording developed market (DM) governments even greater fiscal space. I quantify the extent of fiscal space available to governments: how much debt they can issue without having to run surpluses, or how large a deficit they can sustainably run. But I also show that such borrowing carries risks, and governments must design fiscal rules to ensure their debt is both stable in the short run and sustainable in the long run, and that moreover that such policies are in general not Pareto-improving.
Matching the Wealth Distribution with

Income Inequality & Risk

Lee Tyrrell-Hendry
University of Edinburgh

September 2022
(Click here for latest version)

Abstract

In this paper I show that in contrast to prior research a standard incomplete markets model can broadly match the distribution of wealth observed in the data, including coming close to capturing the extreme concentration at the top, if one accurately models the inequality and risk associated with households’ earnings. Moreover, I show that adding capital income risk does not lead to materially greater concentration at the top, again in contrast to prior research. I lastly show that changes in income inequality and risk since the 1970s in the US can explain perhaps a third of the increase in the concentration of wealth at the top since then.
1 Introduction

Prior research has argued that the standard incomplete markets model cannot match the observed distribution of wealth through income endowment shocks alone. In this short paper I show that in fact it can do so, provided one accurately captures the degree of income inequality and risk observed in the data, which I do using a simple new method using survey panel data. With this approach I am able to very closely match the distribution of wealth for the bottom 95% of households, and also come close to matching the distribution among the wealthiest households as well, at least in the UK, where more fine-grained modelling of the income distribution is possible with survey data.

The main problem with standard income processes – and why this approach works in contrast – is that an agent’s precautionary saving motive is rather weak if their income mean-reverts rather quickly, and in any case cannot reach extremes – as is the case AR1 income processes with normally-distributed shocks – because of standard permanent-income logic. However, the actual risk and extremity observed in labour income data is such that high productivity workers earn extraordinary sums for short periods of time, with considerable downside risk, motivating intense precautionary saving.

Prior research has matched the wealth distribution primarily by incorporating at least one of the following two features:

1. Risky capital income in place of or on top of risky labour/endowment income. However, I also show that adding capital income risk on top of the labour income process described above risk does not materially concentrate wealth further over a baseline model with a realistic labour income process.

2. Exotic and particular modelling structures that strengthen incentives to save for high-income agents and/or weaken them for low-income households. I describe a variety of such approaches below in more detail, but suffice it to say that the simplified approach of this paper allows one to come near to matching the desired features of the data.
without somewhat opaque modelling structures that may strain our economic intuition as well as the limits of our computational power.

Lastly, I use long-running US income panel data since the 1970s to show that changes in the apparent income risk faced by households can explain perhaps a third of the increase in the concentration of wealth at the top since then.

Huggett (1993) and Aiyagari (1994) were among the first to quantitatively explore heterogeneous agent models. But those papers – which were not focussed on explaining the wealth distribution per se – used an income process with few states, and in the latter case assuming normally-distributed shocks, with the variance estimated from the PSID, and consequently could not match well the observed features of the distributions of either income or wealth, particularly their heavy Pareto tails and hence the high concentration among the top few percent, as discussed explicitly in Castaneda et al. (1998) and Stachurski & Toda (2019).

Macroeconomists have since taken a number of different approaches to try to explain this observed concentration of wealth. What all successful approaches have in common is that they give some households an extreme desire to save and others to dissave. This may be achieved through manipulating the utility from saving with heterogeneous and/or idiosyncratically-risky discount factors, like Krusell & Smith (1998) and Carroll et al. (2017), and originally postulated by Ramsey (1928) in a casual final comment of his landmark paper.

Alternatively, one can manipulate the returns to saving, like Benhabib et al. (2011) & (2016), and Khieu & Walde (2018). The latter paper is notable in that despite its partial equilibrium setting and simple employment-unemployment labour income process, heterogeneous returns are sufficient to match the wealth distribution, but incorporating risky capital income requires counterfactual “superstar” states. Benhabib et al. (2017) and Stachurski & Toda (2019) furthermore show that the standard Aiyagari (1994) model (with infinitely-lived households) cannot generate a Pareto-distributed wealth distribution with heavier tails than that of the income process, which is what we see in the data. I do not argue with this result, but merely claim that even without a heavy, Pareto-tailed wealth
distribution, one can still match many features of the data with the standard model, which may be sufficient for certain kinds of policy analysis. Moreover, I show that reasonable calibrations of the process governing capital income risk do not materially concentrate wealth over my baseline approach with labour income risk alone.

A final approach is to give certain households access to a superior, but somehow restricted saving technology. One economic story usually offered – which also fits the data on the occupations of the extremely wealthy – is entrepreneurship, with occupational choice and credit constraints limiting access to the technology among low-productivity, low-wealth households, e.g. Quadrini (2000) and Cagetti & De Nardi (2006). A variation on this is to incorporate illiquid assets, as in Kaplan & Violante (2014), Kaplan et al. (2018) and Luetticke (2021), where adjustment costs disincentivise low-income and low-wealth households from saving in the higher-return illiquid asset.

The alternate approach taken in this paper is to design the income process to generate the appropriate precautionary saving behaviour.

Others have taken this approach too, the simplest of which is simply to calibrate the income process to match the moments of the wealth distribution, as in Castaneda et al. (2003), who use a productivity process with an “awesome state” – reached by 0.04% of the workforce – more than 1000x that of the bottom state, representing the bottom 60% of workers. The downside is that they match the distribution of earnings only in a very coarse sense; their process features counter-factually extreme income inequality at the very top end that is an order of magnitude higher than that observed.

An alternative that relies on a much more sophisticated income process is due to Lise (2012), who endogenises income risk and inequality by incorporating on-the-job search. This generates precautionary savings behaviour through a “job ladder” that offers small incremental gains as workers climb the ladder, with a small risk of a large drop in income if they “fall off”, i.e. lose their jobs. Those at the top of the job ladder – the richest – have most to lose and therefore save the most. Combined with heterogeneous ability, this model can come
closer to matching the distribution of wealth. The downside here is the complicated modelling structure (with the standard search model baggage), which originally necessitated a partial equilibrium setting and hence limited the applications of the model, particularly on questions of policy. The simplified search process of Krusell et al. (2010), which integrates the standard DMP search model into the incomplete markets setting of Aiyagari (1994) achieves the same goal of endogenising labour income risk, but due to the simplicity of its labour income process (employment vs unemployment), it generates a trivial dispersion of wage income among employees, and hence although not explicitly stated the distribution of wealth matches the data poorly.

Other approaches model the life-cycle and human capital decisions of individuals to explain the increases in the level, variance and skewness of earnings observed among agents over their lifetimes, as in Huggett et al. (2006) & (2011). Although they do not explicitly try to match the concentrations of income and wealth at the top-end of the distribution, this literature provides a more compelling economic story to explain income fluctuations than simple exogenous processes. De Nardi (2004) highlighted the importance of bequests and intergenerational (non-)transmission of human capital in an overlapping generations context, essentially introducing substantial downside-risk by having high-productivity households lose part of their accumulated human capital from one generation to the next, motivating high bequests through otherwise standard precautionary saving/intertemporal substitution mechanisms, albeit typically with non-homothetic preferences for bequests. Without these factors, pure life-cycle savings behaviour is insufficient to generate the observed extreme concentrations of wealth, as Huggett (1996) showed.

Table 1 summarises certain key contributions to this vast literature in terms of how closely they were able to match the distribution of earnings and wealth in the US (regardless of whether that was their goal). This paper comes close to achieving this aim with a very simple approach and without many of the bells and whistles of those that come closer.
### Table 1: Matching the wealth distribution in the literature

<table>
<thead>
<tr>
<th>Source</th>
<th>Year</th>
<th>Earnings</th>
<th>Wealth</th>
</tr>
</thead>
<tbody>
<tr>
<td>US economy (2013)</td>
<td></td>
<td>0.67</td>
<td>0.85</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.9</td>
<td>-0.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>37.2</td>
<td>62.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18.8</td>
<td>35.5</td>
</tr>
<tr>
<td>Aiyagari (1994)</td>
<td></td>
<td>0.10</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32.5</td>
<td>17.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.5</td>
<td>13.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6.8</td>
<td>3.2</td>
</tr>
<tr>
<td>Huggett (1996)</td>
<td></td>
<td>0.42</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>9.8</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>22.6</td>
<td>33.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>13.6</td>
<td>11.1</td>
</tr>
<tr>
<td>Castaneda et al (1998)</td>
<td></td>
<td>0.30</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td></td>
<td>20.6</td>
<td>32.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10.1</td>
<td>7.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.0</td>
<td>1.7</td>
</tr>
<tr>
<td>Krusell &amp; Smith (1998)</td>
<td></td>
<td>0.82</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>55.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>24.0</td>
<td></td>
</tr>
<tr>
<td>Quadrini (2000)</td>
<td></td>
<td>0.45</td>
<td>0.74</td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt; 36.0</td>
<td>&lt; 16.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>18.2</td>
<td>45.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>7.9</td>
<td>24.9</td>
</tr>
<tr>
<td>Castaneda et al (2003)</td>
<td></td>
<td>0.63</td>
<td>0.79</td>
</tr>
<tr>
<td></td>
<td></td>
<td>3.7</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>32.6</td>
<td>48.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>14.9</td>
<td>29.9</td>
</tr>
<tr>
<td>Cagetti &amp; De Nardi (2006)</td>
<td></td>
<td>0.80</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>&lt; 6.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>60.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>31.0</td>
<td></td>
</tr>
<tr>
<td>Benhabib et al (2011)</td>
<td></td>
<td>0.69</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.0</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>52.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>34.1</td>
<td></td>
</tr>
<tr>
<td>Lise (2012)</td>
<td></td>
<td>0.62</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>36.3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>19.3</td>
<td></td>
</tr>
<tr>
<td>This paper</td>
<td></td>
<td>0.87</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1.2</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>56.6</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>20.3</td>
<td></td>
</tr>
</tbody>
</table>

Source: 2013 data on US earnings and wealth distribution taken from Kuhn & Rios-Rull (2016)

## 2 Baseline model

The model is that of Aiyagari (1994) with idiosyncratic but no aggregate risk, recast in continuous time, à la Achdou et al. (2022), but with idiosyncratic income shocks modelled as a discrete, multi-state Poisson process, to allow jumps and hence to match observed characteristics of income shocks. There are $J$ idiosyncratic income states and hence $J(J - 1)$ Poisson shocks; a jump from each state to every other state.

**Estimation of income process.** Since the model is otherwise standard, I first describe the novel process I use to estimate the income process.
One problem with the typical income processes used in standard incomplete markets models is that they either feature normally-distributed shocks, or only a few states, that capture reasonably well the body of the income distribution, but not the tails, and which do not allow for the skewness and leptokurtosis of income shocks, as shown in Guvenen et al. (2021). Consequently, the income share of the top 1% is usually too low. Instead, I explicitly model the income of high earners, including the top 1%, and where data allows, the top 0.1%. I show this is able to accurately capture the right tail of the income distribution. Moreover, I allow for (sometimes large) jumps in income, rather than a smooth process, which can capture the skewness and leptokurtosis. Incorporating both these features is crucial to generating the additional concentration of wealth seen in the data.

To model the income process I use survey data to separate households into different bins based on their labour income in each year of the survey. Each bin represents a fractile of the income distribution. I choose uneven fractiles, with a coarser gradation of productivity at low income levels, but explicitly incorporating a state for the top 1% of earners and (where data allows) the top 0.1%. I estimate the mean income for each fractile (which will stand in for their productivity) and estimate the transition probabilities between fractiles directly from the survey data. I show the estimated annual discrete-time transition matrix in the Appendix. This contrasts with the more typical approach of using the Simulated Method of Moments to match a few moments (e.g. skewness and kurtosis of income changes) of artificial data to those observed in actual data. My approach is roughly analogous to the “Empirical Calibration” approach of Civale et al. (2017), although they apply their method to synthetic data generated by a discrete time AR(1) process in order to estimate an income process using a Simulated Method of Moments approach, whereas as I apply my method to actual data.

The result of this estimation is a transition matrix (shown in the Appendix, Figure 6), governing the probability of transitioning from each income fractile from one year to the next. I convert this transition matrix into its continuous time analogue by taking the matrix log. This results in some negative probabilities in off-diagonals, an issue known as the
embeddability problem, which occurs because the transition probabilities for certain fractiles estimated from the sample data are (close to) zero (Israel et al. 2001). I correct for this with the diagonal adjustment method of Inamura (2006): set the negative values to zero and adjust their diagonals so that each row sums to zero. I also show the continuous time transition matrix in the Appendix, Figure 7.

I use the PSID for the US and the Understanding Society survey (which I call by its former acronym, BHPS) for the UK. Since I do not model retirement, I use only data on labour income. The BHPS has since 2016 accurately captured the income shares of the top 1% and 0.1%, albeit with a small sample size for the latter, as shown in Table 3. However, the PSID under-reports both shares, and given the small sample it is not possible to capture the top 0.1% at all. Consequently, I recalibrate the estimated top 1% income level to match the top 1% share observed in the Survey of Consumer Finances (SCF) and administrative data. The SCF accurately captures the income shares of top earners, but usually does not track individuals over time, so it is not possible to calculate transition probabilities using this source. Although I can therefore match the income share of the top 1% using supplemental SCF data, there is nothing I can do to adjust the estimated transition probabilities for the top 1%, thus potentially introducing error if their true income risk is very different from that observed in the PSID. If for example one takes the view that the top 1% of earners in the PSID are actually the second 1% (for example, if the top 1% are missed from the survey entirely) then the observed transition probabilities may understate their true income risk, and hence the strength of their precautionary saving motive (UK data suggests the top 1% and in particular the top 0.1% face greater income risk than those just below them). To compare the model results with wealth concentration data, I use the Wealth & Asset Survey for the UK, supplemented with the results of Alvaredo et al. (2018) to capture the top end, and the SCF for the US.

Although the SCF usually does not track individuals over time, there was a brief panel for 2007-09, for which Kuhn & Rios-Rull (2016) report an almost-complete transition matrix.
I also report results for the baseline model using this transition matrix, although this data is also not ideal, for reasons I discuss below.

Recap of standard incomplete markets model. The household’s problem is to maximise intertemporal utility subject to their budget constraint (wealth grows by labour income, less taxes plus transfers, plus capital income less consumption), the exogenous process for income and an exogenous borrowing limit:

\[
V(a_0, z_j) = \max_{\{c_t\}_{t \in [0, \infty)}} \mathbb{E}_0 \left[ \int_0^\infty e^{-\rho t} u(c_t) dt \right]
\]

\[
\text{s.t. } \quad \dot{a}_t = ra_t + wz_j(1 - \tau) + T_j - c_t \\
\quad a_t \geq a = -\phi
\]

This can be re-written as the Hamilton-Jacobi-Bellman equation, the continuous time analogue of the Bellman equation:

\[
\rho V(a, z_j) = \max_c u(c) + \partial_a V(a, z_j)(ra + wz_j(1 - \tau) + T_j - c) + \sum_{k=1}^J \lambda_{jk} V(a, z_k)
\]

(1)

where \(\lambda_{jj}\) equals the negative of the sum of the other \(\lambda_{jk}\), such that each row of the transition matrix \(\Lambda\) sums to zero. This matrix governs the transitions between states and will be estimated directly from panel income data, as discussed below. The budget constraint includes state-dependent transfers, \(T_j\), a flat income tax, \(\tau\), and a borrowing rate (not shown) given by the lending rate, \(r\), plus a spread, \(\chi\). This HJB equation can be derived from a discrete-time Bellman equation, for a demonstration of which see Achdou et al. (2022).

The household’s optimality conditions yield policy functions for consumption, \(c(a, z)\), and saving, \(\dot{a}(a, z)\), which given the continuous time setting hold with equality everywhere in the interior of the state space, with the boundary condition \(\partial_a V(a, z_j) \geq u'(ra + wz_j(1 - \tau) + T_j)\)
holding at the borrowing constraint:

\[ c(a, z_j) = (u')^{-1} \left[ \partial_a V(a, z_j) \right] \]
\[ \dot{a}(a, z_j) = ra + wz_j(1 - \tau) + T_j - c(a, z_j) \]

The evolution of the joint distribution of income (productivity) and wealth is governed by the Kolmogorov Forward Equation below, and a stationary distribution, if one exists, is found by setting this equal to zero:

\[ \dot{g}(a, z_j) = -\partial_a [\dot{a}(a, z_j)g(a, z_j)] - \sum_{\substack{k=1 \\ k \neq j}}^J [\lambda_{jk}g(a, z_j) - \lambda_{kj}g(a, z_k)] = 0 \quad (2) \]

The first term in this expression captures the change in the density from households increasing or decreasing their net worth by saving. The second term (first part in square brackets) captures the decrease in the density at \((a, z_j)\) from households with productivity \(z_j\) being hit by an income shock at Poisson rate \(\lambda_{jk}\) and moving to productivity \(z_k\), and the third term (second part in square brackets) captures the increase in the density at \((a, z_j)\) from households with productivity \(z_k\) being hit with an income shock at rate \(\lambda_{kj}\) and moving to productivity \(z_j\).

I close the model in the standard fashion of Aiyagari (1994), with aggregate assets in equilibrium equal to the aggregate capital stock used in production, and with the standard firm optimality conditions under competitive factor markets, where I normalise \(z_{ave}\) to 1 and where \(e\) denotes the employment rate, meaning the measure of all except those in the lowest income state:

\[ K = \sum_{j=1}^J \int_a^\infty ag(a, z_j)da \quad r = \alpha z_{ave} \left( \frac{K}{e} \right)^{\alpha - 1} - \delta \quad w = (1 - \alpha) z_{ave} \left( \frac{K}{e} \right)^\alpha \]

The stationary equilibrium of the model is characterised by the value function of the household, \(V(a, z)\), derived from the HJB, and their resulting policy functions, \(c(a, z_j)\) and \(\dot{a}(a, z_j)\), which maximise their lifetime utility subject to prices, \(r\) and \(w\), which in turn clear
all markets given the joint density function \( g(a, z) \), determined from the KFE. Since the model is standard, I refer the reader to Achdou et al. (2022) for a discussion of the numerical approach I use to solve for the model equilibrium, that being finite difference methods.

**Matching the UK wealth distribution.** Starting with the UK, for which more refined survey data are available, the estimated productivity grid is highly skewed. I use 20 uneven quantiles to target income shares, including the top 0.1% and next 0.9%. The income state for the top 0.1% has productivity more than 60x that of the mean, and the next 0.9% 5x the mean. I have rebased the grid so average productivity is 1. I reserve the bottom state for households with zero or negative earnings, which is around 15% of the total. I interpret this state as non-employment, although this is not strictly the case as a small number of households have negative income because of self-employment or business losses.

The transition probabilities are also skewed, displaying significant downside risk for the top 1% and 0.1%, which will be crucial for generating the savings rates among wealthy individuals required to achieve a highly concentrated wealth distribution. The distribution of income growth observed in the UK data is also highly-non-normal, displaying negative skewness of -0.5 and high leptokurtosis of around 15.

I emphasise that I have estimated the income process directly from the data, with no adjustments to transition probabilities or income levels for the UK; moreover I have not calibrated the income process at all to match any features of the wealth distribution. Consequently, the fact that I am able to match closely the wealth distribution is truly a result of income inequality and income risk alone.

Aside from the income process, I calibrate the discount rate to 2% per quarter to match a capital/output ratio of around 3x. I set the depreciation rate to 1.7% per quarter to match the annual investment-to-output ratio of around 20% and set the capital share to 40%. I choose an unsecured borrowing limit of around one quarter of annual output, in line with Kaplan et al. (2018) and choose a spread of the borrowing over the lending rate of 1.65% per quarter to match the share of households with negative wealth of around 9%, slightly
higher than in the latest WAS data but close to other estimates in the literature (Crawford et al. 2016). I use a CRRA utility function with a coefficient of relative risk aversion of 1. I set a government transfer for each income quantile equal to that observed in the 2017 Understanding Society survey and set a flat tax levied on labour income of 20.9% to balance the government budget. There is no other government spending.

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter</th>
<th>Value</th>
<th>Calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td>ρ</td>
<td>Discount rate</td>
<td>2%</td>
<td>~3x capital/output ratio</td>
</tr>
<tr>
<td>γ</td>
<td>Coefficient of risk aversion</td>
<td>1</td>
<td>Standard</td>
</tr>
<tr>
<td>δ</td>
<td>Depreciation rate</td>
<td>1.7%</td>
<td>~20% investment/output ratio</td>
</tr>
<tr>
<td>α</td>
<td>Capital share of income</td>
<td>40%</td>
<td>~40% capital income share</td>
</tr>
<tr>
<td>φ</td>
<td>Borrowing constraint</td>
<td>4.4</td>
<td>25% annual output</td>
</tr>
<tr>
<td>χ</td>
<td>Borrowing spread</td>
<td>1.65%</td>
<td>~9% share with negative wealth</td>
</tr>
<tr>
<td>τ</td>
<td>Income tax rate</td>
<td>20.9%</td>
<td>Balanced government budget</td>
</tr>
</tbody>
</table>

The model matches the earnings, income and wealth shares across most of the distribution almost exactly, and comes close to matching the top 1% share of wealth of 20%, as shown in table 3. The top 0.1% hold around 3% of wealth in the model; there is limited data available from surveys, but Alvaredo et al. 2018 estimate this share at around 7.5-8%, which is also consistent with a power law ratio of around 2.5:1 between the top 10:1% and the top 1:0.1% shares. The model appears to understate this share somewhat. Potentially this could be resolved by including a state for the top 0.01%, data permitting, but this is speculation.

It is also possible that the small sample size for the top 0.1% in the BHPS survey data led to error in the estimation of the transition probabilities, specifically underestimating the probability of negative shocks and hence underestimating the true strength of the precautionary saving motive. Nevertheless, the accuracy with which the model is able to match the bulk of the income and wealth distribution – and specifically the top wealth shares – compared to previous studies does highlight the importance of accurately capturing the inequality and risk of the income process, specifically the negative skewness and leptokurtosis of income shocks.
Table 3: UK labour earnings, total income and wealth shares vs model

<table>
<thead>
<tr>
<th></th>
<th>Gini</th>
<th>&lt; £0</th>
<th>50%</th>
<th>50-90%</th>
<th>10%</th>
<th>1%</th>
<th>0.1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings (model)</td>
<td>48.4</td>
<td>0</td>
<td>17</td>
<td>50.5</td>
<td>32.5</td>
<td>10.9</td>
<td>6</td>
</tr>
<tr>
<td>Earnings (BHPS data)</td>
<td>48.4</td>
<td>0.1</td>
<td>17.2</td>
<td>50.5</td>
<td>32.3</td>
<td>10.8</td>
<td>5.9</td>
</tr>
<tr>
<td>Income (model)</td>
<td>35.7</td>
<td>0</td>
<td>27.2</td>
<td>44.3</td>
<td>28.5</td>
<td>8.1</td>
<td>3.9</td>
</tr>
<tr>
<td>Income (BHPS data)</td>
<td>36.9</td>
<td>0</td>
<td>25.5</td>
<td>45.5</td>
<td>29</td>
<td>10</td>
<td>5.4</td>
</tr>
<tr>
<td>Wealth (model)</td>
<td>71.5</td>
<td>9.2</td>
<td>4.6</td>
<td>40</td>
<td>55.3</td>
<td>15.5</td>
<td>2.7</td>
</tr>
<tr>
<td>Wealth (WAS data)</td>
<td>67.5</td>
<td>1.6</td>
<td>6.5</td>
<td>41.6</td>
<td>51.9</td>
<td>19.9</td>
<td>7.5</td>
</tr>
</tbody>
</table>

Source: Bottom 90% wealth shares taken from 2012 WAS; top 10% taken from Alvaredo et al. 2018

The model also accurately matches the shares of total income earned by the prime age population, albeit slightly undershooting for those at the top end, because the model understates their shares of wealth and hence capital income. The Lorenz curves for earnings (labour income), total income and wealth are also shown below, for both the data and the model.

Figure 1: Baseline model Lorenz curves vs UK data

Comparison to other standard income processes. I now compare the estimated Poisson process above with two other income processes commonly used in the literature: one with mean-reverting and one with permanent (log-)normally-distributed shocks. In continuous
time these are referred to respectively as an Ornstein-Uhlenbeck (OU) process and a geometric Brownian motion (GBM). The stochastic processes for productivity are for an OU process (in logs):

$$dz = (-\theta \ln(z) + \frac{1}{2} \sigma^2) z dt + \sigma z dW,$$

and for a GBM:

$$dz = \mu z dt + \sigma z dW.$$

I calibrate the mean-reverting OU process with $\theta = 0.3$ to match the persistence of earnings estimated in the BHPS data and a variance of $\ln(z)$ of 0.5 to most closely match the Lorenz curve for earnings (this is slightly lower than the measured variance of log earnings of around 0.7), otherwise the calibration is the same, although I drop the taxes and transfers and of course the notion that some agents are unemployed; this has a negligible impact on the distributional moments. I do not recapitulate the household’s HJB equation, or the KFE governing the evolution of the joint distribution of wealth and income. Suffice it to say they are standard.

I calibrate the GBM process with $\mu = 0.01$, reflecting average real income growth of around 1%/year (a little too low), and $\sigma = 0.12$. The variance of log income grows over time at rate $\sigma^2 t$, giving a conditional variance of log labour income among 65 year-olds (45 in model years) of around 0.65, which is broadly in line with Deaton & Paxson (1994). Note also that in this case I use a perpetual youth framework, to ensure a stationary distribution of income and wealth exists; I assume agents die at a constant rate $d = 1/45$, for an expected lifetime of 45 years. I furthermore assume agents are reborn with no wealth and initial productivity of $z^* = 0.5$; living agents are assumed to buy competitive annuities that pay them $da$ throughout their life in exchange for giving up their assets to the insurer upon their death. In this case, as is well-known (e.g. see Gabaix (2009)), the marginal distribution of income has a double-Pareto distribution, with the tails given by $\zeta = \left(-\mu \pm \sqrt{\mu^2 + 2\sigma^2 d}\right)/\sigma^2$. My calibration gives $\zeta^+ \simeq 1.6$, which is roughly the observed Pareto right-tail of the income distribution in the UK.

I show below the wealth shares for various earnings, income and wealth fractiles compared to the baseline model, as well as the Lorenz curves compared to the actual data for the UK. The model with mean-reverting shocks performs poorly. Despite closely matching the
earnings distribution, the lower earnings risk among the rich means they do less precautionary saving and hence there is much less concentration of wealth, so the model falls far short of matching the observed distribution of wealth. The model with permanent shocks performs somewhat better. In particular, the share of wealth of the top 1% is close to the data, although these numbers are sensitive to the (somewhat arbitrary) calibration. Moreover, much of the concentration of wealth in this model arises because a handful of elderly households accrue enormous wealth over 100+ years, which is obviously counterfactual. A more realistic ageing and bequest process can address this issue, as in De Nardi (2004).

Another advantage of the income process in the Baseline model is that it can generate an average annual marginal propensity to consume on the order of 20%, more than double that in the model with either mean-reverting or permanent normally-distributed shocks, though still short of the 50% or so seen in the data.

<table>
<thead>
<tr>
<th></th>
<th>Gini</th>
<th>&lt; £0</th>
<th>50%</th>
<th>50-90%</th>
<th>10%</th>
<th>1%</th>
<th>0.1%</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Earnings</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>48.4</td>
<td>0</td>
<td>16.9</td>
<td>49.3</td>
<td>33.9</td>
<td>11.1</td>
<td>6.1</td>
</tr>
<tr>
<td>Mean-reverting</td>
<td>40.8</td>
<td>0</td>
<td>18.4</td>
<td>51.8</td>
<td>29.8</td>
<td>5.4</td>
<td>0.7</td>
</tr>
<tr>
<td>Permanent (analytical)</td>
<td>50.8</td>
<td>0</td>
<td>19.6</td>
<td>36.1</td>
<td>44.3</td>
<td>18.8</td>
<td>7.9</td>
</tr>
<tr>
<td>Permanent (discretised)</td>
<td>48.4</td>
<td>0</td>
<td>20.8</td>
<td>39.5</td>
<td>39.7</td>
<td>8.9</td>
<td>1</td>
</tr>
<tr>
<td><strong>Income</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>35.7</td>
<td>0</td>
<td>26.3</td>
<td>45.2</td>
<td>28.5</td>
<td>8.3</td>
<td>4.2</td>
</tr>
<tr>
<td>Mean-reverting</td>
<td>38.4</td>
<td>0</td>
<td>23.5</td>
<td>49.5</td>
<td>27</td>
<td>4.5</td>
<td>0.6</td>
</tr>
<tr>
<td>Permanent</td>
<td>50.9</td>
<td>0</td>
<td>19.6</td>
<td>38</td>
<td>42.4</td>
<td>11.5</td>
<td>1.6</td>
</tr>
<tr>
<td><strong>Wealth</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>71.5</td>
<td>9.2</td>
<td>4.6</td>
<td>40.1</td>
<td>55.3</td>
<td>15.5</td>
<td>2.7</td>
</tr>
<tr>
<td>Mean-reverting</td>
<td>54.5</td>
<td>10.9</td>
<td>11.4</td>
<td>54.6</td>
<td>34</td>
<td>5.3</td>
<td>0.7</td>
</tr>
<tr>
<td>Permanent</td>
<td>65.4</td>
<td>0</td>
<td>9.7</td>
<td>36.1</td>
<td>54.2</td>
<td>20.4</td>
<td>4.9</td>
</tr>
</tbody>
</table>
Comparison to a model with capital & labour income risk. Prior literature, for example Benhabib et al. (2016), have highlighted the role of capital income risk in generating a Pareto tail in the wealth distribution. They show analytically in a model without labour income that capital income risk alone generates a Pareto distribution for wealth. In combination with risk from uncertain lifetimes in a perpetual youth model, capital income risk generates a two-tailed “double Pareto” distribution for wealth. Although the precise model specification (no labour income) seems peculiar, the results still apply to more general models with labour income and labour income risk for the right tail of the wealth distribution, since in the limit as wealth grows, labour income becomes small by comparison.

I argue however that the role of capital income risk may have been somewhat overstated in the literature. For example, Achdou et al. (2022) show in an Appendix that a model with capital income risk and a simple two-state labour income process can generate a wealth distribution in line with the data. However, I will show that in a quantitative model with the realistic labour income process estimated above, adding capital income risk does not result in a materially greater concentration wealth over the baseline model, and arguably fits the bulk
Matching the Wealth Distribution

Lee Tyrrell-Hendry

of the distribution worse.

Households can now invest in two assets; a riskless bond with return \( r \), and a risky asset—held in proportion \( \omega \), chosen optimally—the return of which follows a Brownian motion:

\[ Rdt + \sigma dW. \]

Consequently, total assets now evolve stochastically:

\[ da = (ra + (R - r)\omega a - c)dt + \sigma \omega adW. \]

With such permanent wealth shocks, it is necessary again to include some feature to induce a stationary distribution of wealth; I use the same approach as above: a perpetual youth framework where agents die at a constant rate, \( d \), and are reborn with no wealth. The stationary distribution of wealth again follows a Pareto distribution in the right tail:

\[ g = ca^{-\zeta - 1}, \]

with the heaviness of the right tail, \( \zeta \), given by:

\[ \zeta = \frac{\Gamma \pm \sqrt{\Gamma^2 + 2\omega^2\sigma^2d}}{\omega^2\sigma^2}, \]

where

\[ \Gamma = \frac{r - \rho - d}{\gamma} + \frac{\gamma}{2} \left( \frac{R - r}{\gamma \sigma} \right)^2 \]

and \( \omega = \frac{R - r}{\gamma \sigma^2}. \)

\( \zeta \approx 1.3 \) roughly matches the US data, and \( \zeta \approx 1.5 \) the UK.

For completeness, I show the HJB and KF equations (suppressing notation):

\[ \rho V_j = \max_{c,\omega} \left( u(c) + \partial_a V_j(wz_j(1 - \tau) + T_j + ra + (R - r)\omega a - c) + \frac{1}{2} \partial_{aa} V_j \sigma^2 \omega^2 a^2 + \sum_{k=1}^{J} \lambda_{jk} V_k \right) \]

\[ \dot{g}_j = -\partial_a [a(a, z_j)g_j] + \frac{1}{2} \partial_{aa} [\sigma^2 \omega^2 a^2 g_j] - \sum_{k=1}^{J} [\lambda_{jk} g_j - \lambda_{kj} g_k] = 0 \]

The market clearing conditions are also slightly different for this model. Here I interpret the risky asset as capital and the risk-free asset as a bond in positive net supply of around 80% of GDP, i.e. the level of government debt in the UK.

\[ K = \sum_{j=1}^{J} \int_{\frac{a}{2}}^{\pi} \omega a g(a, z_j)da \]

\[ B = \sum_{j=1}^{J} \int_{\frac{a}{2}}^{\pi} (1 - \omega) a g(a, z_j)da \]

The resulting income and wealth shares are close to the data and wealth is slightly more concentrated than in the baseline model, although the low/middle part of the distribution fits the data slightly worse.
Matching the Wealth Distribution

Lee Tyrrell-Hendry

Table 5: UK earnings, income and wealth shares vs labour and capital income risk model

<table>
<thead>
<tr>
<th></th>
<th>Gini</th>
<th>&lt; £0</th>
<th>50%</th>
<th>50-90%</th>
<th>10%</th>
<th>1%</th>
<th>0.1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings (model)</td>
<td>45.9</td>
<td>0</td>
<td>16.9</td>
<td>50.7</td>
<td>32.5</td>
<td>10.9</td>
<td>6</td>
</tr>
<tr>
<td>Earnings (BHPS data)</td>
<td>48.4</td>
<td>0.1</td>
<td>17.2</td>
<td>50.5</td>
<td>32.3</td>
<td>10.8</td>
<td>5.9</td>
</tr>
<tr>
<td>Income (model)</td>
<td>48.5</td>
<td>0</td>
<td>16.8</td>
<td>50.2</td>
<td>33</td>
<td>10.5</td>
<td>5.7</td>
</tr>
<tr>
<td>Income (BHPS data)</td>
<td>36.9</td>
<td>0</td>
<td>25.5</td>
<td>45.5</td>
<td>29</td>
<td>10</td>
<td>5.4</td>
</tr>
<tr>
<td>Wealth (model)</td>
<td>84.8</td>
<td>5.8</td>
<td>0.4</td>
<td>26.3</td>
<td>73.3</td>
<td>31.3</td>
<td>6.8</td>
</tr>
<tr>
<td>Wealth (WAS data)</td>
<td>67.5</td>
<td>1.6</td>
<td>6.5</td>
<td>41.6</td>
<td>51.9</td>
<td>19.9</td>
<td>7.5</td>
</tr>
</tbody>
</table>

Source: Bottom 90% wealth shares taken from 2012 WAS; top 10% taken from Alvaredo et al. (2018)

Figure 3: Lorenz curves of model with labour and capital income risk vs UK data

Matching the US wealth distribution. Moving on to the US, the estimated productivity grid is also highly skewed, although not to the extent of Castaneda et al. (2003): productivity in the top state is more than 20x the mean. Due to the lack of good panel data, the top income state is for the top 1%, not the 0.1%. The estimated productivity grid and semi-annual transition matrix for the US is shown again in the Appendix, Figure 8.

I calibrate the semi-annual discount rate to 5.8% to match the US capital/output ratio of around 3x. I set the depreciation rate to 3% to match the investment-to-output ratio of around 18% and set the capital share to 40%. I choose an unsecured borrowing limit of
around one quarter of average annual income, in line with Kaplan et al. (2018) and choose a spread of the borrowing over the lending rate of 6% (semi-annualised) to match the share of households with negative wealth of around 10%. I use a CRRA utility function with a coefficient of relative risk aversion of 1. I set a government transfer for each income fractile equal to that observed in the 2014 PSID data and set a flat tax levied on labour income of 6.1% to balance the government budget. There is no other government spending.

Table 6: US model calibration

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter</th>
<th>Value</th>
<th>Calibration</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>Discount rate</td>
<td>5.8%</td>
<td>$\sim$3x capital/output ratio</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Coefficient of risk aversion</td>
<td>1</td>
<td>Standard</td>
</tr>
<tr>
<td>$\delta$</td>
<td>Depreciation rate</td>
<td>3%</td>
<td>$\sim$18% investment/output ratio</td>
</tr>
<tr>
<td>$\alpha$</td>
<td>Capital share of income</td>
<td>40%</td>
<td>$\sim$40% capital income share</td>
</tr>
<tr>
<td>$\phi$</td>
<td>Borrowing constraint</td>
<td>1.3</td>
<td>25% annual output</td>
</tr>
<tr>
<td>$\chi$</td>
<td>Borrowing spread</td>
<td>6%</td>
<td>$\sim$10% share with negative wealth</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Income tax rate</td>
<td>6.1%</td>
<td>Balanced government budget</td>
</tr>
</tbody>
</table>

The model matches the wealth shares of the bottom 99% of households somewhat accurately, and can generate a share of wealth for the top 1% of 18%, well below their true share of more than 30%, but still far greater than previous studies have managed to reproduce with a normally-distributed income process. It is very likely that the shortfall among the top percentiles is due to the poor data on their transition probabilities. Although we can match their income shares using additional data from the SCF, we cannot accurately capture their income risk since the PSID does not capture these individuals sufficiently well.

Comparing the US transition matrix to that for the UK, the top 1% appear to face as little as half the risk of falling out of the top bracket, likely because the top 1% in the PSID dataset are not the true 1%. Although it is also conceivable that for whatever reason the US top 1% simply do face less income risk, and save for reasons other than precautionary, for example those discussed in the literature review.

To add some credence to this, below I also include the results of a second model, where the transition matrix for earnings has been interpolated from that shown in Kuhn & Rios-Rull.
Matching the Wealth Distribution

Lee Tyrrell-Hendry

(2016), calculated from 2007-09 SCF data. As discussed above, the SCF captures the upper tail of the earnings distribution much better than does the PSID, so I do not have to adjust the productivity levels for the top 1%. As the results show, model 2 is better able to match the lower ~95% of the wealth distribution than model 1, and also gets slightly closer in matching the upper tail. For comparability, I use the same calibration for model 2 as model 1.

Unfortunately there is no data on the transfers received by each fractile, so I must drop the taxes and transfers from this model and give the lowest fractile (the bottom quintile) a nominal amount of income to ensure the model solves. This explains the large fraction with negative wealth. Moreover, this data is for the Great Recession period 2007-09, so the transition probabilities may not be representative of the typical idiosyncratic risk faced by households throughout the business cycle.

Table 7: US labour earnings, total income and wealth shares vs model

<table>
<thead>
<tr>
<th></th>
<th>Gini</th>
<th>&lt; £0</th>
<th>50%</th>
<th>50-90%</th>
<th>10%</th>
<th>1%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Earnings (model 1)</td>
<td>60.5</td>
<td>0</td>
<td>12.1</td>
<td>41.8</td>
<td>46.1</td>
<td>18.9</td>
</tr>
<tr>
<td>Earnings (model 2)</td>
<td>61.7</td>
<td>0</td>
<td>2.3¹</td>
<td>50.4¹</td>
<td>47.3</td>
<td>19.3</td>
</tr>
<tr>
<td>Earnings (SCF data)</td>
<td>67</td>
<td>&lt;1</td>
<td>2.9¹</td>
<td>46.7¹</td>
<td>49.6</td>
<td>18.8</td>
</tr>
<tr>
<td>Income (model 1)</td>
<td>55.6</td>
<td>0</td>
<td>13.9</td>
<td>45.3</td>
<td>40.9</td>
<td>15</td>
</tr>
<tr>
<td>Income (model 2)</td>
<td>62.8</td>
<td>0</td>
<td>9.3</td>
<td>42.8</td>
<td>47.9</td>
<td>15.3</td>
</tr>
<tr>
<td>Income (SCF data)</td>
<td>58</td>
<td>0</td>
<td>9.5¹</td>
<td>37.5¹</td>
<td>47</td>
<td>19.7</td>
</tr>
<tr>
<td>Wealth (model 1)</td>
<td>75</td>
<td>10.6</td>
<td>4.1</td>
<td>34.1</td>
<td>61.7</td>
<td>17.9</td>
</tr>
<tr>
<td>Wealth (model 2)</td>
<td>87.5</td>
<td>31.7</td>
<td>-1.2</td>
<td>25.5</td>
<td>75.7</td>
<td>20.3</td>
</tr>
<tr>
<td>Wealth (SCF data)</td>
<td>85</td>
<td>10</td>
<td>0</td>
<td>25</td>
<td>75</td>
<td>35.5</td>
</tr>
</tbody>
</table>

¹Bottom 40% and 40-90th percentile
Accounting for the increase in top shares of wealth since the 1970s. Income risk alone accounts for some of the additional share of wealth of the top 1% relative to income in the US compared to other countries, but it can also explain around a third of the increase in top 1% share since the 1970s. I show this by performing the same exercise using PSID/SCF income quantiles from 1975 and PSID transition probabilities between 1969 and 1975.

Two differences of note between the income process estimated from 1970s data vs 2010s data are the lower share of non-employed workers and the lower share of income earned by the top 1% in the earlier period. Also noteworthy is the fact that the 1% appeared to face greater risk of falling out of the top bracket in the 1970s compared to today. This would suggest less precautionary saving and hence a lower wealth share today, however this may just be due to errors in estimating transition probabilities due to poor data. Nevertheless, it does tentatively suggest other motives for saving than precautionary are important for explaining the observed high concentration of wealth relative to income. Other than the income process I use the same calibration as above, to isolate the effects of the change in the income process.

The income share of the top 1% was around 10% in 1975 (the PSID under-reports this
by half) and their wealth share around 22%. Today their wealth share is closer to 35%, an increase of 15pp. Matching the income share as above I can generate a share of wealth among the top 1% of around 10% for the 1970s calibration, compared to 15% for the 2010s calibration. Thus the model suggests that around a third (5pp) of the increase in the top 1% wealth share can be explained by changes in the income process, including a greater income share for the top 1%.

| Table 8: Model labour earnings, total income and wealth shares for 1975 calibration |
|---------------------------------|---------|---------|---------|---------|---------|
| Earnings (model)                | 47.5    | 0       | 18.3    | 53.4    | 28.3    | 8.1     |
| Income (model)                  | 44.5    | 0       | 19.6    | 49.4    | 31      | 8.3     |
| Wealth (model)                  | 57.9    | 3.7     | 12.4    | 44.5    | 43.1    | 10.6    |

I demonstrate this transition of inequality more clearly by introducing a one-time, unexpected permanent change in the income process in 1975. The top 1% share rises from 10% to 15%, with much of the transition occurring in the late 70s and 1980s, as in the data, although the increase in concentration trails off after this, in contrast with the US experience. The top 10% share rises from 40% to nearly 60%; the transition is slower than that for the top 1% and much larger than that observed in the data.

The more unequal income process also implies a fall in the real interest of nearly 0.5pp. The level of interest rates is far higher than observed in the data, since the precautionary savings motive results in a larger capital stock for a given interest rate, so to match the observed capital-output ratios one must raise the discount rate. This issue can be addressed by including a risky asset, as above, which gives the safe asset a liquidity premium, lowering the interest again. Nevertheless, this exercise does give a sense of the potential contribution of higher inequality towards secular stagnation and lower equilibrium interest rates, on top of, for example, the demographic, productivity and fiscal factors identified in Eggertsson et al. (2019) and Rachel & Summers (2019).

Incidentally, the fall in interest rates goes some way to offsetting the rise in inequality that
would have occurred under the transition to the new income process under partial equilibrium. The top 10% share would have risen to more than 70% and that of the 1% to almost 20%, had interest rates not fallen.

Figure 5: Increase in top wealth shares due to increased income inequality and risk

3 Conclusion

I have hopefully shown that labour income risk alone, properly measured, particularly for high earners, can help one match the observed wealth distribution in a standard incomplete markets model to a higher degree of accuracy than past literature suggests. In particular I used a new method of modelling labour income risk as a multi-state Poisson process by estimating carefully chosen income fractiles and transition probabilities between said fractiles directly from survey data. This approach in my opinion compares favourably to other approaches that use an income process with normally-distributed shocks, albeit perhaps not to more complicated models that introduce other motivations for saving by the wealthy.

I furthermore showed that adding capital income risk to a standard model with such an income process does not materially further concentrate wealth over the baseline model,
casting some doubt over the typical belief espoused in the literature that such investment risk is crucial to matching the wealth distribution. Moreover, the observed change in income inequality and risk alone can also explain a substantial part – perhaps a third – of the increase in wealth owned by the top 1% since the 1970s in the US, as well as part of the decline in interest rates.

My conclusion is simple: if one wants to build a heterogeneous agent model that broadly matches the data on the distribution of income and wealth, one can do so relatively simply by taking the approach of this paper and using a labour income process that can be easily estimated from survey data. One does not necessarily need to add in features like capital income risk, entrepreneurship and financial frictions, life-cycles and bequests, or any of the myriad other approaches taken in the literature, which all have their own merits, but which introduce additional mechanisms and complications beyond the standard model.
References


Appendix: Estimated transition matrices

Figure 6: Estimated UK productivity states and discrete time annual transition matrix

<table>
<thead>
<tr>
<th>Earnings percentile</th>
<th>z</th>
<th>T</th>
<th>( \Lambda )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>Non-employment</td>
<td>0.44</td>
<td>0.44</td>
<td>0.44</td>
</tr>
<tr>
<td>-15-30</td>
<td>0.32</td>
<td>0.36</td>
<td>0.32</td>
</tr>
<tr>
<td>30-40</td>
<td>0.54</td>
<td>0.26</td>
<td>0.54</td>
</tr>
<tr>
<td>40-50</td>
<td>0.72</td>
<td>0.18</td>
<td>0.72</td>
</tr>
<tr>
<td>50-55</td>
<td>0.87</td>
<td>0.14</td>
<td>0.87</td>
</tr>
<tr>
<td>55-60</td>
<td>0.95</td>
<td>0.13</td>
<td>0.95</td>
</tr>
<tr>
<td>60-65</td>
<td>1.08</td>
<td>0.11</td>
<td>1.08</td>
</tr>
<tr>
<td>65-70</td>
<td>1.17</td>
<td>0.10</td>
<td>1.17</td>
</tr>
<tr>
<td>70-75</td>
<td>1.27</td>
<td>0.09</td>
<td>1.27</td>
</tr>
<tr>
<td>75-80</td>
<td>1.43</td>
<td>0.08</td>
<td>1.43</td>
</tr>
<tr>
<td>80-82.5</td>
<td>1.53</td>
<td>0.07</td>
<td>1.53</td>
</tr>
<tr>
<td>82.5-85</td>
<td>1.61</td>
<td>0.09</td>
<td>1.61</td>
</tr>
<tr>
<td>85-87.5</td>
<td>1.70</td>
<td>0.07</td>
<td>1.70</td>
</tr>
<tr>
<td>87.5-90</td>
<td>1.84</td>
<td>0.06</td>
<td>1.84</td>
</tr>
<tr>
<td>90-92.5</td>
<td>1.98</td>
<td>0.06</td>
<td>1.98</td>
</tr>
<tr>
<td>92.5-95</td>
<td>2.17</td>
<td>0.05</td>
<td>2.17</td>
</tr>
<tr>
<td>95-97.5</td>
<td>2.55</td>
<td>0.07</td>
<td>2.55</td>
</tr>
<tr>
<td>97.5-99.9</td>
<td>3.23</td>
<td>0.06</td>
<td>3.23</td>
</tr>
<tr>
<td>99.9-100</td>
<td>6.10</td>
<td>0.08</td>
<td>6.10</td>
</tr>
</tbody>
</table>

Source: BHPS and author's calculations
### Figure 7: Estimated UK productivity and continuous time quarterly transition matrix

<table>
<thead>
<tr>
<th>Earnings percentile</th>
<th>A</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-employment</td>
<td>0</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>−15−30</td>
<td>0.32</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30−40</td>
<td>0.53</td>
<td>0.26</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40−50</td>
<td>0.71</td>
<td>0.18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50−60</td>
<td>0.88</td>
<td>0.14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60−70</td>
<td>0.94</td>
<td>0.13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70−75</td>
<td>1.06</td>
<td>0.11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80−82.5</td>
<td>1.15</td>
<td>0.10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>82.5−85</td>
<td>1.25</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85−87.5</td>
<td>1.35</td>
<td>0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>87.5−90</td>
<td>1.45</td>
<td>0.07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90−92.5</td>
<td>1.55</td>
<td>0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>92.5−95</td>
<td>1.65</td>
<td>0.05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95−97.5</td>
<td>1.75</td>
<td>0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>97.5−99</td>
<td>1.85</td>
<td>0.03</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>99.9−100</td>
<td>2.00</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: BHPS and author’s calculations
Figure 8: Estimated US productivity and continuous time semi-annual transition matrix

<table>
<thead>
<tr>
<th>Earnings percentile</th>
<th>z</th>
<th>T</th>
<th>λ</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-employment</td>
<td>0</td>
<td>0.13</td>
<td>1</td>
<td>-12.9%</td>
<td>11.9%</td>
<td>0.4%</td>
<td>0.4%</td>
<td>0.0%</td>
<td>0.1%</td>
<td>0.2%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>~0.17 – 0.3</td>
<td>0.15</td>
<td>0.07</td>
<td>2</td>
<td>8.3%</td>
<td>-35.2%</td>
<td>20.4%</td>
<td>2.1%</td>
<td>2.5%</td>
<td>0.6%</td>
<td>0.7%</td>
<td>0.5%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>0.3 – 0.4</td>
<td>0.36</td>
<td>0.04</td>
<td>3</td>
<td>2.8%</td>
<td>15.2%</td>
<td>-43.4%</td>
<td>21.4%</td>
<td>1.8%</td>
<td>1.7%</td>
<td>0.4%</td>
<td>0.1%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
<tr>
<td>0.4 – 0.5</td>
<td>0.53</td>
<td>0.05</td>
<td>4</td>
<td>2.1%</td>
<td>3.4%</td>
<td>10.3%</td>
<td>-35.2%</td>
<td>16.9%</td>
<td>1.2%</td>
<td>1.1%</td>
<td>0.0%</td>
<td>0.2%</td>
<td>0.0%</td>
</tr>
<tr>
<td>0.5 – 0.6</td>
<td>0.67</td>
<td>0.04</td>
<td>5</td>
<td>1.5%</td>
<td>1.5%</td>
<td>1.6%</td>
<td>19.4%</td>
<td>-42.1%</td>
<td>15.3%</td>
<td>2.1%</td>
<td>0.7%</td>
<td>0.1%</td>
<td>0.0%</td>
</tr>
<tr>
<td>0.6 – 0.7</td>
<td>1.04</td>
<td>0.04</td>
<td>6</td>
<td>1.5%</td>
<td>0.7%</td>
<td>1.4%</td>
<td>0.0%</td>
<td>20.5%</td>
<td>-40.2%</td>
<td>14.8%</td>
<td>0.9%</td>
<td>0.5%</td>
<td>0.0%</td>
</tr>
<tr>
<td>0.7 – 0.8</td>
<td>1.23</td>
<td>0.04</td>
<td>7</td>
<td>2.1%</td>
<td>0.1%</td>
<td>0.2%</td>
<td>1.1%</td>
<td>0.0%</td>
<td>14.1%</td>
<td>-32.2%</td>
<td>14.0%</td>
<td>0.7%</td>
<td>0.0%</td>
</tr>
<tr>
<td>0.8 – 0.9</td>
<td>1.71</td>
<td>0.04</td>
<td>8</td>
<td>1.5%</td>
<td>0.6%</td>
<td>0.0%</td>
<td>0.3%</td>
<td>0.7%</td>
<td>0.0%</td>
<td>8.3%</td>
<td>-19.8%</td>
<td>8.4%</td>
<td>0.0%</td>
</tr>
<tr>
<td>0.9 – 0.99</td>
<td>2.71</td>
<td>0.04</td>
<td>9</td>
<td>1.4%</td>
<td>0.2%</td>
<td>0.2%</td>
<td>0.0%</td>
<td>0.3%</td>
<td>0.1%</td>
<td>0.3%</td>
<td>6.7%</td>
<td>-10.6%</td>
<td>1.7%</td>
</tr>
<tr>
<td>0.99 – 1</td>
<td>17.24</td>
<td>0.03</td>
<td>10</td>
<td>4.8%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>0.0%</td>
<td>1.1%</td>
<td>0.3%</td>
<td>0.9%</td>
<td>0.2%</td>
<td>8.8%</td>
<td>-16.0%</td>
</tr>
</tbody>
</table>

Source: Author’s calculations using PSID and SCF data
Should I Stay (in School) or Should I Go (to Work)

Lee Tyrrell-Hendry
University of Edinburgh

September 2022
(Click here for latest version)

Abstract

I explore optimal education subsidies and progressivity of labour taxes in a model with stochastic human capital accumulation and incomplete markets, endogenous labour supply and an education choice modelled as a stopping time problem, where agents choose an optimal number of years to study before starting work. In a purely analytical Baseline model with tight borrowing constraints on students, which leads to a no-trade equilibrium without savings, the government pays for education via transfers to students or – equivalently – via grants to universities. The social welfare-maximising policy features generous student transfers and highly progressive labour taxes, much more so than currently seen in the US or Europe. This result is robust to myriad extensions, including a Quantitative model with relaxed financial frictions where students can borrow to finance their education, and where hence the equilibrium features extensive precautionary saving by workers.
1 Introduction

This paper is concerned with two key macroeconomic and public policy questions:

1. How should a government calibrate redistributive fiscal policy to maximise welfare, weighing the benefits of redistribution, social insurance against adverse shocks and revenue generation, against the deleterious effects of progressive taxation on incentives to work and to accumulate human capital?

2. What manner and magnitude of policy to finance higher education best optimises welfare? a) market-set tuition fees with transfers to students; b) fee caps/subsidies to universities (set by assumption to ensure equilibrium in the schooling market and avoid rationing of education); or c) student loans, potentially alongside transfers.

These questions have been tackled before, but much of the existing literature that explicitly focuses on education and optimal policy primarily features either a time-allocation problem, à la Ben-Porath (1967), for example Benabou (2002), or an otherwise highly-stylised education choice, for example Heathcote et al. (2017). These approaches alternately miss two important features of educational choice: (1) that it is often a one-time choice, and it can be difficult – and is indeed rare – for people to re-enter higher education later in life; and (2) education has real resource and time costs that must be paid for out of the usually limited resources available to prospective students. The former approach then potentially overstates the ability of workers to insure themselves against adverse shocks – obsolescence of their human capital – by re-entering education, and thus understates the welfare gains from public insurance via redistributive taxation. One can view this as the human capital counterpart of the well-known bug of simple Aiyagari-style incomplete markets models: that households engage in counterfactually extreme precautionary saving in physical capital that renders the welfare effects of adverse shocks – and thus the welfare benefits of public insurance – trivial; see for example Aiyagari (1994), Krusell & Smith (1998), Krusell et al. (2009) and Krusell et al. (2010). In essence, I take the opposite extreme in this paper: I treat education as a stopping time problem,
where it is impossible to re-enter education after leaving, so “precautionary education” is only possible once, before starting work. The second approach – that assumes limited resource and opportunity costs of education and hence assumes away any financial frictions that may limit access to education – likely understates the welfare gains from redistributing resources to students. In the first part of this paper I again take the opposite extreme that students are completely unable to borrow. I relax this subsequently, and allow students to borrow up to a – potentially large – exogenous limit.

To answer these questions and address these limitations of the existing literature, I develop a suite of incomplete markets models with educational choice modelled as an optimal stopping problem. Agents are born as students, accumulate human capital while studying and then optimally choose when to graduate; thereafter, agents also choose labour supply and accumulate human capital while working. In the Baseline model, all agents are ex-ante homogeneous and face constant mortality risk; moreover, I make assumptions that deliver an equilibrium with no trade in bonds, so all agents consume their income. Both students and workers face uninsurable, permanent idiosyncratic shocks to their human capital, although when students face no risk I am able to derive fully analytical results for the value functions, optimal choices, equilibrium prices, stationary distributions of human capital and aggregate social welfare.

This Baseline model is not that realistic, in particular it implies households do not in practice insure themselves against shocks other than by staying in school longer, thus potentially overstating the welfare benefits of subsidising education and of progressive taxation. Moreover, it precludes the main source of financing for higher education in most Anglophone countries: student loans. However, the simple model does clarify some of the mechanisms of the purely quantitative model, which I subsequently introduce, and which does allow students to borrow and features precautionary saving in equilibrium.

I find that optimal policy features highly progressive taxes, far more so than in the US or Europe today, as well as generous transfers to students. These policy conclusions are robust
to a broad range of extensions over the Baseline model. In Figure 1, I summarise the optimal policies in the Baseline model, several diverse extensions, and the full Quantitative model. The Baseline model is represented by the blue dot, and some indicative variations on it in light blue; the analytical extensions that explore for example the effects of decreasing returns to scale, ex-ante heterogeneity in ability, or a different life-cycle process where agents age, are coloured orange; and the full Quantitative model is in red.

In this Baseline model with no borrowing by students allowed, and hence no trade in bonds among workers either, the government must finance education via direct transfers to students or, equivalently, via grants to universities. There are large welfare gains from doing so: optimal policy features generous transfers, funded by highly progressive taxes. The gains from transfers come in part from overcoming the onerous financial friction facing students. The gains from progressive taxation come both from effectively insuring workers against repeated bad shocks, and from redistributing from rich to poor, and these benefits to a great extent outweigh the costs of disincentivising study and work when labour supply is plausibly elastic. To show this, I first shut down idiosyncratic risk among workers, which implies only modestly less generous transfers and less progressive taxes are optimal; then I shut down growth in human capital while working, so there is no inequality among workers, which implies substantially lower (though still positive) transfers and regressive labour taxes are optimal; in that case progressive taxes simply disincentivise education and work without any benefit. When the returns to schooling are risky, the benefits from transfers to students are smaller, because some students get little out of schooling yet are still incentivised to attend, so optimal transfers are slightly lower.

The main downside of the Baseline model is that the financial friction imposed on students is arguably counterfactually strict: students do typically borrow to finance their education (although arguably this is in many cases only possible due to government intervention). A side-effect of this (and some other assumptions discussed in the body of the text) is that the equilibrium of the Baseline model features no trade in bonds, so despite in principle some
workers wishing to borrow to smooth consumption, or others wishing to insure themselves against risk, interest rates adjust so that none in practice do so.

In the Quantitative model, I relax the borrowing constraint on students, which means not only that workers must build wealth to pay off their debts, but that in equilibrium workers do in fact borrow to smooth consumption and save in precaution against risk, which thus gives rise to more interesting, diverse and perhaps realistic behaviour at the individual level, and a different outcome at the aggregate level. Nevertheless, optimal transfers in the Quantitative model – though smaller than in the Baseline model – are still comfortably positive. And let me emphasise that even though students are able to borrow to finance several years of schooling and consumption, the socially optimal government policy is still to entirely pay for all students’ education, plus a little extra. Optimal labour taxes also remain highly progressive, much more so than the current tax schedules in the US (shown by the dashed line in the figure below) or Europe.

![Figure 1: Summary of optimal policies in core models and extensions](image)

Given the core of the paper is to find some optimal degree of government intervention, let me outline the justifications for government intervention, and the limitations I place on the
possible forms of intervention.

There are two justifications for intervention in the model, related to equity and efficiency:

1. **Equity.** Governments in the model are assumed to operate under a Utilitarian social welfare function, so regardless of any inefficiencies present, the government may still wish to intervene to transfer resources to less well-off members of society, be they poorer workers or students. In this regard the paper contrasts with Benabou (2002), for example, who primarily focuses on intervention to pursue efficient allocations, leaving equity to one side.

2. **Efficiency.** There are two forms of inefficiency present in the model:

   (a) Incomplete markets. Agents are unable to fully insure themselves against idiosyncratic risks, both as workers – where idiosyncratic shocks to one’s human capital directly affect one’s wage income – but also as students, where idiosyncratic shocks to human capital affect one’s future wage income and hence the option value of graduating.

   (b) Borrowing constraints. In the Baseline model, students are completely disallowed from borrowing; even when borrowing is allowed, there is still a limit. Since students must pay for their education up front, this friction is particularly constricting. Financing higher education publicly through transfers to students or grants for educational institutions can thus serve to alleviate the borrowing constraint.

I do not however allow the government free reign to intervene in markets. I restrict the government’s role to that of writing the “rules of the game”: designing the tax and transfers system with constant policies in order to maximise aggregate steady state welfare, measured in a Utilitarian sense. I furthermore take the Ramsey (1927) approach in restricting government taxes and transfers to a specific parametric class: an isoelastic tax schedule for workers (also used in for example Heathcote et al. 2017 recently) and homogeneous, lump-sum transfers to students. I also allow for tuition fee caps, with subsidies to universities then determined
endogenously to balance supply and demand in the schooling market, but I show these are in fact equivalent to transfers in this simplified setting. I do not allow the government to engage in direct lump-sum transfers between workers. I do not take the Mirrlees (1971) approach of designing an incentive-compatible tax system for an information-constrained government. And I do not explore either the first-best or the constrained efficient solution of a social planner (as in Dávila et al. 2012 and Nuño & Moll 2018 under incomplete markets), instead focusing on the optimal policy regime conditional on individual households optimising separately and transacting freely.

**Literature review.** There is a rich literature incorporating endogenous human capital accumulation into models with incomplete markets. Notable contributions include Huggett et al. (2006), (2011), which explore life-cycle human capital accumulation in a Ben-Porath (1967) model with idiosyncratic ability in order to discuss the extent to whether the ability agents are born with, or the human capital they accumulate during their lifetimes, explains the distribution of income in the US, but they say little about policy or welfare. Alon et al. (2020) explore the effect of student debt on on-the-job human capital accumulation in a combination Ben-Porath and Roy model, which they take to the data. They find that those with higher student debt earn more initially, but their returns to experience are lower in subsequent years. They argue that credit constrained individuals select into occupations with less scope for on-the-job human capital accumulation.

The literature on optimal tax & transfer policy is far too broad in scope to cover here. Instead I highlight some key papers that feature education or skill acquisition prominently. One line of research explores parental investment in children’s education, in part to explain the persistence of earnings across generations, and also to discuss its implications for optimal education policy, using an overlapping generations structure with explicit parental altruism over children’s welfare, starting with Loury (1981), later Caucutt & Kumar (2003) and Restuccia & Urrutia (2004), and recently Abbott (2022). I abstract away from these features and instead concern myself more with higher education choices where agents have more
autonomy and typically rely less on parental resources. This may overlook an important channel for investment in human capital; however, at least Restuccia & Urrutia (2004) and Abbott (2022) find the welfare gains from subsidies to parental investment in college-age children to be minimal. The slightly more distantly related work of Krueger & Ludwig (2013) and (2016) also features parental altruism and investment in education, but takes a highly quantitative approach and is thematically closer to my paper, exploring optimal progressivity of taxation and subsidies for education, though their approach is distinct from mine. They also find highly progressive taxes and generous tuition subsidies are optimal, but moderated somewhat by the welfare costs of transitioning to such a policy.

Analytical approaches that feature no-trade equilibria or otherwise dispense with savings are a common feature in this literature, even in the presence of uninsurable risks, so by construction may overstate the welfare gains from social insurance. Benabou (2002) explores the welfare effects of education finance policy under incomplete markets, modelling schooling choice as a convex time allocation problem rather than a stopping time problem, à la Ben-Porath (1967), which as discussed may understimate the welfare costs of human capital losses. Heathcote et al. (2017) explore optimal progressivity of taxation, in a model that features both insurable and uninsurable idiosyncratic risk, endogenous labour supply, and a stylised educational choice at the start of agents’ lives. My paper explores similar thematic ground but in a very different framework. In particular, I treat education very differently: in their model, skill investment has no resource cost or opportunity cost, only a utility cost, which as discussed above has limitations. As well as capturing these aspects of education, I can also discuss the welfare effects of student loans and policies relating to education finance, which have no role in their paper. Furthermore, in the body of the paper I compare their findings for optimal policy with my own, and note that subtle modelling choices in their paper, such as a flat life-cycle earnings profile, limit the extent of intergenerational inequality and hence the welfare gains from redistribution via progressive taxation, even if they do capture the welfare gains from social insurance against uninsurable private risk that progressive taxes can
provide. Inequality in their model arises more due to unequal skill acquisition; progressive taxation dampens this variation, but at the cost of lowering output. Moreover, both these papers only explore no-trade equilibria, so are mute on the implications of precautionary saving by workers or borrowing by students for welfare and optimal policy. In the first half of my paper I also take the no-trade analytical approach, but in the second half I allow students to borrow, which results in precautionary saving by workers in equilibrium to clear bond markets. I show that this does not change the core policy conclusions of the Baseline model: that progressive taxation and generous transfers to students enhance welfare.

Treating education as a stopping time problem is relatively uncommon in the literature, but not without precedent. Card (2001) was among the first to use this approach, albeit for the vastly different purpose of estimating the returns to schooling. Hogan & Walker (2003), (2007) is perhaps methodologically the most similar to my paper: using a purely analytical approach they discuss the effects of policy on schooling decisions in partial equilibrium, but do not explicitly consider distributional matters or the welfare effects of policy. My paper develops these aspects further, as well as supplementing the analytical work with a quantitative approach to explore the role of student debt and precautionary saving.

Mellior (2021) may be the most recent paper to explore the welfare consequences of different education financing policy regimes under incomplete markets, in a purely quantitative paper that uses a similar approach to the second half of my paper. However, the economic environment, modelling choices and the policy regimes explored are quite different, so my analysis should be seen as complementing his, rather than extending it. For example, in his paper agents study for the chance to graduate with a degree and a job, which gives them a higher (fixed) wage and lower chance of becoming unemployed; whereas in my paper agents accumulate human capital while studying (and more slowly while working). Moreover, his paper is very careful to model precise features of real-world student loan policies, and compares the welfare consequences of each, finding that when education is as lengthy and costly as it is currently, government-guaranteed income-contingent loans coupled with tuition fee subsidies
are the optimal policy. My paper simplifies but broadens the policy space, exploring optimal progressivity of taxation and generosity of direct transfers to students, alongside student loans and fee subsidies.

2 Baseline Model: No-Trade Equilibrium

The model has an overlapping generations structure in continuous time, with educational choice modelled as a stopping time problem: agents choose early in their life to school for a certain number of years, $S$, and then start working. After graduation, workers choose labour supply, $n$, only and cannot re-enter education. In principle, workers and students can also choose how much to consume, $c$, and save, $\dot{a}$, however, in the Baseline model I make some assumptions to deliver a no-trade equilibrium with tractable analytical results. First, I assume everyone is born with zero wealth ($a_0 = 0$). For standard intertemporal substitution reasons, this implies students would like to borrow – provided the transfer they receive to pay for their education is not too large – but, secondly, I assume they cannot do so. Thus they will carry zero wealth through to graduation. Thereafter, in principle I allow working households to borrow or save. However, because, thirdly, there is no capital and bonds are in zero net supply, and, fourthly, because their income process features stochastic growth with permanent – not mean-reverting – shocks, the interest rate will adjust to clear bond markets so that they do not wish to do so, implying they simply consume their post-tax income. Thus, despite workers’ desire to engage in precautionary saving to self-insure against the idiosyncratic risk they face from their stochastic accumulation of human capital, $h > 0$, in practice they do not do so.

Households

In the simple model, households are ex-ante homogeneous and face two sources of idiosyncratic risk: stochastic human capital accumulation, both while a student and while working, and
uncertain lifetimes. Agents die randomly, with constant mortality rate, $\lambda > 0$, à la Blanchard (1985); in an extension I explore a fixed, finite lifetime.

Human capital, $h$, is accumulated stochastically through schooling and work experience:

$$
\begin{align*}
\frac{dh}{dt} &= \begin{cases} 
(\mu_s h^{\beta_s} - \delta_s h) dt + \sigma_s h dW & \text{if schooling} \\
\mu_w h dt + \sigma_w h dW & \text{if working}
\end{cases}
\end{align*}
$$

where $\mu_i \geq 0$, $\beta_s \leq 1$ and $\delta_i \geq 0$. When $\beta_s < 1$, there are decreasing returns to learning; and with $\delta_i > 0$ in addition, there is a maximum level of human capital attainable, $\bar{h} = \left(\frac{\mu_s}{\delta_s}\right)^{1/\beta_s}$, attained in the Faustian limit if one stays in school forever. When $\beta_s = 1$, as it is for workers, the human capital process is a geometric Brownian motion with drift $\mu_s - \delta_s$, and one can derive analytical solutions for the value function, though for students this case is somewhat pathological and leads to extreme policy implications, so I do not focus on it. When $\sigma_s = 0$ also, one can derive analytical expressions for the distributions of human capital and hence the equations determining the equilibrium, which I make use of later.

Household income consists of capital income, $r_a$, and their post-tax and transfer labour earnings, which for students is simply their transfer net of tuition fees, $T = \tilde{T} - F$, and which for workers is an isoelastic function of pre-tax earnings, $\tau_0 (wnh)^{1-\tau_1}$, where pre-tax earnings are the product of a piece-rate wage, $w$, hours worked, $n$, and human capital, $h$. This tax function dates back at least as far as Feldstein (1969) and has been used frequently since then, including in Heathcote et al. (2017) recently. $\tau_1 \in [-1, 1]$ is a measure of the progressivity of taxes: when $\tau_1 = 0$, taxes are proportional with tax rate $1 - \tau_0$; when $\tau_1 > 0$, taxes are progressive (and regressive when $\tau_1 < 0$), with transfers to lower income households, and $\tau_0 > 0$ determining the level of taxation required to balance the budget against whatever other government spending there is. For simplicity I assume there is no other government spending besides transfers to workers, students or educational institutions.

Labour supply is also endogenous, allowing households a second margin of adjustment in response to progressive labour taxation – reducing the number of hours they work – in
addition to spending less time in education. With endogenous labour supply, one faces the
delicate question of preference specification. If income effects are too strong, for example
with separable preferences with CRRA utility of consumption with risk aversion $\gamma > 1$, as
in MaCurdy (1981), hours worked is negatively related to and convex in human capital. As
this gives the broadly counterfactual result that higher income people work fewer hours, this
seems like an implausible calibration. On the other hand, weak income effects that generate a
plausibly positive and concave relationship between hours worked and human capital require
implausibly low risk aversion, $\gamma < 1$. I thus focus solely on balanced growth preferences where
income and substitution effects cancel, which in a no-trade equilibrium without savings implies
constant labour supply irrespective of human capital. In particular, the utility function that I
use is a refinement of the specification of King et al. (1988), and was also used by Trabandt
& Uhlig (2011), and which nests both extremes of inelastic labour supply with CRRA utility
and the infinite-elasticity Cobb-Douglas specification, while retaining both constant relative
risk aversion, $\gamma > 0$, and constant, finite Frisch elasticity of labour supply, $\varphi \geq 0$, without
restricting either parameter:

$$u(c, n) = \begin{cases} 
\frac{c^{1-\gamma}v(n)-1}{1-\gamma} & \text{where } v(n) = \left(1 - \psi(1 - \gamma)\frac{n^{1+1/\varphi}}{1+1/\varphi}\right)^{\gamma} \text{ if } \gamma \neq 1 \\
\ln(c) - \psi\frac{n^{1+1/\varphi}}{1+1/\varphi} & \text{if } \gamma = 1
\end{cases}$$

where $\psi \geq 0$ is a shifter ensuring labour is a bad. When $\psi = 0$ or $\varphi \to 0$, labour supply is
inelastic and utility is standard CRRA in consumption only. When $\varphi \to \infty$ - and $\psi = -\frac{1}{1-\gamma}$
when $\gamma > 1$ or $\psi = \frac{1}{1-\gamma}$ when $\gamma < 1$ - the utility function reduces to Cobb-Douglas preferences:

$$u(c, n) = \frac{c^{1-\gamma}(1-\psi(1-\gamma)n)^{\gamma} - 1}{1-\gamma}.$$ 

**The workers’ problem.** Workers choose consumption, $c$, hours worked, $n$, and savings,
$da$, to maximise utility over their remaining lifetime, subject to their budget constraint, the
exogenous process for human capital and an exogenous borrowing limit, $a < 0$:

$$V^w(a_S, h_S) = \max_{\{c_t, n_t\} \in [S, \infty)} \mathbb{E}_S \left[ \int_S^\infty e^{-(\rho + \lambda)(t-S)} u(c_t, n_t) dt \right]$$

s.t. 

$$da_t = (ra_t + \tau_0(wn_t h_t)^{1-\tau_1} - c_t) dt$$

$$dh_t = \mu_w h_t dt + \sigma_w h_t dW_t$$

$$a \geq a$$

The workers’ problem can also be re-written as a Hamilton-Jacobi-Bellman (HJB) equation:

$$(\rho + \lambda)V^w(a, h) = \max_{c, n} u(c, n) + V^w_a \left( ra + \tau_0(wn h)^{1-\tau_1} - c \right) + V^w_h \mu_w h + \frac{1}{2} V^w_h \sigma_w^2 h^2$$

We cannot in the general case derive a closed-form solution to this HJB equation, so I proceed quantitatively, except in some important special cases. In particular, in the absence of an asset in positive net supply, and with students unable to borrow, there exists a no-trade equilibrium in which all households holds consume their post-tax income and hold no wealth. For a fuller discussion of this no-trade equilibrium, see the Appendix.

Given this no-trade equilibrium, we can derive workers’ optimal labour supply by their static first-order condition: $u_n = 0$, after substituting in $c(h) = \tau_0(wn h)^{1-\tau_1}$ from the budget constraint. This yields a labour supply that is independent of human capital and decreasing in the progressivity of labour taxes:

$$n = \left\lfloor \frac{(1 - \tau_1)(1 + 1/\varphi)}{\psi \gamma (1 + 1/\varphi) + (1 - \tau_1) \psi (1 - \gamma)} \right\rfloor^{1+1/\varphi}$$

**The students’ problem.** All newborns start as students with the same level of human capital, $h_0 = 1$. Students are assumed not to be able to work while studying, $n = 0$, so their only choice is how much to save/consume and when to stop education and become a worker to maximise intertemporal utility subject to the exogenous process for human capital
accumulation:

\[
V^s(a_0, h_0) = \max_{S, \{c_t\}_{t \in [0, S)}} \mathbb{E}_0 \left[ \int_0^S e^{-(\rho + \lambda)t} u(c_t, n_t) dt + e^{-\rho S} V^w(a^*(S), h^*(S)) \right]
\]

s.t. \[
da_t = (ra_t + T - c_t) dt
\]
\[
dh_t = \left( \mu_s h_t^{\beta_s} - \delta_s h_t \right) dt + \sigma_s h_t dW_t
\]
\[
n_t = 0 \quad a_0 = 0 \quad a \geq a^* \quad h_0 = 1
\]

The optimal stopping time, \( S \in [0, \infty) \), is associated with a threshold of values of wealth and human capital, \((a^*, h^*)\), outside of which the student will choose to graduate, and within which they will choose to stay in school. Denote by \( E \) (for education) the region of \((a, h)\)-space bounded below by the borrowing constraint \( a = a^* \) (ignoring for a moment the no-trade equilibrium) & \( h = 0 \), and above by the threshold curve \( h^*(a^*) \). For \((a, h) \in E\), students prefer to stay in education, so their value function exceeds that of workers; outside this region their value functions are equal:

\[
V^s(a, h) \geq V^w(a, h) \quad (a, h) \in E
\]
\[
V^s(a, h) = V^w(a, h) \quad (a, h) \notin E
\]

The students' problem, while they remain students, with \((a, h) \in E\), can be described by the HJB equation:

\[
(\rho + \lambda)V^s(a, h) = \max_c u(c) + V^s_a (ra + T - c) + V^s_h \left( \mu_s h^{\beta_s} - \delta_s h \right) + \frac{1}{2} V^s_{hh} \sigma^2 h^2
\]  

(2)

which holds with inequality for \((a, h) \notin E\). These equations and inequalities can be written...
more compactly as the following HJB variational inequality:

\[
(\rho + \lambda) V^*(a, h) = \max \left\{ \max_c u(c) + V^*_a (ra + T - c) + V^*_h \left( \mu_s h^{\beta_s} - \delta_s h \right) + \frac{1}{2} V^s_{hh} \sigma_s^2 h^2, \right. \\
\left. (\rho + \lambda) V^w(a, h) \right\}
\]  

As alluded to above, both the workers’ and students’ problems can be simplified considerably by making the appropriate assumptions to deliver a no-trade equilibrium.

**Theorem 1.** A no-trade equilibrium exists in which all agents consume their income and hold zero wealth.

**Proof.** See Appendix.

Predicating on the no-trade equilibrium, wealth can be dropped as a state variable from both students’ and workers’ problems. The threshold at which students choose to graduate thus becomes a single point, \(h^* \in [h_0, \overline{h}]\), where recall \(\overline{h}\) was the maximum attainable human capital if one stays in school forever.

**Lemma 1.** In the Baseline model with a no-trade equilibrium, the workers’ value function is:

\[
V^w(h) = \begin{cases} \\
\frac{1}{\kappa} \left( \frac{\tau_0(whh)^{1-\gamma_1} - 1 - \gamma v(n)}{1-\gamma} \right) - \frac{1}{(\rho+\lambda)(1-\gamma)} & \text{if } \gamma \neq 1 \\
\frac{1}{\rho+\lambda} \left[ \ln(\tau_0(whh)^{1-\gamma_1}) - \psi \frac{n^{1+1/\phi}}{1+1/\phi} \right] + \frac{1-\gamma_1}{(\rho+\lambda)^2} (\mu_w - \sigma^2_w / 2) & \text{if } \gamma = 1
\end{cases}
\]  

where \(\kappa = \rho + \lambda - (1 - \tau_1)(1 - \gamma_1)\mu_w - \frac{1}{2} (1 - \tau_1)(1 - \gamma_1)(1 - \gamma_1) \left[ (1 - \tau_1)(1 - \gamma) - 1 \right] \sigma^2_w\), and where \(n\) is the optimal labour supply, given above.

Moreover, the students’ value function takes the form:

\[
V^s(h) = \frac{u(T)}{\rho + \lambda} + E(h)
\]

where \(E(h)\) satisfies the following ODE: \((\rho + \lambda) E(h) = E'(h) \left( \mu_s h^{\beta_s} - \delta_s h \right) + \frac{1}{2} E''(h) \sigma_s^2 h^2\)

**Proof.** Plug these into the workers’ and students’ HJB equations.
The parameter $\kappa$ in the workers’ value function is a kind of discount rate that accounts for the growth and variability of workers’ future income, and is an inverse quadratic function of tax progressivity. Note, in the students’ value function, the first term – the discounted utility from student grants net of fees ($T = \tilde{T} - F$) – represents the value of staying a student forever, and the second term represents the option value of graduating.

As is standard in stopping time models, the threshold $h^*$ is determined by the value matching and smooth pasting conditions that the level of the students’ and workers’ value functions – and their slopes – should be equal at the cut-off:

Value matching:
$$V^w(h^*) = V^s(h^*) \iff \frac{1}{\kappa} \left( \frac{\tau_0 (wnh^*)^{1-\gamma}}{1-\gamma} v(n) \right) = \frac{u(T)}{\rho + \lambda} + E(h^*)$$

Smooth pasting:
$$V^h_w(h^*) = V^s(h^*) \iff \frac{1-\tau_1}{\kappa} \left( \frac{\tau_0 (wnh^*)^{1-\gamma}}{1-\gamma} v(n) h^{*-1} \right) = E'(h^*)$$

When the cut-off is below the starting level of human capital ($h_0 = 1$), students will choose to start work immediately, so $h^* = 1$.

**Distribution of human capital.** Now that we have solved the households’ problem, i.e. solved for the threshold level of human capital, we can solve for the distribution of human capital, and the distribution of students and workers. The distributions of human capital across students and workers is governed by the Kolmogorov Forward Equations (KFE):

$$\dot{g}^s(h) = -\partial_h \left[ (\mu_s h^{\beta_s} - \delta_s h) g^s(h) \right] + \partial_{hh} \left[ \frac{1}{2} \sigma^2_s h^2 g^s(h) \right] - \lambda g^s(h) \quad h \in (0, h^*) \setminus h_0 \quad (6)$$

$$\dot{g}^w(h) = -\partial_h \left[ \mu_w h g^w(h) \right] + \partial_{hh} \left[ \frac{1}{2} \sigma^2_w h^2 g^w(h) \right] - \lambda g^w(h) \quad h \in (0, \infty) \setminus h^* \quad (7)$$

The first terms in each equation represent the growth in human capital, the second terms its variation. The final terms represent the “outflows”, i.e. the change in the measure of agents owing to their deaths. Each group also has “inflows” of new agents at certain points: newborn students at the initial level of human capital, $h_0$, and new workers at the graduation threshold, $h^*$. In the steady-state, these equations will equal 0, giving rise to a stationary
distribution. When the human capital of students is stochastic, there is to my knowledge no closed-form expression for its stationary distribution; I discuss the deterministic case with closed-form solutions later.

**Supply side & market clearing**

Consumption goods and education goods are supplied by competitive firms using labour supplied by households and a constant returns to scale production function, $f_i(L^i)$. One can also consider the slightly more general specification where production is allowed to face decreasing returns, and where the resulting profits are assumed to be either paid to some unmodelled factor of production, or distributed to foreign owners that otherwise take no part in the economy. The difference is not qualitatively important for the equilibrium or policy implications, and comes at the expense of two extra equilibrium equations, so I focus on the CRS case for simplicity, but naturally this will mean that changes in policy have no effect on relative prices: wages and tuition fees. I consider the DRS case in an extension.

**Educational institutions.** Educational institutions produce schooling using the constant returns to scale technology, $S = A^E L^E$, charge fees $F$ and receive government grants $G$ proportional to the amount of schooling supplied, and choose the amount of labour to employ to maximise profits. For simplicity I assume workers with different quantities of human capital are infinitely substitutable, so:

$$\pi^E = \max_{L^E} (F + G)A^E L^E - wL^E$$

The schools’ first-order condition implies labour is paid its marginal product: $w = (F + G)A^E$. In the free market equilibrium, relative tuition fees, $F$, adjust so that, conditional on any grants paid by the government, total demand for schooling across all agents equals that
supplied. Our second equilibrium condition:

\[
S^S \equiv A^E L^E = \int_{0}^{h^*} g^*(h)dh \equiv S^D
\]  

(9)

One can also imagine a scenario where the government caps or fixes tuition fees, and instead sets grants \( G \) to clear the market for schooling. Below we will see these are equivalent and do not change the equilibrium. This assumption that grants clear the market when fees cannot is convenient, but overlooks one possible consequence of tuition fee caps, which is that subsidies from government fall short of that need to equilibrate markets, and hence education is rationed. I am currently exploring this idea in separate work.

**Firms.** Firms produce consumption goods, maximising profits in a competitive environment, operating a potentially different CRS technology to educational institutions, but where the price of the consumption good is normalised to 1:

\[
\pi^F = \max_{L^F} A^F L^F - wL^F
\]  

(10)

Likewise, firms pay labour – which is free to move across sectors – their marginal product, implying the marginal revenue product is equalised across sectors, yielding production efficiency and also pinning down the equilibrium wage and tuition fee:

\[
w = A^F \quad \text{and} \quad F = \frac{A^F}{A^E} - G
\]

Note, the model can explain the rise in tuition fees in the US as a form of Baumol’s cost disease: productivity in the non-education sector rising relative to the education sector. It cannot however incorporate demand-led relative price rises, where for example more generous tuition fee subsidies/student transfers (or student loans) lead to higher tuition fees; this would require decreasing returns to scale, which I explore as an extension.

Effective total labour supply is given by the total human capital of all workers – again
workers with different levels of human capital are perfect substitutes – multiplied by their working hours. Labour market equilibrium is thus given by:

\[ L^S \equiv \int_0^\infty nhg^w(h)dh = L^F + L^E \equiv L^D \] (11)

**Bond market.** Capital is not used in production, so households’ net wealth is entirely saved in bonds, which are risk-free and in zero net supply, as in Huggett (1993), with the interest rate, \( r \), clearing the market:

\[ B^S \equiv \int adG(a, h) = 0 \] (12)

Given the assumptions made above, a no-trade equilibrium prevails, in which all agents hold zero wealth, hence bond market clearing holds trivially. I show in the Appendix that the interest rate that clears the market is

\[ r = \rho + \lambda + \gamma(1-\tau_1)\mu_w - \frac{1}{2}\gamma(1-\tau_1)(1+\gamma(1-\tau_1))\sigma_w^2 \] (= \( \kappa \)).

**Government.** The government receives tax revenue from labour taxes, and redistributes it as subsidies for students, \( T + F \), per-student grants for educational institutions, \( G \), and potentially also as transfers to low-income workers if taxes are progressive. There is no other government spending or borrowing, so the government budget constraint is as follows:

\[ \text{Tax revenue} \equiv \int_0^\infty \left( whnh - \tau_0(wnh)^{1-\tau_1} \right) g^w(h)dh = S^D(T+F+G) \equiv \text{Student transfers} \] (13)

where the level of taxation, \( \tau_0 \), balances this equation.

**Social welfare.** Welfare in the steady state of this economy is given by the utilitarian social welfare function:

\[ W \equiv \int_0^{h^*} V^s(h)g^s(h)dh + \int_0^\infty V^w(h)g^w(h)dh \] (14)
Stationary equilibrium and social optimum of the Baseline model. A stationary equilibrium of the Baseline model is a set of value functions \( \{ V_s(h), V_w(h) \} \), policy functions \( \{ c_s, c_w(h), n, h^*, S \} \), distributions \( \{ g^s(h), g^w(h) \} \), prices \( (w, F, r) \) and government policies \( (\tau_0, \tau_1, T, G) \) such that:

- Students (2 & 3) and workers (1) maximise their lifetime utility given prices and policies;
- Firms (10) and educational institutions (8) maximise profits given prices and policies;
- The labour market (11), schooling market (9), bond market (12) and goods market all clear and the government budget constraint holds (13); and
- The distributions of the human capital found from the Kolmogorov forward equations of students (6) and workers (7) are both stationary over time.

Since I assumed constant returns to scale production, the wage \( w \) and tuition fee \( F \) are pinned down by technology; the real interest rate \( r \) that clears the market and attains the no-trade equilibrium can be found from the workers’ optimality conditions; the choice of consumption is hence trivial – all agents consume their (post-tax) income; the choice of hours, \( n \), is constant and depends only on the progressivity of taxes and preferences; and \( S \) is isomorphic with \( h^* \).

We thus require just two equilibrium equations in two unknowns, \( h^* (\tau_1, T, G) \) and \( \tau_0 (\tau_1, T, G) \), for which the value matching/smooth pasting/student HJB equation and the government budget constraint shall suffice.

To my knowledge, the equilibrium \( h^* \) and \( \tau_0 \) are unique for a given government policy \( (\tau_1, T, G) \) (recall \( \tau_0 \) is pinned down by the government budget constraint). I assume these policy parameters are chosen ex-ante. I search over this parameter space to find the social optimum, according to the welfare criteria above (14).

Analytical special case: \( \sigma_s = 0 \)

Before discussing the results of the full Baseline model, I first consider a special case of the model above that yields fruitful analytical results where the human capital of students
is assumed to be deterministic, so $\sigma_s = 0$. With this assumption we can derive analytical expressions for the value functions of students as well as workers, the distributions of human capital across workers and students, as well as equilibrium prices and taxes and aggregate social welfare.

In particular, we can now use the value matching and smooth pasting conditions to derive an implicit expression for the graduation threshold, $h^*$. The function $E(h)$ in the value function of the students is still unknown, but since we know that at the threshold the level and slope of the students’ value function is the same as that of the workers’ – which we know – we can derive an implicit expression for the graduation threshold. Note we could only do this when $\sigma_s = 0$ because the smooth pasting and value matching conditions do not provide any information about the second derivative of the value function, which appears in the students’ HJB equation when $\sigma_s > 0$. This will serve as our first equation describing the equilibrium.

**Lemma 2.** The optimal graduation threshold, $h^*$, is determined by the following implicit equation:

$$
\frac{1}{\kappa} \left( \frac{(\tau_0 (wh)^{1-\gamma})^{1-\gamma} v(n)}{1-\gamma} \right)_{V^w(h^*)} = \frac{1}{\rho + \lambda} \left[ \frac{T^{1-\gamma} + 1 - \tau_1}{\kappa h^*} \left( \tau_0 (wh)^{1-\gamma} v(n) \left( \mu_s h^* - \delta h^* \right) \right) \right]_{V^s(h^*)}
$$

(15)

**Proof.** Use the value matching and smooth pasting conditions to plug the workers’ value function, evaluated at the optimal threshold, into the students’ HJB equation. For the case when $\gamma = 1$, see the Appendix.

Since the student’s problem is now deterministic, the rate at which they accumulate human capital is now given by $\dot{h} = \mu_s h^\beta - \delta h$, which is a Bernoulli differential equation with a closed-form solution: $h(t) = \left[ h_0^{1-\beta} e^{-\delta(1-\beta)t} + \frac{\mu_s}{\delta} \left( 1 - e^{-\delta(1-\beta)t} \right) \right]^{1-\beta}$. Thus using $h^* = h(S)$, we can find a closed-form expression for the optimal stopping time, which is now a constant rather than a distribution as it is in the stochastic case, as a function of $h^*$: $S(h^*) = -\frac{1}{\delta(1-\beta)} \ln \left( \frac{h_0^{1-\beta} e^{-\delta(1-\beta)t} + \frac{\mu_s}{\delta}}{h_0^{1-\beta} - \frac{\mu_s}{\delta}} \right)$.

Going back to the original definition of the value function as the expectation of discounted future utility of consumption, we can find a semi-closed-form expression for the students’
value function, in terms of the student’s current human capital and the threshold level:

**Lemma 3.** The students’ value function can be described in terms of their current human capital, $h$, and the graduation threshold, $h^*$:

$$V_s(h) = \frac{u(T)}{\rho + \lambda} + \left( \frac{h^{1-\beta} - \frac{\mu_s}{\delta}}{h^{1-\beta} - \frac{\mu_s}{\delta}} \right)^{-\frac{\rho + \lambda}{\delta(1-\beta)}} \left[ V_w(h^*) - \frac{u(T)}{\rho + \lambda} \right]$$

(16)

**Proof.** Use the integral definition of $V_s(h(t)) = \int_t^S e^{-(\rho + \lambda)(\tau-t)} u(T) d\tau + e^{-(\rho + \lambda)(S-t)} V_w(h^*)$; evaluate the integral and plug in the values of $h(t)$ and $S(h^*)$ found above.

As explained before, the value of being a student comprises two terms: the first the value of being a student forever, the second the option value of graduating, previously denoted $E(h)$.

**The distributions of human capital.** Since workers’ human capital evolves according to a geometric Brownian motion, its stationary distribution, found from equation 7, has the form of a double-Pareto distribution split around the injection point, the cut-off level of human capital, $h^*$, as discussed in Gabaix et al. (2016).

**Lemma 4.** The density of human capital of workers is:

$$g^w(h) = \begin{cases} c \left( \frac{h}{h^*} \right)^{-\zeta_+ - 1} & \text{for } h < h^* \\ c \left( \frac{h}{h^*} \right)^{-\zeta_- - 1} & \text{for } h > h^* \end{cases}$$

where $\zeta_-$ and $\zeta_+$ are the roots of $\frac{1}{2} \sigma_w^2 \zeta^2 + \left( \mu_w - \frac{1}{2} \sigma_w^2 \right) \zeta - \lambda = 0$ and where $c = \frac{\zeta_+ \zeta_-}{(\zeta_- - \zeta_+)(h^*)^2} \left( \frac{h^{1-\beta} - \frac{\mu_s}{\delta}}{h^{1-\beta} - \frac{\mu_s}{\delta}} \right)^{\frac{\lambda}{\delta(1-\beta)}} \frac{1}{\lambda^2\mu_s^2}.$

The density of human capital of students is:

$$g^s(h) = \frac{\lambda}{\mu_s h^{\beta_s - \delta_s h}} \left( \frac{h^{1-\beta_s} - \frac{\mu_s}{\delta_s}}{h^{1-\beta_s} - \frac{\mu_s}{\delta_s}} \right)^{\frac{1}{\delta_s(1-\beta_s)}} h \in [h_0, h^*)$$

**Proof.** That the density of a geometric Brownian motion with a stabilising force (here: death)
follows a double Pareto distribution is well-known in the literature, e.g. Gabaix et al. (2016), and can be derived by guessing that the density has the form of a Pareto distribution, \( ch^{-\zeta - 1} \), and evaluating the stationary Kolmogorov forward equation. That will yield the quadratic equation above, verifying the guess and determining the tail parameters of the two Pareto distributions, which are the roots of the quadratic.

The constant \( c \) is pinned down by the condition that the sum of the two distributions (i.e. the total measure of all workers and students) equals 1: \( \int_0^\infty g^s(h) + g^w(h)dh = 1. \) Moreover, the total measure of workers is simply those who live long enough for their human capital to reach the graduation threshold, \( h^* \), which is exponentially distributed due to the Poisson death rate. Thus the total measure of workers is \( e^{-\lambda S} \) and that of students \( 1 - e^{-\lambda S} \):

\[
\int_0^\infty g^w(h)dh = e^{-\lambda S} = \left( \frac{h^{s1-\beta} - \frac{\mu_s}{\delta_s}}{h_0^{1-\beta} - \frac{\mu_s}{\delta_s}} \right) \frac{\lambda}{\delta_s(1-\beta)}
\]

Hence we can find \( c \) by evaluating the integral on the left over the two parts of the density and setting it equal to the expression on the right.

The density of students over human capital can also be found by a change of variable on the density of students of age \( t \), which is exponentially distributed, \( \lambda e^{-\lambda t} \). This density can be checked by substituting it into the students’ Kolmogorov Forward equation.

The value functions and stationary distributions of student and workers for this special case are plotted below:
With the total measure of students and the stationary distribution of workers’ human capital pinned down, we can also now evaluate the government budget constraint explicitly, yielding an expression for $\tau_0$, which serves as our second equilibrium equation:

**Lemma 5.** The parameter governing the level of taxes, $\tau_0$, is determined by the following implicit equation:

$$
ch^* \left\{ \frac{1}{1 - \zeta^-} - \frac{1}{1 - \zeta^+} \right\} - \tau_0(wh^*)^{1-\tau_1} \left\{ \frac{1}{1 - \tau_1 - \zeta^-} - \frac{1}{1 - \tau_1 - \zeta^+} \right\} = (1 - S^D) (T + F + G)
$$

where $c = \frac{\zeta_1 \zeta_-}{(\zeta^- \zeta^+)}$ and $S^D = 1 - \left( \frac{h^*}{h_0} \right)^{\frac{1-\beta}{\mu_x}} \frac{\mu_y}{\mu_x} \frac{\lambda^*}{\lambda (1-\beta_x)}$.

**Proof.** Integrate on the left-hand side of the government’s budget constraint (13) over the density of workers, and integrate the right-hand side over the density of students.

**Calibration**

In the Baseline model, I calibrate the parameters as follows:
Table 1: Calibration of structural parameters in Baseline model

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Parameter</th>
<th>Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho$</td>
<td>Discount rate</td>
<td>0.01</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>Mortality rate</td>
<td>$1/45$</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>Coefficient of relative risk aversion</td>
<td>2</td>
</tr>
<tr>
<td>$\varphi$</td>
<td>Frisch elasticity of labour supply</td>
<td>0.5</td>
</tr>
<tr>
<td>$\psi$</td>
<td>Labour disutility shifter</td>
<td>1</td>
</tr>
<tr>
<td>$\mu_w$</td>
<td>Drift of human capital while worker</td>
<td>0.0125</td>
</tr>
<tr>
<td>$\mu_s$</td>
<td>Drift of human capital while student</td>
<td>0.3</td>
</tr>
<tr>
<td>$\beta_s$</td>
<td>Decreasing returns to learning in school</td>
<td>0.2</td>
</tr>
<tr>
<td>$\delta_s$</td>
<td>Depreciation of human capital while student</td>
<td>0.18</td>
</tr>
<tr>
<td>$\sigma_w$</td>
<td>Volatility of human capital while worker</td>
<td>0.1</td>
</tr>
<tr>
<td>$\sigma_s$</td>
<td>Volatility of human capital while student</td>
<td>0</td>
</tr>
<tr>
<td>$A^F$</td>
<td>Firm productivity</td>
<td>1</td>
</tr>
<tr>
<td>$A^E$</td>
<td>Education productivity</td>
<td>1</td>
</tr>
</tbody>
</table>

My calibrations of the discount rate and coefficient of relative risk aversion are relatively standard. I choose a mortality rate of $1/45$, to match the average working life. I set the Frisch labour supply elasticity to 0.5, in common with the microeconometric literature. I set the labour shifter to $\psi = 1$, which normalises labour supply to 1 when taxes are proportional, even with positive elasticity, and it is also the value it takes in the Cobb-Douglas limit when $\gamma = 2$, since $\psi = -\frac{1}{1-\gamma} = 1$. I normalise productivity of the representative firm and educational institutions to 1 for simplicity. This results in a tuition fee equal to the wage of an “unskilled” full-time worker – meaning one with the initial level of human capital, $h_0$ – which is at least the right order of magnitude in the US, for example.

The parameters of the workers’ stochastic process, $\mu_w$ and $\sigma_w$, combined with the mortality rate, $\lambda$, pin down the Pareto tails of the pre-tax income distribution. I therefore calibrate these parameters to match the right-tail of the US income distribution, which has $\zeta_+ \simeq 1.5$, and one additional moment, the estimated growth in the cross-sectional dispersion of log income over the life cycle ($\text{Var}(\ln(h_t) - \ln(h^*)) = \sigma_w^2(t-S)$), which grows by around 1-2% per year, according to Deaton & Paxson (1994), implying $\sigma_w = 0.1 - 0.14$. The lower bound $\sigma_w = 0.1$ is consistent with $\zeta_+ \simeq 1.5$ when $\mu_w = 0.0125$, so I use these values in my calibration. This
implies expected growth in incomes while working, \( E_S [\ln(h_t) - \ln(h^*)] = (\mu_w - \frac{1}{2} \sigma_w^2) (t - S) \), of 0.75% per year. In the data, incomes tend to grow faster earlier in the working life, and plateau around middle age (which is not compatible with a stochastic growth process like a geometric Brownian motion) and roughly double over the working life, implying a constant growth rate of around 1.5% per year, or \( \mu_w = 0.02 \) when \( \sigma_w = 0.1 \). However, given the nature of the coupled exponential ageing process and stochastic income growth process, which leads to a counterfactually large number of very old people with high incomes, a lower growth rate seems the most reasonable trade-off to match the cross-sectional moments. Regardless, I also compare results when \( \mu_w = 0 \) and \( \sigma_w = 0 \). I set \( \sigma_s = 0 \) in order to achieve a fully analytical solution, but explore \( \sigma_s = 0.1 \) in a quantitative extension as a robustness check.

I calibrate the parameters of the students’ deterministic human capital accumulation process, \( \mu_s, \beta_s \) and \( \delta_s \), to the returns to higher education estimated in the literature. These estimates vary wildly by country, cohort and estimation methodology; and of course averaging over an entire population as I must do here obscures vast differences in returns along lines of gender, race, field of study and quality of institution, all which I must abstract from for simplicity but which nevertheless could have implications for the optimal financing of higher education. Nevertheless, estimated returns are approximately 8-10% per year of education on average for a Bachelor’s degree (for example see Blundell et al. 2000, 2001, 2005), and perhaps a few percent at most for each year of a PhD, if not negative (for example see van der Steeg et al. 2014, Britton et al. 2020). I thus choose \( \mu_s = 0.3, \delta_s = 0.18 \) and \( \beta_s = 0.2 \), which deliver around 12% growth in human capital per year initially, rapidly decreasing to around 6% per year after 3 years, or around 9% per year on average for the first 3 years of higher education – a typical Bachelor’s degree – before decreasing to around 2.5% per year after 8 years of study.

Equilibrium equations and welfare

**Proposition 1.** The steady state equilibrium of the analytical special case of the Baseline model when \( \gamma \neq 1 \) is given by the following equations governing the optimal human capital
graduation threshold and the government budget constraint:

\[
\begin{align*}
    h^* & : \frac{1}{\kappa} \left( \tau_0 (\omega h^*)^{1-\tau_1} \right)^{1-\gamma} v(n) = \frac{1}{\rho + \lambda} \left[ \frac{T^{1-\gamma}}{1-\gamma} + \frac{1-\tau_1}{\kappa h^*} (\tau_0 (\omega h^*)^{1-\tau_1})^{1-\gamma} v(n) \left( \mu_s h^* \delta_s - \delta_s h^* \right) \right] \tag{18a}
\end{align*}
\]

\[
\begin{align*}
    \tau_0 & : \omega h^* \left[ \frac{\zeta + \zeta_-}{(1-\zeta_-) (1-\zeta_+)} \right] - \tau_0 (\omega h^*)^{1-\tau_1} \left[ \frac{\zeta + \zeta_-}{(1-\tau_1 - \zeta_-) (1-\tau_1 - \zeta_+)} \right] \right] \left( \frac{h^{1-\beta} - \mu_s}{h_0^{1-\beta} - \mu_s} \right)^{\frac{\lambda}{\delta_s (1-\beta)}} \tag{18b}
\end{align*}
\]

where recall that \( w = A^F \) and \( F = \frac{A^F}{A^x} - G \) due to constant returns to scale production, \( n = \left[ \frac{(1-\tau_0) (1+1/\phi)}{\psi (1+1/\phi) + (1-\tau_0) \psi (1-\gamma)} \right]^{1/\gamma} \).

A corner solution arises when \( h^* < h_0 \), in which case the optimal graduation threshold is \( h_0 = 1 \), and the optimal tax level is \( \tau_0 = \frac{(1-\tau_1 - \zeta_-) (1-\tau_1 - \zeta_+) (\omega n)^{1-\tau_1}}{(1-\zeta_-) (1-\zeta_+)} \).

Proof. For the proof and for the same system of equations when \( \gamma = 1 \), see the Appendix.

In equilibrium, higher transfers typically lead to a higher graduation threshold, \( h^* \), which implies higher average human capital of workers, but fewer workers. In most of the region in which an equilibrium exists, higher transfers lead to higher \( \tau_0 \), and so higher take-home pay for a given level of human capital. However, equilibrium \( \tau_0 \) eventually decreases for very high \( T \), as the revenue demands of extra schooling outweigh the human capital gains and average tax rates therefore must increase.

An equilibrium may not exist for very high \( T \), since it becomes optimal to remain a student forever. To see this, take the simple case where workers retain the \( h \) they graduate with until they die (\( \mu_w = \sigma_w = 0 \)); ignoring taxes and assuming exogenous labour supply (normalised to 1), it is optimal never to graduate whenever \( T \geq h^* \) (since \( w = 1 \)), so this cannot be an equilibrium. Due to decreasing returns to learning, \( h^* \) is bounded above by \( \bar{h} \), which in my calibration is around 1.9, so clearly \( T \geq 1.9 \) cannot be an equilibrium; the presence of growth in human capital while working, risk, taxes and endogenous labour supply mean in fact \( T \)
must be somewhat less than 1.9 in order for an equilibrium to exist.

![Figure 3: Equilibrium graduation threshold, $h^*$, and tax level, $\tau_0$](image)

**Corollary 1.** Aggregate social welfare in the analytical special case of the Baseline model is given by:

\[
W^s = \frac{T^{1-\gamma} - 1}{(\rho + \lambda)(1 - \gamma)} \left[ 1 - \left( \frac{h^*}{h_0} \right)^{\frac{\lambda}{\delta_s}} \right] \left[ \frac{1}{\kappa} \left( \frac{\tau_0/\rho}{1 - \gamma} \right) - \frac{T^{1-\gamma}}{(\rho + \lambda)(1 - \gamma)} \right] \frac{\lambda}{\rho} \left( \frac{h^*}{h_0} \right)^{\frac{\lambda}{\delta_s}} \frac{\lambda}{\delta_s} \left[ 1 - \left( \frac{h^*}{h_0} \right) \right]^{\frac{\lambda}{\delta_s}}
\]

\[
W^w = \frac{1}{\kappa} \left( \frac{\tau_0/\rho}{1 - \gamma} \right) \left[ \frac{1}{\kappa} \left( \frac{\tau_0/\rho}{1 - \gamma} \right) \right] \frac{Z}{\kappa} \left( \frac{h^*}{h_0} \right)^{\frac{\lambda}{\delta_s}} \frac{\lambda}{\delta_s} \left[ 1 - \left( \frac{h^*}{h_0} \right) \right]^{\frac{\lambda}{\delta_s}}
\]

\[
W = W^s + W^w
\]

where \( Z = \frac{\zeta + \zeta'}{[(1-\tau_1)(1-\gamma)-\zeta][(1-\tau_1)(1-\gamma)-\zeta']} \)

**Proof.** See the Appendix.
Students’ welfare can be decomposed into a sum of two terms: the lifetime utility from being a student forever, multiplied by the measure of students; plus the discounted option value of graduating, averaged over all students.

Workers’ welfare can be decomposed into a product of three terms: the welfare of a graduate, which is also the welfare of the modal worker, since the distribution of workers is split around that point; the measure of workers; and the term $Z$, which accounts for the inequality between workers, i.e. the relative measures of workers either side of the mode. This last term converges to 1 when the distribution of workers collapses to a point ($\zeta_+, \zeta_- \to \infty$, when $\mu_w = \sigma_w = 0$) or when taxes are totally progressive ($\tau_1 = 1$), in which case there is no inequality between workers. The calibration here ($\zeta_+ \simeq 1.5$, $\zeta_- \simeq -3$) is such that, in the entirety of the policy space explored (i.e. without very regressive taxes, $\tau_1 \ll 0$, or totally progressive taxes $\tau_1 = 1$) we have that $Z < 1$, so inequality implies the welfare of the average worker, and hence the aggregate, is below that of the modal worker.

Figure 4: Welfare by tax progressivity and level of student transfers in the Baseline model

Results: Baseline model. The key results from this benchmark case are:

1. The level of university grants, $G$, has no effect on the equilibrium; tuition fees fall 1:1,
so government expenditure and hence equilibrium tax rates are unchanged.

2. If tuition fees, $F$, are capped by the government, but $G$ is appropriately set to balance the supply and demand for schooling – i.e. such that $F + G = \frac{A^E}{\lambda^r}$ – the equilibrium is also unchanged.

Thus, financing higher education through tax-funded university grants or through privately-funded tuition aided by tax-funded student transfers are equivalent. Of course, students require some level of transfers to pay for consumption.

3. Social welfare is maximised with highly progressive labour taxes, $\tau_1 \gg 0$, and relatively high net student transfers, $T$. Typically, for a given progressivity of taxes, higher transfers improve welfare up to a point. As transfers increase, the optimal progressivity of taxes also increases, up to the global maximum (the red dot).

4. However, when net student transfers are very low, the equilibrium is degenerate, with no students. In this case, welfare is independent of transfers. The optimal degree of progressivity in this region is similar to the global maximum. This region is larger the more progressive are taxes.

5. Moreover, when high student transfers are very high, an equilibrium may not exist, as it becomes optimal never to graduate. Extremely progressive taxes and extremely regressive taxes both reduce the maximum student transfers the government can offer.

**Welfare of students vs workers.** Transfers to students naturally have no direct effect on the welfare of workers, so only affect it through changing the equilibrium $h^*$ and $\tau_0$. These have no effect on hours worked, only consumption and the measure of workers.
**Insurance vs redistribution.** Recall there are three distinct motivations for policy intervention in this economy: social insurance against risk; redistribution; and financing of higher education (by redistributing to students). We can disentangle these competing motives by setting $\sigma_w = 0$, to shut down the insurance motive, as shown in Figure 6. As one might expect, when there is no risk – and hence no inequality among workers of a given age – there is no benefit from social insurance and less benefit from redistribution, so optimal taxes are distinctly less progressive, though notably still more so than the current US schedule.
We can shut down the redistribution motive entirely by setting $\mu_w = 0$, so workers no longer accumulate human capital on the job and thus there is no income inequality at all between workers. The only “redistributive” goal that remains is towards students, i.e. in funding their higher education and thus overcoming the financial friction (that they cannot borrow). In this case, optimal taxes are somewhat regressive. Note also that eliminating wage growth while working substantially reduces the level of student transfers at which it becomes optimal never to graduate, thus substantially curtailing the region in which an equilibrium can exist.
Figure 7: Welfare with no inequality among workers, $\sigma_w = 0$ & $\mu_w = 0$

Clearly then, while in this model social insurance against risk provides a compelling case for progressive taxation, in common with much of the literature on this topic, not least for example Heathcote et al. (2017), we also see that intergenerational inequality – which that paper does not feature due to its assumed flat life-cycle earnings profile – also motivates progressive taxation. This explains much of the disparity in the results between my paper and their’s.

**Stochastic learning.** We can examine numerically the slightly more general version of the Baseline model where human capital accumulates stochastically while studying, $\sigma_s > 0$, using the finite difference method of Achdou et al. (2022). We can see from Figure 8 that qualitatively and quantitatively the results from the analytical Baseline model go through largely unchanged: optimal transfers are high and optimal taxes are progressive.
Figure 8: Welfare in the Baseline model with risky learning while studying, $\sigma_s = 0.1$

3 Analytical Extensions

Supply side general equilibrium effects

A concern one may have is that expanding access to higher education raises costs and hence tuition fees, and so in general equilibrium subsidies end up costing a great deal and imposing a much greater tax burden. The Baseline model neuters this effect by assuming constant returns to scale production, so that policy has no effect on relative prices. I relax this assumption by instead assuming that production of both consumption goods and schooling faces decreasing returns to scale, $f_i(L_i) = A_i L_i^{1-\alpha_i}$. One can ascribe this for example to the presence of some fixed factor of production. For simplicity, I assume that any profits or payments to this fixed factor go to overseas owners who do not participate in the economy in any other way. I moreover assume here that both sectors face the same degree of decreasing returns: $\alpha_E = \alpha_F$.

In the Appendix I provide the two additional equations that determine the equilibrium; labour market and schooling market clearing conditions.

Equilibrium wages for the most part move inversely with effective labour supply: as
the progressivity of taxes increases, all agents reduce their labour supply, raising the wage; however, as the generosity of student transfers increases, schooling increases, raising the aggregate effective labour supply and lowering the piece-rate wage. Tuition fees naturally increase with the number of students.

Figure 9: Equilibrium tuition fees, $F$, and wages, $w$, with DRS production

The optimal tax schedule remains highly progressive, compared to the current US system. As for optimal student transfers, from the scale of the diagram below, introducing DRS production appears to substantially reduce the maximum transfers that the government can offer students before an equilibrium ceases to exist. However, this is partly illusory: with DRS production both the piece-rate wage and the equilibrium tuition fee are typically far less than 1 ($= A^F$ and $= \frac{A^F}{A^E}$), as they were in the Baseline model. Consequently, to interpret the values of $T$ on the x-axis correctly, and compare them to the Baseline case, one should scale them by one of either $w$ or $F$; in fact then the optimal student transfers in the model with decreasing returns to scale production are somewhat more generous than in the Baseline model. Thus, the intuition that some may have, that subsidising education induces higher tuition fees and hence offsets any welfare gains from the subsidies, appears to be misguided, at least in this model.
Ex-ante heterogeneity

The Baseline model ignores apparent heterogeneity, for example in ability and hence in the returns to education, and thus may overstate the gains from flat subsidies to all students. In the Appendix I discuss in more depth a very general analysis allowing for ex-ante heterogeneity in learning ability and disutility from labour. For now, I focus on heterogeneous ability. In particular, I envision “learning ability” as a parameter, $z$, that augments ($z > 1$) or diminishes ($z < 1$) the speed with which an agent accumulates human capital, either while studying or working. Human capital thus accumulates like $\dot{h}_t = \mu_s z h_t^{\beta_s} - \delta_s h_t$ while a student, and $dh_t = \mu_w z^\alpha h_t dt + \sigma_w h_t dW_t$ while working, where $\alpha \in [0, 1]$ parameterises how strongly school-learning ability translates into on-the-job learning ability.

Here I focus on a special case, with just two types of agents who differ only in their ability to learn while in school (so with $\alpha = 0$): one with normal/high ability $z_1 = 1$ – as in the Baseline model – and one with low ability $z_0 = \bar{z}$, where $\bar{z}$ is low enough that these agents will never attend university, i.e. $S(\bar{z}) = 0$ or $h^*(\bar{z}) = h_0$. For sake of argument, we can assume $\bar{z} = 0$, so such students are incapable of learning while at university, and where hence those
agents will never attend university, regardless of how generous tuition subsidies are. For the avoidance of any foundational mathematical ambiguity, let’s regard \( Z^0 = 1 \), so such agents learn on-the-job at the standard rate.

This set-up is convenient as it allows us to incorporate the fact that only around half of students go to university, so let’s assume half of agents are type \( Z_1 \) and half \( Z_0 \). Of course, with \( Z = 0 \), the fraction who attend university will be either half or zero, depending on how generous student transfers are, but one can also generalise the set-up to explain the increasing share of students attending university.

The effect of introducing low-skilled learners is to broaden slightly the range of possible equilibria. Since a large fraction of people never attend university, they need no subsidies, so a given level of subsidies for those that do attend university requires lower average taxes, \( \tau_0 \), meaning the value of working is higher, thus raising the level of student transfers at which it becomes optimal never to graduate, and at which an equilibrium therefore cannot exist. In addition, optimal taxes are more progressive, although in general the effect of introducing this particular form of ex-ante heterogeneity is small.

Figure 11: Welfare with ex-ante heterogeneous ability while studying

\[ \text{Figure 11: Welfare with ex-ante heterogeneous ability while studying} \]

\[ \begin{array}{c}
\text{Welfare, } W \\
\text{Tax progressivity, } \tau_1 \\
\text{Transfer, } T \\
\end{array} \]
Ageing

The Baseline model captures the time opportunity cost of schooling only in a very crude way: the perpetual youth structure means spending an extra year in school has a uniform time cost, in the form of a higher discount rate, since after the extra year one still has the same life expectancy as before ($1/\lambda$). I thus consider a version of the model where agents live a certain and finite number of years, $Y$, instead of dying at a constant rate, so spending an extra year in school means one has one year less to live and hence earn, limiting the value of an extra year in education to accumulate a marginal amount of additional human capital. Such an alteration likely better reflects the opportunity cost of schooling.

This change substantially reduces the average welfare gains to workers from higher transfers to students, $T$, simply because once the students actually graduate they have fewer years of life to enjoy their higher earnings. Optimal student transfers are therefore substantially lower than in the Baseline model. Optimal taxes are also significantly less progressive, likely because there is a much smaller right-tail of workers with very high incomes. As is well-known, the income distribution of a perpetual youth model with permanent income shocks is Pareto-distributed because there is a tail of workers who live inordinately long and so accumulate enormous human capital; capping their lifespan cuts off this fat right tail. Thus the government raises less tax revenue from the super-rich, and this necessitates higher average taxes (lower $\tau_0$) from the bulk of earners, lowering the welfare gains from progressive taxes. It should be pointed out however, that the income distribution of course does have a thick right tail in the data, it simply arises by other means not captured by the model (e.g. see Gabaix et al. 2016), so this version of the model may understate the welfare gains from progressive taxation. In any case, optimal taxes are still more progressive than the current US system.
4 Quantitative Model: Introducing Student Debt

I now make one subtle change over the Baseline model: students can borrow. I set a borrowing limit of 5\times the tuition fee. In addition to any borrowings, students still receive a gross transfer \( \bar{T} = T + F \) from the government. This small change means the “no trade” result is no longer an equilibrium, and individuals’ policy functions no longer have closed-form solutions. I thus proceed entirely numerically.

Introducing student loans – and hence necessitating that workers build wealth, both to pay off their debts and to ensure bond markets clear – can change individual behaviour substantially. Households with low-to-moderate wealth and income behave similarly to the no-trade equilibrium, but those with low wealth and high income save considerable sums – on the order of 30-40\% of their income – in part to insure themselves against income risk, while those with high wealth and low income dissave. The new equilibrium also evidences “precautionary working”, where low-wealth, high-income households work longer hours – again partly to insure themselves – while wealthy households work less, owing to standard wealth
effects.

Figure 13: Consumption, labour & saving in the Quantitative model vs Baseline no trade

Since both the workers’ and students’ problems are otherwise the same as before, I do not recapitulate them here. For completeness I show below the Kolmogorov Forward equations that describe the evolution of the joint distributions of human capital and wealth for both students and workers. The stationary distribution is found by setting them equal to zero (here suppressing function notation):

\[
\begin{align*}
\dot{g}^s &= -\partial_a [\dot{a} g^s] - \partial_h \left[ (\mu_s h \delta_s - \delta_s h) g^s \right] - \lambda g^s + \lambda \delta(a) \delta(h - h_0) = 0 \quad (a^*, 0] \times [h_0, h^*) \\
\dot{g}^w &= -\partial_a [\dot{a} g^w] - \partial_h [ \mu_w h g^w] + \frac{1}{2} \partial_{hh} \left[ \sigma_w^2 h^2 g^w \right] - \lambda g^w = 0 \quad [a, \infty) \times (0, \infty) \setminus (a^*, h^*)
\end{align*}
\]

Compared to the Baseline model there is one additional term in each equation; the first, representing the change in the measures of students and workers as they save. I have also included the term in the first equation that represents inflows of newborn students at the point \((0, h_0)\); newborns inherit zero wealth and start with human capital \(h_0\), which is captured by the Dirac measures at \(h = h_0\) and \(a = 0\) in the final term.

The supply side and market clearing conditions and social welfare functions are again
largely unchanged from the Baseline model, so I do not recapitulate them all here.

**Stationary equilibrium of the Quantitative model.** A stationary equilibrium of the Quantitative model is a set of value functions \( \{V^s(a, h), V^w(a, h)\} \), policy functions \( \{c^s(a, h), c^w(a, h), n(a, h), h^s(a^*)\} \), distributions \( \{g^s(a, h), g^w(a, h)\} \), prices \( (w, F, r) \) and government policies \( (\tau_0, \tau_1, T) \) such that students & workers maximise lifetime utility and firms & educational institutions maximise profits given prices and policies; all markets clear and the government budget constraint holds; and the distributions of the human capital are stationary over time. \( w \) and \( F \) are pinned down by technology, and \( \tau_0 \) is restricted by the government budget constraint, so the triple \( (h^*(a^*), \tau_0, r) \) completely describes an equilibrium for a given policy \( (\tau_1, T) \). I search over \( (\tau_1, T) \) to find the policy that maximises social welfare.

One key difference in the resulting equilibria of the Quantitative model compared to the Baseline is that the region of the policy space with no students is now much smaller, even when \( T \) is small but positive. There are numerical challenges to solving the model when \( T \leq 0 \) (e.g. when \( \tilde{T} = 0 \) and so students receive no (gross) transfer), since those who are at the borrowing constraint and hence unable to borrow more would be unable to consume. Economically, it is clear how to resolve this: students who have borrowed up to the limit and who can hence no longer afford to pay tuition fees would “graduate” immediately and start working. However, it turns out that the region of the policy space that (at least locally) maximises social welfare features positive transfers, so the problem is moot.

**Results: Quantitative model.** The optimal tax and transfer schedule is qualitatively and quantitatively similar to the Baseline model and its myriad extensions. Optimal taxes are highly progressive, and only modestly less so than in prior iterations of the model, with optimal \( \tau_1 \simeq 0.5 \). Perhaps surprisingly, the optimal transfer, for essentially any degree of tax progressivity, is overwhelming positive, albeit smaller than in the Baseline model. Recall moreover that \( T \) is the net transfer, in excess of that needed to cover tuition fees, even with student loans available. Naturally, this is far above the levels of grants available to most
Should I Stay (in School) or Should I Go (to Work) 

Lee Tyrrell-Hendry 

prospective students in most countries.

Figure 14: Welfare in Quantitative model

5 Conclusion

In this paper I have explored the question of how to calibrate income tax progressivity and higher education subsidies to balance the welfare gains from social insurance, overcoming financial frictions and redistribution, with the welfare losses from disincentivising work and study that such policies can bring about. To do so while overcoming some of the shortcomings of the prior literature, I set out an incomplete markets model with an education choice modelled as a stopping time problem, with endogenous labour supply and saving choices. I have shown that in such a model, there are substantial welfare gains from both highly progressive taxation of labour income and high subsidies for higher education, compared to a baseline of no intervention.

Optimal policy in the model features more progressive taxes than currently observed in the US or Europe, as well as large subsidies, even when students can borrow to pay for their education. I have further shown that these welfare conclusions are robust to myriad extensions.
of the model, including diseconomies from scale in education, where expanding access to education makes it more costly; heterogeneity in the returns to schooling, i.e. in ability; a more realistic life-cycle profile; and less severe financial frictions, where student can borrow to some extent, which also features a more realistic equilibrium with extensive precautionary saving by workers.

Notably, the form of subsidy appears not to be particularly important for welfare: I have shown that direct transfers to students, who then pay tuition fees set freely by the market, are equivalent to per-student government grants to universities, even when governments explicitly cap tuition fees, provided the grant is sufficient to balance supply and demand for schooling and avoid rationing of education. Financing education with student loans also does not change the optimality of highly progressive taxation and substantial subsidies for education.

Naturally, further research ought to consider additional channels by which such policies may have adverse effects, for example a micro-founded model of on-the-job human capital accumulation that arises from optimising behaviour, where obviously incentives to accumulate human capital are crucial, but which in this paper is modelled as an exogenous ad hoc process. Moreover, one might consider incorporating other features of the education decision, such as the field of study, or the quality of institution, or calibrating the financing structure to more closely match real-world policies. In particular, an interesting extension, which I am exploring in further work, is to allow for rationing of education when tuition fees are capped by the government, hence breaking the equivalence between university grants and student transfers in the present model.

Regardless, I believe the model in this paper offers a compelling case for recalibrating tax and education finance policy.
References


Mellior, G. (2021), Higher education funding, welfare and inequality in equilibrium.


Proof of Theorem 1: Existence of no-trade equilibrium.

Proof. To prove the existence of a no-trade equilibrium, I take the same approach as Heathcote et al. (2017) in showing that the policy functions described in the body of the text satisfy all the consumers’ optimality conditions and give rise to aggregate outcomes that satisfy all market clearing conditions. First, recall the workers’ HJB equation:

\[
(\rho + \lambda) V^w = \max_{c,n} \left( \frac{c^{1-\gamma} v(n) - 1}{1 - \gamma} + V^w_a (ra + \tau_0 (wh)^{1-\tau_1} - c) + V^w_h \mu_w h + \frac{1}{2} V^w_{hh} \sigma_w^2 h^2 \right)
\]

I conjecture that workers consume their post-tax income, \( c = \tau_0 (wh)^{1-\tau_1} \), and supply constant labour of \( n = \left[ \frac{(1-\tau_1)(1+\psi)}{\psi(1+1/\psi)+(1-\tau_1)\psi(1-\gamma)} \right]^{\frac{1}{1+1/\psi}} \), the latter derived in the text from the workers’ labour optimality condition assuming the former holds and workers all hold zero wealth. The workers first-order conditions are as follows:

\[
c^{-\gamma} v(n) = V^w_a
\]

\[
- \frac{c^{1-\gamma}}{1 - \gamma} v'(n) = V^w_a \tau_0 (1 - \tau_1)(wh)^{1-\tau_1} n^{-\tau_1}
\]

Combining these two expressions and substituting in the assumed policy function for consumption reveals that labour supply is indeed given by the prior expression.

Secondly, applying the envelope condition to the workers’ HJB equation yields:

\[
(\rho + \lambda) V^w_a = V^w_{aa} \dot{a} + V^w_a r + V^w_{ha} \mu_w h + \frac{1}{2} V^w_{haa} \sigma_w^2 h^2
\]

where the first term on the right-hand side is zero under the assumed policy functions, provided wealth is also zero. Recognising by Young’s Theorem that \( V_{ha} = V_{ah} \) and \( V_{hha} = V_{ahh} \), and hence differentiating the consumption first-order condition (again under the assumed policy
function), we find:

$$V_{ah}^w = -\gamma c(h)^{-\gamma - 1} c'(h) v(n) \quad V_{ab}^w = -\gamma c(h)^{-\gamma - 1} c''(h) v(n) + \gamma (1 + \gamma) v(n) c(h)^{-\gamma - 2} (c'(h))^2$$

Substituting these into the envelope condition yields:

$$(\rho + \lambda - r) V_a^w = -\gamma c^{-\gamma - 1} c'(h) v(n) \mu_w h + \frac{1}{2} \left( -\gamma c^{-\gamma - 1} c''(h) v(n) + \gamma (1 + \gamma) v(n) c^{-\gamma - 2} (c'(h))^2 \right) \sigma_w^2 h^2$$

where $c(h) = \tau_0 (wnh)^{1-\tau_1}$, $c'(h) = \tau_0 (1 - \tau_1) (wn)^{1-\tau_1} h^{-\tau_1} = c(h) \tau_1 h^{-1}$ and $c''(h) = -\tau_0 (1 - \tau_1) \tau_1 (wn)^{1-\tau_1} h^{-\tau_1-1} = -c(h) (1 - \tau_1) \tau_1 h^{-2}$. Substituting these into the expression for $V_a^w$, and that into the consumption first-order condition yields:

$$(\rho + \lambda - r) c^{-\gamma} v(n) = -\gamma c^{-\gamma} (1 - \tau_1) v(n) \mu_w h + \frac{1}{2} \gamma c^{-\gamma} (1 - \tau_1) \tau_1 v(n) \sigma_w^2 h^2 + \frac{1}{2} \gamma (1 + \gamma) v(n) c^{-\gamma} (1 - \tau_1)^2 \sigma_w^2$$

Making the appropriate cancellations and rearranging reveals that this equation holds when the interest rate prevailing in the market is:

$$r = \rho + \lambda + \gamma (1 - \tau_1) \mu_w - \frac{1}{2} \gamma (1 - \tau_1) (1 + \gamma (1 - \tau_1)) \sigma_w^2$$

as conveyed in the text. To check that this interest rate clears the bond market is trivial since $a = 0$ for all agents. Moreover, individual budget constraints are all respected and imply zero savings provided all workers graduate with zero wealth, as the assumption of no student debt guarantees. The aggregate resource constraint/goods market clearing condition thus clearly holds too. The workers’ value function subsequently derived in the body of the text also satisfies the workers’ HJB equation, again given that under the assumed policy function savings are everywhere zero, thus verifying the guess. Aggregate labour demand and the supply of university places are both infinitely elastic, given the assumption of constant returns to scale production, so will adjust to clear those markets provided prices $(w, F)$ satisfy the supply side optimality conditions. The government budget constraint holds when $\tau_0$ takes
the equilibrium value described in Proposition 1. Thus, no-trade respects all households’ optimality conditions and clears all markets and is hence an equilibrium. □

Proof of Corollary 1.

Proof. Starting from the general expression for aggregate welfare:

\[ W_s^* \equiv \int_0^{h^*} V^s(h) g^s(h) dh \]
\[ W_w^* \equiv \int_0^{h^*} V^w(h) g^w(h) dh \]
\[ W = W_s^* + W_w^* \]

Plugging in the expressions for the workers’ value function (4), worker welfare becomes:

\[ W_w = \int_0^{h^*} \left( \frac{\tau_0(wnh)^{1-\tau}}{\kappa(1-\gamma)} \right)^{1-\gamma} v(n) c \left( \frac{h}{h^*} \right)^{-\zeta_-^{-1}} dh + \int_{h^*}^{\infty} \left( \frac{\tau_0(wnh)^{1-\tau}}{\kappa(1-\gamma)} \right)^{1-\gamma} v(n) c \left( \frac{h}{h^*} \right)^{-\zeta_-^{-1}} dh - AG_w^w \]

\[ = \frac{(\tau_0(wn)^{1-\tau})^{1-\gamma} v(n)}{\kappa(1-\gamma)} c \left( \frac{h}{h^*} \right)^{-\zeta_-^{-1}} dh \left( h^* \int_0^{h^*} h(1-\tau)(1-\gamma)-\zeta_-^{-1} dh + h^* \int_{h^*}^{\infty} h(1-\tau)(1-\gamma)-\zeta_-^{-1} dh \right) - AG_w^w \]

\[ = \frac{(\tau_0(wn)^{1-\tau})^{1-\gamma} v(n)}{\kappa(1-\gamma)} c \left( \frac{h}{h^*} \right)^{-\zeta_-^{-1}} dh \left[ h^* \int_0^{h^*} \frac{1}{(1-\tau)(1-\gamma)-\zeta_-} h(1-\tau)(1-\gamma)-\zeta_-^{1-\gamma} h^* \right] - AG_w^w \]

\[ = \frac{(\tau_0(wnh^*)^{1-\tau})^{1-\gamma} v(n)}{\kappa(1-\gamma)} c \left( \frac{h}{h^*} \right)^{-\zeta_-^{-1}} dh \left[ \frac{1}{(1-\tau)(1-\gamma)-\zeta_-} - \frac{1}{(1-\tau)(1-\gamma)-\zeta_+} \right] - AG_w^w \]

\[ = \left[ \frac{(\tau_0(wnh^*)^{1-\tau})^{1-\gamma} v(n)}{\kappa(1-\gamma)} Z - A \right] \left( \frac{h^* \gamma_\beta - \mu_\gamma}{\gamma_\beta - \mu_\gamma} \right) \frac{\lambda}{\sigma_\gamma(1-\beta_\gamma)} \]

where \( A = \frac{1}{\gamma_\beta + \lambda(1-\gamma)} \), \( Z = \frac{\zeta_- \zeta_+}{(1-\tau)(1-\gamma)-\zeta_-[(1-\tau)(1-\gamma)-\zeta_+]} \), where the last line uses the definition of \( c \) and where \( G_w^w \) represents the total measure of workers, \( G_w^w = \left( \frac{h^* \gamma_\beta - \mu_\gamma}{\gamma_\beta - \mu_\gamma} \right) \frac{\lambda}{\sigma_\gamma(1-\beta_\gamma)} \).

The welfare of students is slightly more taxing to calculate, as we must account not only for the value of being a student, but also the option value of graduating, and the effective discount factor of that option value before the students chose to graduate, averaged across the population of students. Recall that from the definition of the value function when
human capital accumulation during school is deterministic, \( V_s(h) = \int_t^S e^{-(\rho+\lambda)(\tau-t)}u(c^*)d\tau + e^{-(\rho+\lambda)(S-t)}V_w(h^*) \), the integral is:

\[
\int_t^S e^{-(\rho+\lambda)(\tau-t)}u(c^*)d\tau = \left[ -\frac{1}{\rho+\lambda} e^{-(\rho+\lambda)(\tau-t)} \frac{T^{1-\gamma} - 1}{1-\gamma} \right]_t^S
\]

\[
= \frac{T^{1-\gamma} - 1}{(\rho+\lambda)(1-\gamma)} \left( 1 - e^{-(\rho+\lambda)(S-t)} \right)
\]

The first term here represents the value of being a student forever. The second term depends on \( S \), which in turn depends on \( h \), and is the discounted value of remaining a student that one gives up when graduating. Thus we have:

\[
W_s = \int_0^{h^*} \left\{ \frac{u(T)}{\rho+\lambda} + e^{-(\rho+\lambda)(S-t)} \left[ V_w(h^*) - \frac{u(T)}{\rho+\lambda} \right] \right\} g_s(h)dh
\]

\[
= \int_0^{h^*} \left\{ \frac{T^{1-\gamma} - 1}{(\rho+\lambda)(1-\gamma)} + e^{-(\rho+\lambda)(S-t)} \left[ \frac{(\tau_0(wnh^*)^{1-\gamma})^{1-\gamma} v(n)}{\kappa(1-\gamma)} - \frac{T^{1-\gamma}}{(\rho+\lambda)(1-\gamma)} \right] \right\}
\]

\[
\cdot \frac{-\lambda}{\delta_s \left( h_0^{1-\beta_s} - \frac{\mu_s}{\delta_s} \right) \left( h_0^{1-\beta_s} - \frac{\mu_s}{\delta_s} \right)^{\frac{\lambda}{\delta_s(1-\beta_s)^{-1}}} h^{-\beta_s} dh
\]

where in the second line I have plugged in the density and value function at graduation, and adjusted the lower bound of integration to account for the fact that under deterministic human capital accumulation, no student will have less human capital than they start with.

Since students with different levels of human capital are at varying distances from the human capital (and hence graduation) threshold, when discounting their option value of graduating we must account for their current level of human capital, and hence distance from the threshold, by replacing \( e^{-(\rho+\lambda)(S-t)} = \left( \frac{h^{1-\beta_s} - \frac{\mu_s}{\delta_s}}{h_0^{1-\beta_s} - \frac{\mu_s}{\delta_s}} \right)^{\frac{\rho+\lambda}{\delta_s(1-\beta_s)}} \) before integrating over \( h \) using u-substitution with \( u = h^{1-\beta_s} - \frac{\mu_s}{\delta_s} \).
where again the first terms represents the value of being a student forever, multiplied by the measure of students; the second term is the option value of graduating, discounted by a factor inversely proportional to the length of time to graduation for students at different levels of human capital, averaged across all students, i.e. \( \int_{h_0}^{w^*} e^{-(\rho + \lambda)(S-t)} g^s(h) dh \).

When \( \gamma = 1 \), the optimal graduation threshold is determined implicitly by the joint value matching and smooth pasting conditions, shown here again substituted into the students’ value function evaluated at the threshold:

\[
\begin{align*}
\ln \left( \frac{\tau_0 (w h^*)^{1-\gamma} v(n)}{\kappa (1-\gamma)} \right) &= \psi \frac{n^{1+1/\phi}}{1 + 1/\phi} + \frac{1 - \tau_1}{\rho + \lambda} \left( \mu_w - \frac{\sigma_w^2}{2} \right) = \ln(T) + \frac{1 - \tau_1}{\rho + \lambda} \left( \mu_s h^{s-1} - \delta \right)
\end{align*}
\]
Appendix 2: Extensions

Supply side general equilibrium effects

The first extension I explore is to introduce decreasing returns to scale production of goods and education, \( f_i(L_i) = A_i(L_i)^{1-\alpha_i} \), to allow for what one might call more “interesting” supply-side general equilibrium effects. With this, relative prices \( F \) and \( w \) are no longer pinned down by technology, so we need two additional equilibrium equations, for which the schooling and labour market clearing conditions will suffice:

\[
F : \quad A^E \left( \frac{(1 - \alpha_E)FAE}{w} \right)^{\frac{1-\alpha_E}{\alpha_E}} = 1 - \frac{h^*1-\beta - \mu_s}{h_0^{1-\beta} - \mu_s} \left( \frac{\lambda}{\delta s} \right)
\]

\[
w : \quad \left( \frac{(1 - \alpha_F)AF}{w} \right)^{1/\alpha_F} + \left( \frac{(1 - \alpha_E)FAE}{w} \right)^{1/\alpha_E} = \frac{\zeta + \zeta - n h^*}{(1 - \zeta_-(1 - \zeta_+)} \left( \frac{h^*1-\beta - \mu_s}{h_0^{1-\beta} - \mu_s} \right) \left( \frac{\lambda}{\delta s} \right)
\]

Ex-ante heterogeneity

The second extension incorporates ex-ante heterogeneity, namely in ability, \( z_i \), which here will enhance the learning capabilities of students and workers; and labour disutility, represented by the shifter \( \psi_i \), where recall that workers’ disutility from labour is represented by the function \( v_i(n) = (1 - \psi_i(1 - \gamma)\frac{n^{1+\psi}}{1+\psi})^\gamma \). The latter, on top of making some agents more willing to work for a given wage, will also mean they wish to finish studying earlier and enter the workplace, relative to those more distasteful of work, increasing the size of the workforce. I index both these variables with \( i \) to emphasise that they vary between individuals but are nevertheless constant.
The workers’ and students’ problems become:

\[
(p + \lambda) V^w_i(h) = \frac{(\tau_0(w_i h)^{1-\gamma_i})^{1-\gamma_i} v_i(n_i)}{1-\gamma_i} + V^w_i(h) \mu_i z_i h + \frac{1}{2} V^w_i(h) \sigma^2_i \]

\[
(p + \lambda) V^s_i(h) = \frac{T^{1-\gamma_i} - 1}{1-\gamma_i} + V^s_i(h) \left( \mu_i z_i h^{\delta_i} - \delta_i h \right) + \frac{1}{2} V^s_i(h) \sigma^2_i h^2
\]

Workers’ optimal labour is given by: \( n_i = \left( \frac{(1-\tau_i)(1+\varphi)}{\psi_i(1+\varphi) + (1-\tau_i) \psi_i(1-\gamma)} \right) \frac{1}{\varphi_i} \) and their value function by:

\[
V^w_i(h) = \begin{cases} 1 \kappa_i \left( \frac{(\tau_0(w_i h)^{1-\gamma_i})^{1-\gamma_i} v_i(n_i)}{1-\gamma_i} - \frac{1}{(\rho + \lambda)(1-\gamma)} \right) & \text{if } \gamma \neq 1 \\ \frac{1}{\rho + \lambda} \left[ \ln \left( \tau_0(w_i h)^{1-\gamma_i} \right) - \frac{\varphi_i(1+\varphi)}{1+\varphi} \right] + \frac{1}{\varphi_i(\rho + \lambda)} \left( z_i^0 \mu_i - \frac{\sigma_i^2}{2} \right) & \text{if } \gamma = 1 \end{cases}
\]

while that of students is:

\[
V^s_i(h) = \frac{u(T)}{\rho + \lambda} + \left( \frac{h^{1-\beta_i} - \frac{z_i \mu_i}{\delta_i}}{h^{1-\beta_i} - \frac{z_i \mu_i}{\delta_i}} \right)^{\frac{\rho + \lambda}{\delta_i(1-\beta_i)}} \left[ V^w_i(h_i) - \frac{u(T)}{\rho + \lambda} \right]
\]

I assume for simplicity that innate characteristics are discretely distributed according to the probability mass function, \( m_i \). The distributions of human capital across students and workers have the same form as in the Baseline model:

\[
g^w_i(h) = m_i \frac{\lambda}{z_i \mu_i h^{\delta_i} - \delta_i h} \left( \frac{h^{1-\beta_i} - \frac{z_i \mu_i}{\delta_i}}{h^{1-\beta_i} - \frac{z_i \mu_i}{\delta_i}} \right)^{\frac{\lambda}{\delta_i(1-\beta_i)}} \quad h_0 < h < h^*_i
\]

\[
g^s_i(h) = \begin{cases} m_i c_i \left( \frac{h}{h_i^*} \right)^{-\zeta_i^*-1} & h < h^*_i \\ m_i c_i \left( \frac{h}{h_i^*} \right)^{-\zeta_i^+} & h > h^*_i \end{cases}
\]

where \( c_i = \frac{\zeta_i^+ \zeta_i^-}{(\zeta_i^- - \zeta_i^+) h_i^*} \left( \frac{h_i^*}{h_0^{1-\beta_i}} \right)^{\frac{\lambda}{\delta_i(1-\beta_i)}} \) and where \( \zeta_i^+ \) and \( \zeta_i^- \) are the positive and negative roots of \( \frac{1}{2} \sigma^2_i \zeta_i^2 + \left( z_i^0 \mu_i - \frac{1}{2} \sigma_i^2 \right) \zeta_i = \lambda \).

The equilibrium equations now become:

54
Aggregate welfare has the same form, but I do not repeat the whole expression:

$$W = \sum_i \left[ \int_{h_0}^{h_i} V^w_i(h) g^w_i(h) dh \right] m_i + \sum_i \left[ \int_0^\infty V^w_i(h) g^w_i(h) dh \right] m_i$$

**Ageing**

The final extension introduces finite lifetimes, as opposed to a perpetual youth, which gives the trade-off of spending an extra year in school real bite and more closely reflects the opportunity cost of schooling. Treating death as deterministic at age $Y$ allows us to retain an analytical solution. The workers' HJB equation now becomes:

$$\rho V^w(h, y) = u(\tau_0(wnh)^{1-\tau_1}, n) + V^w_h(h, y) \mu_w h + \frac{1}{2} V^w_{hh}(h, y) \sigma_w^2 h^2 + V^w_y(h, y)$$

And so with CRRA utility as before, their value function becomes:

$$V^w(h, y) = \frac{1}{\kappa} \left( \frac{\tau_0(wnh)^{1-\tau_1}}{1-\gamma} v(n) \right) \left( 1 - e^{\kappa(y-Y)} \right) - \frac{1}{\rho(1-\gamma)}$$

where $n$ is the same as in the Baseline model but where now $\kappa = \rho - (1-\tau_1)(1-\gamma) \left( \mu_w - \sigma_w^2 \right) - (1-\tau_1)^2(1-\gamma)^2 \sigma_w^2$. Notice that if the age of death $Y$ is taken to be infinite, so workers live forever, then the expression collapses to that of the perpetual youth Baseline model with a
death rate $\lambda = 0$, as one would expect.

The HJB equation for students becomes:

$$\rho V^s(h, y) = u(T) + V^s_h(h, y) \left( \mu_s h^\beta_s - \delta_s h \right) + V^s_y(h, y)$$

Rearranging this and using the smooth pasting conditions to find the value function of the student at the cut-off:

$$V^s(h^*, y^*) = \frac{T^{1-\gamma} - 1}{\rho(1-\gamma)} + \frac{(1-\tau_1)(\tau_0 (\omega h^*)^{1-\tau_1})^{1-\gamma} v(n)}{\rho \kappa h^*} \left( 1 - e^{\kappa (y^*-Y)} \right) \left( \mu_s h^* - \delta_s h^* \right)$$

Equating the value functions of students and workers at the cut-off (value matching), we find an implicit function for the threshold values of human capital and age, $h^*(y^*)$, that represents the boundary of when students decide to graduate. Naturally, since ageing one year now means one year less of life left, and hence a diminished option value of graduating, the function $h^*(y^*)$ is decreasing in age.

The symmetry with the Baseline model should hopefully be apparent. Note again that as $Y \to \infty$, the equation collapses to that of the perpetual youth Baseline model with a death rate $\lambda = 0$. Note also that with a finite $Y$, as students approach the age of death, $y \to Y$, the second term drops out, so the value matching condition implies that the optimal graduation threshold will be such that the flow utility from working equals the flow utility of studying, i.e. $u \left( \tau_0 (\omega h^*)^{1-\tau_1} , n \right) = u(T)$. In other words, near-dead students with no future will choose to graduate when the flow utility they could get from working exceeds the flow utility they get as students, or (ignoring disutility from labour momentarily) when their post-tax wage exceeds the student transfer.

The function $h^*(y^*)$ describes the optimal graduation level of human capital for every age (I denote $y$ with a star only for symmetry with $h^*$). However, in the case where there
is no risk while studying, all students will accumulate human capital at the same rate, 
\[ h(y) = \left( h_0^{1-\beta_s} - \frac{\mu_s}{\delta_s} \right) e^{-\delta_s (1-\beta_s) y} + \frac{\mu_s}{\delta_s} \right]^{\frac{1}{1-\beta_s}}. \] Consequently, they will all reach the threshold at the same age, which I denote \( y^{**} \) and call the equilibrium graduation age, and which is determined by equating these two expressions, \( h^*(y^{**}) = h(y^{**}) \). The associated level of human capital, denoted \( h^{**} \), can be found by plugging \( y^{**} \) into either of these equations. Unfortunately there is no closed-form solution, but the equilibrium is shown below:

![Figure 15: Optimal human capital graduation threshold and accumulation path](image)

The density of all agents over age is uniform and simply equal to \( \frac{1}{Y} \). Thus the total measure of students is equal to the share of agents who have not yet reached the equilibrium graduation age, \( y^{**} \), and is thus simply \( y^{**}/Y \), with the total measure of workers therefore the remainder, \( \frac{Y-y^{**}}{Y} \).

The densities of workers and students evolve as follows:

\[
\dot{g}^s(h, y) = -\partial_h \left[ (\mu_s h^{\beta_s} - \delta_s h) g^s(h, y) \right] - \partial_y \left[ g^s(h, y) \right] \quad h \in (h_0, h^*)
\]

\[
\dot{g}^w(h, y) = -\partial_h \left[ \mu_w h g^w(h, y) \right] + \partial_{hh} \left[ \frac{1}{2} \sigma^2 \partial_h^2 g^w(h, y) \right] - \partial_y \left[ g^w(h, y) \right] \quad h \in (0, \infty)
\]
In the stationary equilibrium these will equal 0. They can thus be rearranged as follows:

$$\partial_y [g^s(h, y)] = -\partial_h \left[ \left( \mu_s h^{\beta_s} - \delta_s h \right) g^s(h, y) \right] \quad h \in (h_0, h^*)$$  
$$\partial_y [g^w(h, y)] = -\partial_h \left[ \mu_w h g^w(h, y) \right] + \partial_{hh} \left[ \frac{1}{2} \sigma_w^2 h^2 g^w(h, y) \right] \quad h \in (0, \infty)$$

The first equation is simply the advection partial differential equation, and the second is the advection-diffusion PDE.

The solution to the first is relatively simple. The density of students aged \( y > y^{**} \) is of course zero. The density of students aged \( y < y^{**} \) but with human capital not equal to \( h(y) \) is also zero. Thus the marginal density of students aged \( y < y^{**} \) with human capital equal to \( h(y) \) is thus equal to the total density of students aged \( y < y^{**} \), or \( 1/y \). In words, students of a given age all have the same level of human capital, because they accumulate it at the same rate, thus the density of students at that age with that level of human capital is simply the density of students at that age. A change of variable gives the joint density across human capital and age:

$$g^s(h, y) = \begin{cases} \frac{1}{Y} \frac{1}{\mu_s h^{\beta_s} - \delta_s h} & \text{if } h = h(y) \text{ & } y < y^{**} \\ 0 & \text{otherwise} \end{cases}$$

The second PDE also has a well-known solution that human capital is log-normally distributed, with the variance increasing in age: \( \frac{h}{h^{**}} | y \sim LN \left( \left( \mu_w - \frac{\sigma_w^2}{2} \right) (y - y^{**}), \sigma_w^2(y - y^{**}) \right) \). We can thus use the property that the moments of a log-normally distributed variable are given by \( \mathbb{E} [h^n | y] = (h^{**})^n e^{n(\mu_w - \frac{1}{2} \sigma_w^2)(y-y^{**}) + \frac{1}{2} \sigma_w^2(y-y^{**})} \) to find aggregate variables. For example, tax
The equilibrium is thus fully determined by 3 equations:

\[
\text{Tax revenue} = \int_{y^{**}}^{Y} \int_{0}^{\infty} \left( whn - \tau_0 (wnh)^{1-\tau_1} \right) g^w(h, y) dh dy
\]

\[
= \int_{y^{**}}^{Y} \mathbb{E} \left[ whn - \tau_0 (wnh)^{1-\tau_1} | y \right] \frac{1}{Y} dy
\]

\[
= \int_{y^{**}}^{Y} \left\{ whn^{**} e^{\mu w (y-y^{**})} - \tau_0 (wnh^{**})^{1-\tau_1} e^{(1-\tau_1) (\mu w - \frac{1}{2} \tau_1 \sigma_w^2) (y-y^{**})} \right\} \frac{1}{Y} dy
\]

\[
= \frac{whn^{**}}{\mu_w Y} (e^{\mu w (Y-y^{**})} - 1) - \frac{\tau_0 (wnh^{**})^{1-\tau_1}}{(1-\tau_1)(\mu_w - \frac{1}{2} \tau_1 \sigma_w^2) Y} \left( e^{(1-\tau_1)(\mu_w - \frac{1}{2} \tau_1 \sigma_w^2) (Y-y^{**})} - 1 \right)
\]

The equilibrium is thus fully determined by 3 equations:

\[
h^*(y^*) : \frac{(\tau_0 (wnh^*)^{1-\tau_1})^{1-\gamma} v(n)}{\kappa (1-\gamma)} (1 - e^{\kappa (y^*-Y)}) = \frac{T^{1-\gamma}}{\rho (1-\gamma)} - \frac{(\tau_0 (wnh^*)^{1-\tau_1})^{1-\gamma} v(n)}{\rho (1-\gamma)} e^{\kappa (y^*-Y)}
\]

\[
+ \frac{(1-\tau_1) (\tau_0 (wnh^*)^{1-\tau_1})^{1-\gamma} v(n)}{\rho h^*} (1 - e^{\kappa (y^*-Y)}) \left( \mu_s h^* \beta_s - \delta_s h^* \right)
\]

\[
y^{**} : h^*(y^{**}) = \left[ \left( h_0^{1-\beta_s} - \frac{\mu_s}{\delta_s} \right) e^{-\delta_s (1-\beta_s) y^{**}} + \frac{\mu_s}{\delta_s} \right]^{1-\tau_s}
\]

\[
\tau_0 : \int_{y^{**}}^{Y} \int_{0}^{\infty} \left( whn - \tau_0 (wnh)^{1-\tau_1} \right) g^w(h, y) dh dy = \frac{y^{**}}{Y} (T + F + G)
\]

Since solving for this equilibrium is not as trivial as in the Baseline model – which simply involves solving a system of non-linear equations – let me briefly outline the (pseudo-)algorithm I use to solve it (for a given \(\tau_1\) and \(T\)):

1. Create a grid for \(y\) and guess an initial \(\tau_0\)

2. Find the value functions, \(V^w(h, y)\) and \(V^s(h^*, y^*)\), using the formulae above

3. For each gridpoint in \(y\), find the \(h^*\) that satisfies the value matching equation using a root-finding algorithm, yielding a vector for \(\hat{h}^*\)

4. Interpolate the discrete mapping \(y \mapsto \hat{h}^*\) (I use splines), yielding the continuous function

59
5. Find the $y^{**}$ that equates $h^*(y^*)$ and $h(y)$ using a root-finding algorithm, and hence find $h^{**}$ from $h^*(y^{**})$ or $h(y^{**})$

6. Calculate the government budget surplus using the equation above, and update $\tau_0$ (I use bisection) until it balances to within some tolerance

Aggregate worker welfare is given by:

\[
W^w = \int_{y^{**}}^{Y} \int_{0}^{\infty} V^w(h, y) g^w(h, y) dh dy \\
= \int_{y^{**}}^{Y} \mathbb{E}[V^w(h, y)|y] \frac{1}{Y} dy \\
= \int_{y^{**}}^{Y} \left[ \frac{(\tau_0(wnh^{**})^{1-\gamma})^{1-\gamma} v(n)}{\kappa(1-\gamma)} \left(1 - e^{\kappa(y - Y)}\right) e^{(\rho - \kappa)(y - y^{**})} - \frac{1}{\rho} \right] \frac{1}{Y} dy \\
= \frac{(\tau_0(wnh^{**})^{1-\gamma})^{1-\gamma} v(n)}{\kappa(1-\gamma)} \left[ \frac{1}{\rho - \kappa} (e^{(\rho - \kappa)(Y - y^{**})} - 1) \\
- \frac{1}{\rho} e^{-\kappa(Y - y^{**})} (e^{(Y - y^{**})} - 1) \right] \frac{1}{Y} - \frac{1}{\rho(1-\gamma)} \frac{Y - y^{**}}{Y}
\]

Aggregate student welfare is:

\[
W^s = \int_{0}^{y^{**}} \int_{h_0}^{h^{**}} V^s(h, y) g^s(h, y) dh dy \\
= \int_{0}^{y^{**}} V^s(h(y), y) g^s(y) dy \\
= \frac{T^{1-\gamma} - 1}{1-\gamma} \int_{0}^{y^{**}} \frac{1}{Y} dy + \left[ V^w(h^{**}, y^{**}) - \frac{T^{1-\gamma} - 1}{1-\gamma} \right] \frac{1}{Y} \int_{0}^{y^{**}} e^{-\rho(y^{**} - y)} dy \\
= \frac{T^{1-\gamma} - 1}{1-\gamma} \frac{y^{**}}{Y} + \left[ V^w(h^{**}, y^{**}) - \frac{T^{1-\gamma} - 1}{1-\gamma} \right] \frac{1 - e^{-\rho y^{**}}}{\rho Y}
\]

And total aggregate welfare across all agents is the sum of the two.
Abstract

Why has the real interest rate on public debt, \( r \), recently fallen below the growth rate of the economy, \( g \), and what does it imply for the limits of government borrowing? Idiosyncratic investment risk can motivate precautionary saving in safe assets like government debt, which pushes down interest rates. But I argue two factors – so far mostly overlooked in this literature – explain much of the decline in \( r \) vs \( g \) over recent decades: (1) the increasing presence of emerging market (EM) economies, whose investors face even greater idiosyncratic risk and hence have stronger precautionary saving motives; and (2) limits to the private supply of safe assets. These two elements both afford developed market (DM) governments greater capacity to borrow and dampen the negative side-effects of debt-financed government spending shocks, like the response to the pandemic. Nonetheless, fiscal space is not unlimited, and governments should design fiscal rules that ensure debt is both sustainable in the long-run and stabilising in the short-run.
1 Introduction

When the real interest rate on government debt, $r$, is below the growth rate of the economy, $g$, the traditional arithmetic of government budget constraints ceases to apply: the value of a government’s debt is no longer equal to the present value of its future primary surpluses. Instead, a government can run a permanent deficit or finance a one-off expenditure with debt and roll it over without subsequently running surpluses. Government debt is a bubble, which within limits can be mined for additional spending, in the language of Brunnermeier et al. (2021a). This is the situation developed market (DM) economies of today find themselves in, as exemplified by recent estimates of the natural real interest rate and trend growth rate from Holston et al. (2017), shown in Figure 1. In this paper, I take the view that part of the recent decline in $r$ vs $g$ is due to increasing demand for safe assets from emerging market (EM) economies like China, as well as declines in the private supply of safe assets following the financial crisis, and I explore the implications of these factors for the fiscal space of DM governments.

Figure 1: $r^*$ vs $g$ in developed market economies over the last 50 years

Source: Estimates of $r^*$ and $g$ from Holston et al. (2017) are a GDP-weighted average of their estimates for the US, Canada, UK and Euro area, using OECD estimates of GDP PPP. The pre-1995 Euro area weight is the sum of the weights of the original 11 Euro area countries.
$r < g$ has led some to conclude that governments of advanced economies should spend more (Blanchard 2019; Furman & Summers 2020). Others argue that while increasing spending with debt finance may be feasible, it may not be welfare-enhancing, since the economy is still dynamically efficient, as the marginal product of capital is still above the growth rate (Reis 2021; Ball & Mankiw 2021). Others are agnostic on welfare, but simply try to explain the phenomenon, as well as the related observation that government debt appears to be valued in the market significantly more than the present value of expected future surpluses; this bubble is on the order of 3x GDP, according to Jiang et al. (2021). Two main explanatory frameworks have emerged in the literature: the overlapping generations framework; and, more recently, the idiosyncratic risk framework.

The overlapping generations framework is well-established and goes back to Samuelson (1958) and Diamond (1965), while Tirole (1985) and more recently Farhi & Tirole (2012) in particular address the possibility of asset bubbles in such a framework. In such models, agents’ need to save for their retirement is what drives down interest rates. This framework yields a natural explanation for the recent decline in $r < g$: demographic changes. Today the population is older, spends longer in retirement and perhaps also anticipates less generous public pensions, necessitating higher saving and pushing down interest rates. See Eggertsson et al. (2019) or Auclert et al. (2021) for a discussion of this view. It is this framework that Blanchard (2019) and others have used to explore the implications of $r < g$ for fiscal space.

The second, more recent strand of the literature relies on idiosyncratic risk – and the precautionary saving it motivates – as the key mechanism pushing down interest rates, starting with Brunnermeier et al. (2021a) and (2021b), known henceforth as BMS. These papers also try to establish the limits of fiscal space, extending the fiscal theory of the price level to allow for a bubble in government debt. While the idiosyncratic risk framework can in principle explain why $r < g$, it does not by itself offer a compelling rationale for why $r$ is so much less than $g$ today than it was 50 years ago. One can argue that idiosyncratic risk has increased; this is hard to observe directly, but we can potentially see some of the consequences of this
like higher income and wealth inequality, as discussed for example in Mian et al. (2021c) and (2021b), who furthermore argue that the evidence suggests higher inequality explains the decline in interest rates better than demographics (although they do not attribute this to idiosyncratic risk). We can also point to some tentative causes such as the financial crisis, which obviously precipitated a large fall in real interest rates.

My motivation in this paper then is to supply this rationale explaining the secular *decline* in interest rates – global imbalances and a dearth of safe assets – and hence to strengthen the case for the idiosyncratic risk framework. Moreover I explore the consequences for the fiscal space of governments. I thus extend BMS by incorporating global financial markets and limits on private safe asset supply, although my model is real and not concerned with nominal aspects.

The global imbalances-safe assets hypothesis stems from two stylised facts that have emerged in the last two decades: one relating to the demand for safe assets and the other relating to their supply. On the demand-side: the growth of China and other EM economies. Even 20 years ago, China accounted for less than 3% of global wealth, today it accounts for nearly a fifth (Shorrocks et al. 2021). It has concomitantly become one of the largest holders of US Treasuries and other DM government bonds. EM investors’ demand for safe assets may in part be because they face higher idiosyncratic risk than investors in developed economies, for myriad reasons. On the supply-side: the diminished private supply of safe assets since the financial crisis. Issuance of AAA-rated mortgage-backed securities totalled around $1 trillion a year prior to the financial crisis, only to be subsequently “decisected”—cut in one-tenth. These two factors have been discussed in the literature on global savings gluts, such as Angeletos & Panousi (2011) or Caballero et al. (2008), (2017) and (2020), but so far to my knowledge no one has discussed their implications for the fiscal space of DM governments. In this paper I thus link the literatures on idiosyncratic risk-induced precautionary saving and that on global imbalances and provide a tractable framework within which to analyse the effects of both on the fiscal space of governments.
As in BMS, in this paper investors face idiosyncratic risk on their capital holdings, which drives precautionary saving into risk-free government bonds and pulls the interest rate below the growth rate. My innovation is to nest this model in a two-country framework, reminiscent of Angeletos & Panousi (2011). Like that paper, I take the view that the degree of idiosyncratic risk facing entrepreneurs in a country represents the level of development in that country: less developed countries face higher risk, for example owing to worse legal protections for business owners or less developed financial markets with less capacity for risk-sharing. I novelty use this framework to explore the implications of global imbalances — driven by idiosyncratic risk — for fiscal space. Since investment is riskier, the precautionary saving motive is even stronger in the developing world, so integration of global capital markets results in emerging market (EM) investors lending to developed market (DM) entrepreneurs and governments, pulling down DM interest rates further, consistent with the observation that capital tends to flow from poorer to richer countries (Lucas 1990, Gourinchas & Jeanne 2013).

I introduce two notions of fiscal space in this paper: the “Bubble Bound”—the level of public debt below which the government faces negative interest rates and so can exploit its bubble; and the “Profligacy Peak”—the maximum deficit that the government can sustain. I prove that for DM governments the Bubble Bound is higher under integrated global capital markets than under autarky, and I demonstrate that the Profligacy Peak also tends to be larger; moreover both are, under certain weak conditions, increasing in the degree of idiosyncratic risk at home and abroad.

Integrated financial markets can thus materially expand the amount of fiscal space available to governments in the developed world compared to autarky, all else equal: governments can sustain higher debt relative to GDP before the bubble bursts and they have to start running primary surpluses, and they can sustain higher deficits. However, if both DM and EM governments are attempting to use the maximum amount of fiscal space available, integrated financial markets do not necessarily offer an advantage over autarky. To borrow and extend the colourful metaphor of Brunnermeier et al. (2021a), financial integration allows governments
to mine another country’s bubble, but doesn’t by itself expand the combined size of the two bubbles.

One implication of the above is that integration can pose risks to developing countries, if they are already close to their Bubble Bound. Integrating with a more developed (less risky) country, that in the new equilibrium will be a net borrower, can lead to capital flight and permanently lower output in the EM country, while substantially lowering the amount of fiscal space available to its government. The result that integration can reduce the output of emerging market economies and raise that of developed economies is the reverse of the results in for example Angeletos & Panousi (2011), Corneli (2017) & (2021); it is due to subtle modelling choices, but perhaps demonstrates the fragility of the conventional wisdom and the risks of integrated capital markets. The model in this paper is stylised, but may offer insights into the performance of the Russian economy and that of other former Soviet countries following the collapse of the USSR. Upon integrating with the global economy, these countries experienced significant capital flight, a sharp fall in output and an immediate need to reduce government borrowing, and – when the latter was not achieved – experienced rapid inflation and later a financial crisis.

The second key innovation in this paper over prior research on fiscal space is to incorporate limits on the private supply of safe assets. I model these limits as a collateral-based credit constraint, à la Kiyotaki & Moore (1997), as a reduced-form way to capture constraints on the ability of the financial sector to transform risky business and personal loans into riskless securities. Such constraints seem evident in light of the collapse of the asset-backed securities (ABS) market following the financial crisis, and as shown in Figure 1 it is following that crisis that (natural) real interest rates began to fall substantially below trend growth. In this paper I explore the consequences of such constraints for governments’ fiscal space. I prove that the Bubble Bound is larger when these constraints bind, and show quantitatively that the Profligacy Peak is also higher. Moreover, credit constraints substantially lessen the crowding-out effects of public debt for the country whose entrepreneurs are constrained, since
by construction other investors who could potentially better use the funds are constrained from borrowing. Note that Reis (2021) also features a borrowing constraint, but it plays a slightly different role, which I discuss in depth below.

I am not the first to try to triangulate concrete limits on fiscal space. Mian et al. (2021) also try to quantify the Bubble Bound and Profligacy Peak (under different names). Their paper uses a representative agent framework with bonds in the utility function to incentivise saving. They also consider the impact of the zero lower bound, which has the effect of raising real rates and reducing fiscal space when it binds. Kocherlakota (2021) offers a benchmark case where there are no limits to fiscal space. He considers a variation of an Aiyagari (1994) model with an additional extreme idiosyncratic state, reached with arbitrarily low probability but with arbitrarily high and constant marginal utility of consumption. In this “all-you-can-eat” state, the household feels compelled to consume all their savings, for example because they suffer an adverse health shock and must pay their medical bills, although it is not given such an interpretation in the paper. The presence of such a state yields extreme precautionary saving at the individual level, and a willingness to absorb arbitrarily high quantities of public debt at negative interest rates at the aggregate level—an infinite Bubble Bound.

I also consider the optimal level of public debt, adding to earlier research by Aiyagari & McGrattan (1998), Desbonnet & Kankanamge (2016) or more recently Le Grand & Ragot (2022). They however examine optimal debt levels in the presence of distortionary taxes and government spending, I do the same in a model specifically designed to neutralise the distorting effects of taxes and spending, and hence to isolate the welfare effects of public debt. One should therefore view the welfare results in the body of this paper as a benchmark bound against which to evaluate more quantitatively realistic models. I find the optimal level of public debt is positive and much higher than these prior authors find, and much higher than the levels observed in advanced economies today; several multiples of output. However, debt at these levels is well beyond the Profligacy Peak or Bubble Bound notions of fiscal space discussed above; the bubble in public debt is deflated.
Much of the recent research instead focuses on whether or not debt issuance can be Pareto-improving, like Blanchard (2019), Reis (2021), Aguiar et al. (2021) or Brumm et al. (2021). The latter of which also evaluates the effects of international trade and capital markets, likewise finding they expand the potential for welfare-enhancing debt issuance, although they primarily consider the case of two symmetric countries, rather than that of systemic global imbalances. However, debt issuance in my model is in general not Pareto-improving, even when \( r < g \), since issuing debt likely raises \( r \), which typically crowds out capital and hence lowers workers’ wages. To be Pareto-improving, the debt must be able to fund increases in transfers to workers that more than compensate for their lower wages; this only occurs when \( r \ll g \), much more so than is the case today.

I furthermore explore the dynamics of the model, subjecting the economy to an unexpected, transitory, debt-financed 10% of GDP shock to government spending, similar to the response to the Covid-19 pandemic. I find that financial integration dampens the negative effect of the spending shock on the output of the developed country, relative to autarky; credit constraints further dampen and delay the negative effects relative to the unconstrained case.

But \( r < g \) does not give governments free rein to spend and borrow more. My analysis suggests a number of cautions for governments looking to exploit low interest rates. Governments should:

1. Target a steady state level of debt, rather than the deficit, in order to ensure a steady state equilibrium exists

2. Respond sufficiently aggressively to excessive debt, in order to ensure the steady state is (locally) stable

3. Follow large spending shocks with sufficiently aggressive austerity, in order to ensure the economy returns to the (desired) steady state, as dynamics may be globally unstable even if they are locally stable

While these principles accord, I think, with the conventional wisdom and, arguably, with the
historical record in most developed countries (e.g. Bohn 1998, Chen et al. 2021) and with certain (somewhat arbitrary) fiscal rules currently in place (but only dubiously enforced), like the EU’s Maastricht criteria; they are arguably inconsistent with the current projected path of fiscal policy in the US, UK and other developed economies, as well as with certain other fiscal rules that have been recently proposed in various countries (e.g. McNicol 2017).

Not adhering to these principles may result in the economy “blowing up”, with an equilibrium either not existing, or with the economy diverging away from steady state following a shock. The model in this paper is real, but in a monetary model with nominal debt claims, such a scenario would imply hyperinflation, like the non-monetary steady state of BMS (2021b). A novel result in my paper is that even if the economy avoids hyperinflation in response to large shocks, it may still face a “debt trap”, where it approaches a second steady state featuring higher debt and lower output and remains there for an extended period. It may prove challenging to escape such a debt trap, necessitating sufficiently large austerity to escape and return to the normal steady state within a reasonable timeframe. Note that this applies even when \( r \ll g \).

In sum, the paper demonstrates that accounting for global saving imbalances and constraints on private-sector safe asset creation is critical when assessing the amount of fiscal space available to governments today, and offers a tractable framework within which to do that. The increasing importance of China and the rest of the developing world in global capital markets, as well as the limited ability of the private sector to supply said capital markets with sufficient quantities of safe assets, suggests governments of advanced economies may have considerably more fiscal space available than in the past, though still within clearly defined limits. At the same time, the model suggests caution for governments of emerging economies: integrating with globalised capital markets can lead to capital flight and substantially reduce their ability to borrow.
2 Model of Global Imbalances due to Idiosyncratic Risk

The baseline model is as follows: time is continuous, \( t \in [0, \infty) \), and households comprise a measure one each of infinitely-lived entrepreneurs and workers. Workers are homogeneous and cannot save, so consume all their income from labour, which they supply inelastically in equilibrium. The firm side consists solely of entrepreneurs.

Entrepreneurs each operate their own firm and own capital, \( k > 0 \), and hire labour, \( l > 0 \), at wage, \( w_t \), using which they produce output. They can also purchase or lend risk-free bonds, \( b \in \mathbb{R} \), on which they earn the real interest rate \( r_t \). Their net wealth is thus \( a = b + k \), and evolves over time as \( \dot{a} = \dot{b} + \dot{k} \).

There are two countries, \( i = H, F \), each with its own entrepreneurs, workers and governments. I abstract from international trade in goods and currency markets by having entrepreneurs in both countries produce the same good, which trades at the same price, resulting in currencies that trade at parity.

The framework borrows from Angeletos & Panousi (2011), extended to incorporate a government sector and credit constraints on entrepreneurs. However, unlike that paper, I assume that entrepreneurs themselves do not work, but instead hire workers who cannot save. I make these simplifying assumptions for two reasons: (1) it allows for a closed-form solution for the entrepreneurs’ policy functions for consumption and capital in the presence of collateral-based credit constraints; and (2) it allows the interest rate to be negative, for reasons that will become clear when I analyse the existence of steady state equilibrium in section 3, which is obviously central to my analysis.

Entrepreneurial profits are given by revenue less labour and depreciation costs, subject to an idiosyncratic shock that is proportional to the entrepreneur’s capital stock. Revenue simply equals output, produced according to a constant returns to scale Cobb-Douglas production function of the entrepreneur’s capital and labour demand, as the consumption good is the
numeraire. An entrepreneur’s flow profits are thus:

\[ d\pi_t = \left(k_t^{\alpha} l_t^{1-\alpha} - w_t l_t - \delta k_t \right) dt + \sigma k_t dW_t \]

where \( W_t \) is a Brownian motion, representing the idiosyncratic risk to the entrepreneur’s profits that averages out across agents. One can interpret this most straightforwardly as a depreciation shock to their capital stock. \( \sigma > 0 \) is the standard deviation of this shock. There is no aggregate risk, but I do explore MIT shocks under a perfect foresight equilibrium, so I allow prices, and hence the entrepreneurs’ value functions and optimal policies, to vary over time. \( \alpha \in (0, 1) \) represents the capital intensity of production and \( \delta > 0 \) the (average) rate of depreciation.

The entrepreneur consumes and saves out of profits and interest from their risk-free savings, so their budget constraint is:

\[ da_t = d\pi_t + (r_t b_t - c_t) dt \]

Plugging in profits and substituting out bond holdings gives:

\[ da_t = \left(k_t^{\alpha} l_t^{1-\alpha} - w_t l_t - (r_t + \delta) k_t + r_t a - c_t \right) dt + \sigma k_t dW_t \]

We can immediately solve for optimal labour demand and replace in the budget constraint to simplify:

\[ l_t = \left(1 - \frac{\alpha}{w_t} \right)^{\frac{1}{\alpha}} k_t \]

Entrepreneurs may also face credit constraints, such that they cannot borrow more than a fraction \( \theta \in [0, 1] \) of their capital: \( -b \leq \theta k \), or equivalently that their capital holdings must be no more than a fraction \( \lambda = \frac{1}{1-\theta} \geq 1 \) of their net worth: \( k \leq \lambda a \). In the absence of the collateral constraint, entrepreneurs also face a less strict no-Ponzi condition to rule out infinite accumulation of debt. The precise form this takes is not important – a simple (non-positive) lower bound on net wealth will suffice: \( a \geq a_0 \) – since with Ponzi schemes ruled out, it will
always be optimal to keep wealth positive, given the absence of non-financial income.

**Entrepreneurs’ problem.** The entrepreneurs’ problem is to maximise their lifetime utility, discounted at rate $\rho > 0$, by choosing consumption, capital and labour demand, given their initial wealth at time $t_0$, $a_0 > 0$, and subject to their (stochastic) budget constraint and collateral constraint or no-Ponzi condition:

$$V(t_0, a_0) = \max_{\{c_t, k_t\in [t_0, \infty)\}} \mathbb{E}_t \left[ \int_{t_0}^{\infty} e^{-\rho(t-t_0)} u(c_t) dt \right]$$

subject to

$$da_t = ((R_t - r_t)k_t + r_t a_t - c_t) \, dt + \sigma k_t dW_t$$

where $k_t \leq \lambda a_t$

where $R_t = \alpha \left( \frac{1-\alpha}{w_t} \right)^{\frac{1-\alpha}{\alpha}} - \delta$ is the expected return on capital. I assume entrepreneurs have CRRA preferences, $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$, with $\gamma \geq 1$ the coefficient of relative risk aversion. Note I have already substituted out the entrepreneur’s labour demand problem, so the resulting problem is simply analogous to a Merton model: choose a fraction of wealth to invest in the risky asset, a fraction in the risk-free asset, and a flow consumption rate.

This problem can be more easily solved recursively, and to this end can be expressed as a Hamilton-Jacobi-Bellman equation:

$$\rho V(t, a) = \max_{c, k \leq \lambda a} u(c) + \partial_a V(t, a) ((R_t - r_t)k + r_t a - c) + \frac{1}{2} \partial_{aa} V(t, a) \sigma^2 k^2 + \dot{V}(t, a)$$  \hspace{1cm} (1)

where $\partial$ denotes the derivative operator. The entrepreneurs’ problem can be solved with a single state variable, individual wealth $a$, since capital and bond wealth are perfectly fungible (the problem would still only require one state variable even if capital for example were not perfectly fungible, provided adjustment costs were convex). However, since I allow for MIT shocks and deterministic transition dynamics, the value function and the entrepreneurs’ policy functions depend on prices and hence may change over time. Since entrepreneurs have CRRA preferences, their value function will have that same form – isoelastic in net wealth – and
their policy functions will be linear in net wealth.

The presence of a borrowing constraint of this form presents no issue for finding an analytical solution to the entrepreneurs’ problem. When the constraint does not bind the solution is simply that of the unconstrained problem, and when it does bind, the solution can be found simply by substituting the binding constraint \( k = \lambda a \) directly into the HJB equation. In essence, the entrepreneurs’ optimal investment decision is myopic; they “see” the constraint only when it binds, as shown in Vila & Zariphopoulou (1997).

**Lemma 1.** The entrepreneurs’ value function is as follows:

\[
V(t, a) = \begin{cases} 
\frac{m_t \gamma}{1-\gamma} a^{1-\gamma} & \text{if } \gamma \neq 1 \\
E_t + \frac{1}{\rho} \ln(a) & \text{if } \gamma = 1
\end{cases}
\]

where \( E_t \) is determined by the following ODE:

\[
\dot{E}_t = \rho E_t + 1 - \ln(\rho) - \frac{1}{\rho} \left( (R_t - r_t)\omega_t + r_t - \frac{1}{2} \sigma^2 \omega_t^2 \right)
\]

The entrepreneurs’ policy functions are:

\[
c(t, a) = \begin{cases} 
m_t a & \text{if } \gamma \neq 1 \\
\rho a & \text{if } \gamma = 1
\end{cases}
\]

\[
k(t, a) = \min \left\{ \frac{R_t - r_t}{\gamma \sigma^2}, \lambda \right\} a \equiv \omega_t a
\]

\[
b(t, a) = (1 - \omega_t) a
\]

where their marginal propensity to consume, \( m_t \), is determined by the following ODE:

\[
\frac{\dot{m}_t}{m_t} = m_t - r_t + \frac{r_t - \rho}{\gamma} + \frac{1 - \gamma}{\gamma} (R_t - r_t)\omega_t + \frac{\sigma^2 (\gamma - 1)}{2} \omega_t^2
\]

Their Euler equation is:

\[
\frac{dc_t}{c_t} = \left( \frac{r_t - \rho}{\gamma} + \frac{R_t - r_t}{\gamma} \omega_t + \frac{\sigma^2 (\gamma - 1)}{2} \omega_t^2 \right) dt + \omega_t \sigma dW
\]
The entrepreneurs’ optimal choices of consumption, and hence their optimal paths of wealth given initial wealth \( a_0 \), also satisfy the transversality condition:

\[
\lim_{t \to \infty} \mathbb{E}_{t_0} \left[ e^{-\rho(t-t_0)} u'(c_t)a_t \right] = 0.
\]

**Proof.** See Appendix.

**Representative worker’s problem.** Entrepreneurs do not work in this model, to preserve the analytical solution. Instead I assume that entrepreneurs hire hand-to-mouth workers who are unable to save and who also pay taxes/receive transfers, \( \tau_t \), that are proportional to their income and which will be one possible source of government deficits/surpluses. As a result, the workers’ problem is simple and does not affect the dynamics of the entrepreneurs’ wealth/capital accumulation and consumption. I assume workers supply one unit of labour inelastically. Since workers have no wealth this is equivalent to assuming they have preferences consistent with balanced growth. In this case the workers’ consumption-labour optimality condition would be: \( c_t l_t^\varphi = w_t (1 - \tau_t) \), where \( \varphi \) is the inverse Frisch elasticity of labour supply, which given they consume their income results in constant labour supply \( l_t = 1 \).

### 3 Aggregate Equilibrium

Since we have a representative worker who supplies a unit labour inelastically, aggregate labour supply is also \( L_t = 1 \). Combining this with entrepreneurs’ aggregate labour demand – and hence imposing equilibrium in the labour market – yields an expression for the wage rate in terms of the aggregate capital stock: \( w_t = (1 - \alpha) K_t^\alpha \). Given that all entrepreneurs are ex-ante identical, and given that their policy functions are all linear in their net wealth, aggregation is simple; the aggregate economy simply behaves like a scaled-up version of an individual, with idiosyncratic risk washing out in the aggregate. The corresponding aggregate variables are denoted with capital letters:

\[
C_t = \int_0^\infty c(t,a)g(t,a)da \quad K_t = \int_0^\infty k(t,a)g(t,a)da \quad B_t = \int_0^\infty b(t,a)g(t,a)da
\]
There is a distribution of wealth, $g(t,a)$, which one can characterise, but it has no effect on the aggregate economy. In addition, credit constraints will either bind for all entrepreneurs in a country or none. There are two countries in the model, denoted $H$ and $F$ for Home and Foreign, representing the developed world and the developing world, and from here on I subscript quantities and prices in a given country with their respective index or with $i$ in general.

The model is closed through bond market equilibrium: an Autarkic equilibrium, where each country’s bond market clears separately, with domestic bondholdings, $B_{it}$, equalling domestic public debt outstanding, $D_{it}$; or an Integrated equilibrium, where bond markets clear globally.

**Definition 1.** An Autarkic equilibrium in the economy with hand-to-mouth workers is a sequence for aggregate variables $\{K_{it}, B_{it}, C_{it}\}_{t=0}^{\infty}$ for $i = H,F$, aggregated from policy functions $\{k_i(t,a), b_i(t,a), c_i(t,a)\}_{t=0}^{\infty}$, which maximise agents’ lifetime utility given prices $\{r_{it}, R_{it}, w_{it}\}_{t=0}^{\infty}$, and where all markets clear, including bond markets in each country $B_{it} = D_{it}$.

Note that since workers’ consumption has no effect on aggregate dynamics, I ignore it in discussion of the equilibrium; aggregate consumption refers to that of entrepreneurs alone.

**Definition 2.** An Integrated equilibrium is a sequence for aggregate variables $\{K_{it}, B_{it}, C_{it}\}_{t=0}^{\infty}$ for $i = H,F$, aggregated from policy functions $\{k_i(t,a), b_i(t,a), c_i(t,a)\}_{t=0}^{\infty}$, which maximise agents’ lifetime utility given prices $\{r_{it}, R_{it}, w_{it}\}_{t=0}^{\infty}$, and where all markets clear, including global bond markets $B_{Ht} + B_{Ft} = D_{Ht} + D_{Ft}$.

**Proposition 1.** The aggregate dynamics of the two-country model with hand-to-mouth workers are given by the following system of non-linear ordinary differential equations and static
Global Imbalances & The Limits of Fiscal Space

Lee Tyrrell-Hendry

equations:

\[ \dot{K}_{it} + \dot{B}_{it} = R_{it}K_{it} + \rho_{it}B_{it} - C_{it} \]  
\[ \dot{C}_{it}/C_{it} = \frac{R_{it} - \rho}{\gamma} + \frac{R_{it} - r_{it}}{\gamma}\omega_{it} + \frac{\sigma_{i}^{2}(\gamma - 1)}{2}\omega_{it}^{2} \]  
\[ R_{it} = \alpha K_{it}^{\alpha-1} - \delta \]  
\[ \omega_{it} = \min \left\{ \frac{R_{it} - r_{it}}{\gamma\sigma_{i}^{2}}, \lambda_{i} \right\} \]

where \( i = H, F \) for all equations.

Governments in both countries satisfy the following real budget constraint:

\[ \dot{D}_{it} = r_{it}D_{it} - (T_{it} - G_{it}) \]

where \( T_{it} \) and \( G_{it} \) are determined depending on the specification of the model.

Under Autarky, the local risk-free interest rates adjust to clear local bond markets:

\[ B_{it} = D_{it} \]

Under Integrated financial markets, there is a single risk-free rate \( r_{Ht} = r_{ Ft } = r_{t} \) that 
clears the global market:

\[ B_{Ht} + B_{Ft} = D_{Ht} + D_{Ft} \]

Proof. Equation (2) comes from aggregating the entrepreneurs’ budget constraint. Equation (3) is the aggregate Euler equation, derived by aggregating the individual’s Euler equation 
given above. (4) is an expression for the expected return on capital, derived from that 
expression above (in terms of \( w_{t} \)), with the workers’ labour supply condition substituted in, 
thus it already imposes labour market equilibrium. Equation (5) is the aggregated optimal 
portfolio share in capital.

Theorem 1. Under both Autarky and Integration, an equilibrium price vector exists and is
Proof. Under both Autarky and Integration, the expected return on capital and the wage rate are given by $R_i(K_i) = \alpha K_i^{\alpha - 1} - \delta$ and $w_i(K_i) = (1 - \alpha) K_i^\alpha$. It remains to determine the equilibrium risk-free real interest rate, $r$.

Under Autarky, combining equations 4 and 5 with 7, bond market equilibrium is given by

$$B_i = \left(\frac{\gamma \sigma_i^2}{\alpha K_i^{\alpha - 1} - \delta - r_i} - 1 \right) K_i = D_i,$$

or rearranging, the interest rate is $r_i^A(K_i, D_i) = \alpha K_i^{\alpha - 1} - \delta - \frac{K_i}{K_i + D_i} \gamma \sigma_i^2$.

Under Integration, combining those same equations with 8, bond market equilibrium becomes:

$$B(K_H, K_F, r) = \max \left\{ \frac{\gamma \sigma_H^2}{\alpha K_H^{\alpha - 1} - \delta - r} - 1, -\theta_H \right\} K_H + \max \left\{ \frac{\gamma \sigma_F^2}{\alpha K_F^{\alpha - 1} - \delta - r} - 1, -\theta_F \right\} K_F = D$$

Clearly only the total stock of public debt matters, so the state variables $K_H, K_F, D$ pin down the equilibrium interest rate via the mapping $r^I(K_H, K_F, D)$. This mapping has no closed form, but is unique by the implicit function theorem, since $\frac{\partial}{\partial r} [B(K_H, K_F, r)] \neq 0$ anywhere in the positive orthant of $(K_H, K_F)$. Note that when credit constraints bind, say in the Home country, the equilibrium interest rate becomes $r^{I,CC}(K_H, K_F, D) = \alpha K_F^{\alpha - 1} - \delta - \frac{K_F}{K_F + D + \theta_H K_H} \gamma \sigma_F^2$, with the indices reversed if constraints bind in the Foreign country.

I have yet to specify government taxes and spending, and it is actually not necessary to do so to characterise the steady state given the assumptions I have made; only the primary surplus/deficit and the level of public debt matter for the equilibrium—a kind of reverse Ricardian equivalence. Taxes/transfers to workers are proportional to labour income, and so given the preferences of workers outlined above have no distorting effects since income and substitution effects cancel. Moreover, government spending merely crowds out workers’ consumption, since they pay all the taxes, but because workers cannot save does not by itself distort the equilibrium. This result is peculiar to this specification of the model, where for simplicity I have assumed consumption and capital goods are totally fungible. If there
were adjustment costs on capital, then this would not hold and there would be additional distortions from government spending, as is the case in BMS (2021b).

A Bubble in Government Debt. Integrating the government’s flow budget constraint gives its intertemporal equivalent:

\[
D_t = \int_t^{\infty} e^{-\int_t^s r_v d\tau} (T_s - G_s) ds + \lim_{T \to \infty} e^{-\int_t^T r_v d\tau} D_T
\]

This holds in the absence of aggregate risk, so primary surpluses are discounted at the risk-free rate, and in the absence of seignorage revenues, i.e. when the interest rate on money is the same as that on government debt. The latter likely approximately holds true today in most developed countries, as most central banks now pay interest on reserves – which today comprise the bulk of high-powered money – not far below that on debt.

There is nothing to rule out the existence of a bubble in this model; the last term need not equal zero owing to some transversality or no-Ponzi condition, and in steady state will indeed be positive whenever the interest rate is below the growth rate of the economy, \( r < g \), as discussed in BMS (2021b). In this paper I abstract from long run growth \( (g = 0) \), as a number of sources highlight that factors other than declining growth are most important for explaining the decline in global real interest rates, for example Rachel & Smith (2015), Rachel & Summers (2019a), (2019b), Mian et al. (2021c) or Holston et al. (2017), (2020), who update Laubach & Williams (2003) in estimating the causes of the decline in the natural rate of interest. This is not completely without loss of generality, since when \( \gamma \neq 1 \) \( r \) fluctuates \( \gamma \): 1 with \( g \), but since for much of this paper I will be considering calibrations with \( \gamma = 1 \) or close to 1, the loss is not material. The relevant criterion for a bubble is thus \( r < 0 \), which as shown below holds whenever domestic and or foreign investment is sufficiently risky.

In the steady state, the equation becomes:
where $s = T - G$ now denotes the primary surplus. In the long run, for a steady state to exist these terms must exactly offset each other by having $D = \frac{s}{r}$, or $\dot{D} = 0$ in the government’s flow budget constraint. Thus, a bubble in public debt ($r < 0$) does not mean governments can issue however much they want. In the long run, the real value of public debt must stabilise, but stable public debt is consistent with persistent deficits, equal to $-rD$, if $r$ is negative. Moreover, as Cochrane (2021) highlights, there is no discontinuity at $r = g$; the maximum sustainable primary deficit smoothly increases by $\varepsilon$ as the interest rate moves $\delta$ below 0.

If a government were to maintain a persistent deficit larger than $-r_tD$, government debt would accumulate, and by bond market equilibrium so would entrepreneurs’ bond holdings, creating a persistent crowding-out of capital, until the real value of capital and hence output reached zero. A (positive) steady state equilibrium would not exist. This is because the model here is entirely real. If I were to introduce prices and make government debt a nominal claim, the problem may reconcile itself through hyperinflation, analogous to the non-monetary equilibrium in BMS (2021a, 2021b).

**Steady State**

In the steady state, equations 2, 3 and 6 are equal to zero in both countries. Based on the discussion above, I proceed to analyse the steady state having already imposed $\dot{D} = 0$, implicitly assuming the primary surplus, $s$, adjusts to accommodate changes in $r$ to keep $D$ constant and ignoring out-of-steady-state dynamics. I can also drop equation 2, since it is not necessary to determine the steady steady, and I can thus simplify the other four equations into three equations in $K_H$, $K_F$ and $r$, and ultimately into one equation in $r$ as a function of $D$, as below, also assuming preferences are the same in both countries.

**Proposition 2.** The steady state under both Autarky and Integration is described by the
following equations for the capital stock and bond market equilibrium in terms of the interest rate(s):

\[ \forall i = H, F \quad K_i(r) = \left( \frac{\alpha}{r + \delta + \sigma_i \max \{\mu(r), \mu^c(r)\}} \right)^{\frac{1}{1-\alpha}} \] (9)

where \( \mu(r) \equiv \sqrt{\frac{2\gamma}{1+\gamma} (\rho - r)} \) and \( \mu^c(r) \equiv \frac{\rho - r}{\sigma_i \lambda_i} - \frac{\gamma(\gamma-1)}{2} \sigma_i \lambda_i \).

Bond market clearing under Autarky:

\[ \forall i = H, F \quad B_i(r_i) = D_i \implies \left( \frac{\gamma \sigma_i^2}{\alpha K_i(r_i)^{\alpha-1} - \delta - r_i} - 1 \right) K_i(r) = D_i \]

Bond market clearing under Integrated financial markets:

\[ B_H(r) + B_F(r) = D_H + D_F \implies \max \left\{ \frac{\gamma \sigma_H}{\mu(r)} - 1, -\theta_H \right\} K_H(r) + \max \left\{ \frac{\gamma \sigma_F}{\mu(r)} - 1, -\theta_F \right\} K_F(r) = D_H + D_F \] (10)

Proof. Follows immediately from the steady state of the dynamic equilibrium equations. \( \square \)

Note that \( \mu(r) \) and \( \mu^c(r) \) are also the Sharpe ratios, \( \mu = \frac{R-r}{\sigma_i} \), when credit constraints don’t and do bind, respectively. Note also that under Autarky credit constraints cannot bind, provided public debt is non-negative.

**Does a steady state exist and is it unique?** In the long run, the government has one choice to make: how much debt to issue, or how big a deficit to run; not both. With the choice of one, the steady state budget constraint pins down the other, \( D = \frac{\xi}{r} \). However, although any equilibrium can be attained by either choice, they are not equivalent. Targeting the long-run level of debt is much more likely to pin down a unique steady state; one is guaranteed to exist under very mild conditions, and it is very likely to be unique, as I show below. In targeting the deficit a steady state will almost certainly not be unique, and may not exist at all. Thus I assume henceforth that governments target the long-run level of debt.
In the absence of credit constraints, the existence of a steady state – whether under Autarky or Integration – is relatively easy to prove, and a mild sufficient condition guaranteeing existence is that investment in at least one country is sufficiently safe:

**Theorem 2.** Under both Autarky and Integrated financial markets, in the absence of credit constraints a steady state for a given level of global public debt exists when \( \exists i : \rho + \delta > \frac{\sigma_i^2}{2} \gamma (\gamma - 1) \).

**Proof.** See Appendix.

If this condition is not satisfied, a steady state does not exist for arbitrary levels of public debt, but does exist – in fact two exist – for sufficiently high levels. In characterising the dynamics of the model later on, it will become apparent that the economy will tend towards these high-debt steady states.

Uniqueness of the steady state is considerably harder to establish, even without credit constraints. A sufficient condition is that bondholdings are monotonically increasing in the interest rate, which is likely to be the case when the above condition holds, though hard to prove. Another sufficient but onerous condition under Integrated markets is as follows:

**Corollary 1.** Under Integrated financial markets, in the absence of credit constraints a steady state for a given level of global public debt is unique when one country is a net borrower, i.e. \( \exists i : r^* < \rho - \frac{\sigma_i^2}{2} \gamma (1 + \gamma) \).

**Proof.** See Appendix.

In the presence of credit constraints, proving the existence of a steady state is more challenging; as the below calibrations suggest, as constraints bind more tightly the steady state may no longer be unique (e.g. when \( \lambda_H = 1.25 \) in the calibration below), or may not exist at all (e.g. when \( \lambda_H < 1.2 \), not shown).

Figure 2 below shows some examples of the possible existence and uniqueness or otherwise of steady state in the model. The figure shows aggregate bond holdings in each country.
(Home – the developed market – in blue, and Foreign – the emerging market – in red), and globally (black). The horizontal dashed line represents global public debt; here it is zero, but adjust it up or down freely in your head. For more reasonable calibrations – in particular when investment is not too risky, as specified in Theorem 2 – aggregate bond holdings in a country are (monotonically) increasing in the risk-free interest rate. When this condition is violated, bond holdings are U-shaped. The intersection between each coloured line and the dashed line is the Autarkic steady state in each country; the intersection between the thick and dashed black lines the Integrated steady state. When they do not intersect – as may happen when the condition in Theorem 2 is violated – there is no steady state equilibrium at the chosen level of public debt.

The calibration in Figure 2 is such that parameters are equal in both countries, with $\alpha = 0.36$, $\gamma = 2$, $\rho = \delta = 0.06$ and $\lambda = 1.5$, except $\sigma_F = 0.4$ and $\sigma_H = 0.1$, unless otherwise stated in the figure. Existence of steady state thus requires (ignoring credit constraints for one moment) that for at least one country, $\sigma_i < \sqrt{\rho + \delta}$ $\simeq 0.346$. $\sigma_F$ exceeds this threshold in all three panels; $\sigma_H$ is below it in the first two (and credit constraints are not so strict) so that at least one steady state exists, but exceeds it in the last, so no steady state exists, except for high levels of public debt.
Figure 2: Examples when a steady state exists and is unique, is not unique, or does not exist

Note however that the uniqueness of the steady state here is partly an artifact of imposing $\dot{D} = 0$ before specifying the dynamics of debt and hence ignoring its out-of-steady-state dynamics. In specifying those dynamics, one must take a stand on the “target” steady state – since this determines tax revenues and hence the deficit/surplus – which will pin down a $\dot{D} = 0$ locus through the phase space; the intersection of this locus with the other loci will then determine the steady state. When I come to discuss dynamics later on, it will become apparent that even when the condition in Theorem 2 holds, the steady state may not be unique.

Let us briefly consider the alternative of targeting the long-run deficit/surplus, with the level of debt determined endogenously as $D = \frac{s}{r}$. In this case, steady state is not guaranteed to either exist or be unique, even when global bond holdings are “nicely” behaved, i.e. monotonically increasing in the interest rate. As Figure 3 below suggests, targeting a surplus is consistent with two steady states, one with a positive level of debt and another with negative public debt; a small deficit is also consistent with two steady states, both with positive debt; but for a large deficit no steady state exists. These steady states correspond
to the same ones that exist when choosing the corresponding level of debt, \( D = \frac{\delta}{r} \), but may have different stability properties, which suggests caution for governments that target the deficit. Bad shocks could lead the economy to converge on another steady state with lower output and/or welfare, and a miscalculation of the amount of fiscal space available, or a permanent change in the degree of risk, could lead to no steady state existing at all, with the economy spiraling towards collapse or hyperinflation until a different fiscal rule is chosen. Note Figure 3 applies in Autarky, or under Integration when both countries follow the same debt- or deficit-targeting strategy. When one country targets the deficit but another the debt, the equilibrium will still look like the panel on the right, but with the hyperbola shifted up by the debt target of the second country.

**Figure 3: Steady state may not exist or be unique when targeting the deficit**

Crowding Out or Crowding In? In a standard representative agent Ramsey model with non-distortionary taxation, Ricardian equivalence holds, so issuance of debt does not by itself affect equilibrium and so does not crowd out capital. Here however, public debt may either crowd out or crowd in capital. Consider the locus of the capital stock as a function of the interest rate given in Proposition 2. The point on this curve at which the economy
rests is given by the interest rate that ensures bond market equilibrium. If bondholdings are monotonically increasing in the interest rate, then issuing more public debt results in a higher interest rate. Whether or not this crowds out capital depends on whether the equilibrium lies in the downward-sloping or upward-sloping region of the curve. The inflection point is reached when \( \tilde{r}_i = \rho - \frac{1}{2} \frac{\gamma}{1+\gamma} \sigma_i^2 \); above this interest rate, issuing more debt crowds in capital. There is thus a level of public debt above which issuing more debt crowds in capital, although for sensible calibrations this level is so high – several multiples of GDP – that we can reasonably ignore it.

**Autarky.** Under Autarky, and with no government debt, there is no net borrowing, so entrepreneurs must allocate all their wealth to capital: \( \omega_i = \frac{R_i - r_i}{\gamma \sigma_i^2} = 1 \). Combining with equation 9, this yields an equilibrium interest rate of \( r_{i,ND}^{A} = \rho - \frac{\gamma(1+\gamma)}{2} \sigma_i^2 \), which note puts the economy necessarily to the left of the inflection point, \( \tilde{r}_i \), so issuing debt from this point certainly crowds out capital. In the presence of government debt, the equilibrium interest rate becomes \( r_{i,D}^{A} = \rho - \frac{\gamma}{1+\gamma} (1-\psi)^2 \sigma^2 \), where \( \psi = \frac{D}{K+D} \) denotes the share of government debt in total wealth, which clearly tends to the no-debt interest rate when \( \psi = 0 \) and to \( \rho \) in the limit where government debt comprises all wealth (\( \psi \to 1 \)). Note then that the economy reaches the crowding out-crowding in inflection point when \( \tilde{r}_i = r_{i,D}^{A} \), which occurs when \( \psi = \frac{\gamma}{1+\gamma} \), so for example when \( \gamma = 1 \), government debt must account for half of total wealth, or on the order of three times output if the capital stock is also three times output, as in the data.

We can see this graphically. We can rewrite equation 9 in terms of the capital/output ratio, \( \frac{K}{Y} (r) = \frac{\alpha}{r+\delta+\sigma \mu(r)} \). We can also derive a second equation for the capital stock, conditional on bond market equilibrium for a given level of public debt: \( \frac{K}{Y} (r) = \frac{\alpha}{r+\delta+(1-\psi)\gamma \sigma^2} \). The equilibrium capital stock relative to output under Autarky is given by the intersection of these two curves: \( \frac{K}{Y}^* = \rho+\delta+\sigma^2 ((1-\psi)\gamma - \frac{1}{2} (1-\psi)^2 (1+\gamma)) \), which naturally depends on the level of debt through \( \psi \). Such equilibria are shown in Figure 4: the left panel shows bond market equilibrium for various levels of public debt and the associated levels of capital; the right panel
translates these bond market equilibria into loci of capital, with the intersection denoting the equilibrium. As discussed, issuing debt first crowds out, then crowds in capital beyond a certain point.

Figure 4: Autarky equilibrium: debt crowds out capital, except at high levels

Note that under Autarky and without public debt, the equilibrium capital stock, \( \frac{K}{Y} \), is higher under incomplete markets (\( \sigma > 0 \)) than that under complete markets (\( \sigma = 0 \)) when \( \gamma > 1 \), as the equilibrium lies in the downward-sloping part of the capital locus. This has previously been noted by Angeletos (2007), except there the presence of labour income for entrepreneurs altered the condition under which this holds, so in fact for reasonable calibrations the capital stock was in fact lower under incomplete markets. The contrary result in this paper is somewhat frustrating for the empirical realism of the model, because as Angeletos & Panousi (2011) point out, it means riskier countries – i.e. less developed countries – will have a higher capital stock relative to output than more developed countries under Autarky. In the data the opposite is typically true; richer countries are more capital intensive than poorer ones. The model replicates this stylised fact under Integration, but not under Autarky. Though we do not have particularly good data for countries that are
completely divorced from global capital markets, like North Korea, Somalia or the former Soviet Union, it seems unlikely that such countries would break the international mould and possess higher relative capital stocks. That said, the presence of relatively high public debt in these countries offers a potential route out of the dilemma, and in any case it is questionable that a model based on optimising entrepreneurs operating under free markets should apply to such countries in the first place.

**Integration.** There is unfortunately no closed-form solution for the equilibrium interest rate under Integration. However, we can still characterise the equilibrium: firstly, provided government debt is not too great or too substantially negative, the interest rate under Autarky in the riskier, less-developed country, $r^A_F$, must be lower than that in the developed country, $r^A_H$, for the reasons established above. Secondly, if global bondholdings are well-behaved, i.e. monotonically increasing in the interest rate, then the equilibrium interest rate under Integration must lie between the two interest rates that would prevail under Autarky, $r^A_F < r^I < r^A_H$. Consequently, financial Integration must lower the equilibrium capital stock in the developing country and raise it in the developed country, again provided both countries are in the downward-sloping part of the capital locus. Again, the obverse of this result was noted in Angeletos & Panousi (2011), where their calibration places both economies on the upward-sloping part of the capital locus. This emphasises the fragility of certain properties of equilibrium in this model, but I do not focus on these and do not rely on them in the subsequent discussion on fiscal space.

The presence of two countries now means that it is possible for one country to be a net borrower. For example, without government debt, if Foreign risk is greater than domestic risk then Home entrepreneurs will be net borrowers, since $\omega_H > 1$ requires that the equilibrium interest under financial Integration be strictly less than that under the Autarky $r^{I,ND} < r^{A,ND}_H = \rho - \frac{\gamma(1+\gamma)}{2} \sigma^2_H$. Equivalently, this requires that the Foreign Autarky, No-Debt equilibrium interest rate be below that of Home. Given that the equilibrium interest rate under Integration lies between the Autarky rates, for Home entrepreneurs to be net borrowers
requires simply that $\sigma_H < \sigma_F$. In Figure 5 below I show such an equilibrium: the dashed lines represent aggregate bond holdings, the thick lines the capital stock; the blue lines refer to the developed market, red to the emerging; the thick black line again represents aggregate global bond holdings. For simplicity I assume there is no public debt in either country, so the equilibrium interest rates occur when bond holdings cross the zero horizontal.

Figure 5: Integrated equilibrium with no public debt when $\sigma_H < \sigma_F$

Now that a country can be a net borrower, it is possible for credit constraints to bind. From the steady state Euler equation, one can immediately see that when credit constraints bind the capital stock is strictly lower than it would be in their absence. From the steady state bond market clearing condition, one can see that in the absence of net global government debt ($D_H + D_F = 0$) and with $\sigma_H \neq \sigma_F$, bond holdings can be positive in at most one country; entrepreneurs in one country are net debtors, those in the other net creditors. With positive net public debt, it is possible for entrepreneurs in both countries to be net lenders, but clearly not both net borrowers. Consequently credit constraints will bind in at most one country.
Testing the Model against the Data

The key empirical fact this model is designed to explain is persistent negative real interest rates, and in particular the decline in the real interest rate over the last 20 years. This it naturally does if one considers the growth of China and other emerging markets over that period to be a transition from an equilibrium in which the developed world was essentially in ‘autarky’ – or effectively so given the de minimis wealth held by agents in those emerging economies – to one in which the developed and emerging markets are integrated (however imperfectly) and more balanced in size. The model naturally makes other empirical predictions, which we can test against the data, for which I use the Penn World Tables, version 10.0 (Feenstra et al. 2015), and the World Inequality Database (Chancel et al. 2021).

The most obvious empirical prediction is the global imbalances of the title – that in equilibrium capital flows from developing countries to developed countries. This empirical fact is well-established and indeed was the primary motivation for Angeletos & Panousi (2011). The image below is taken directly from Gourinchas & Jeanne (2013) and shows capital flows against productivity growth over 1980-2000 for 68 non-OECD countries. A few other quasi-anecdotal data points also lend credence to the theory, for example that China is the largest foreign holder of US Treasuries.

However, it should be noted that a related prediction of the model is that rich countries should have trade and net foreign asset deficits, and poor countries surpluses; while this holds for the key motivating examples like the US, UK and China, it does not hold across all countries, with the Germanic countries and oil-rich states being notable exceptions. Moreover, with long-run growth, on a balanced growth path net foreign assets would remain constant as a share of output, so the net lender country (the developing country) would have growing net foreign assets, i.e. would run current account surpluses, while the net borrowing country (developed country) would run current account deficits; this again is true for the US, UK and China, but is at odds with the cross-country correlation of current account balances. I abstract from growth in the model, so current account balances are zero everywhere in steady
The second major empirical prediction of the model is that riskier, less developed countries will have lower capital-output ratios than do safer, developed countries in an equilibrium with globally integrated capital markets. This too we observe in the data (Caselli 2005), and this cannot be explained for example by standard growth theories that explain income differences between countries on a balanced growth path solely through differences in total factor productivity. A positive correlation between incomes and capital intensity is also a prediction of models featuring investment-specific technological growth, but such models also predict lower returns on capital in less developed countries, while my idiosyncratic risk/global imbalances model predicts the opposite; the data is patchier here, and the correlation is weak, but does somewhat support the imbalances hypothesis, although see Gourinchas & Jeanne (2013) and Caselli & Feyrer (2007) for arguments against this view.

I see the risk/global imbalances hypothesis and the OLG/demographics hypothesis as very much complementary explanations for the recent fall in real interest rates. Nevertheless, there
are some revealing patterns in the data that can help us discriminate between the two. The OLG/demographics hypothesis does not to my understanding imply anything in particular about the degree of overall wealth inequality either within or between countries. The global imbalances model however predicts a Pareto right-tail for the wealth distribution, which under integrated bond markets should be identical in both countries when capital is unconstrained; however, when rich countries are borrowing-constrained, the model predicts they should have lower wealth inequality than poorer countries. This is largely borne out by the data, although the correlation is only weakly significant, and there are a few notable exceptions such as the US and several Middle East countries.

Figure 7: Richer countries have higher capital intensity and lower wealth inequality

Source: GDP per worker is the output-side real GDP at current PPPs from the PWT 10.0, divided by employment. The capital/output ratio is the capital stock at current PPPs, divided by the same measure of output, also from the PWT 10.0. The top 1% wealth share is from the World Inequality Database.
4 The Limits of Fiscal Space

It is convenient to express equation 10, which characterises bond market equilibrium solely as a function of the interest rate, in terms of Home Debt/GDP:

\[
\frac{D_H}{Y_H} = \max \left\{ \frac{\gamma \sigma_H}{\mu(r)} - 1, -\theta_H \right\} \frac{\alpha}{r + \delta + \sigma_H \max \{\mu(r), \mu_H'(r)\}} + \left[ \max \left\{ \frac{\gamma \sigma_F}{\mu(r)} - 1, -\theta_F \right\} \frac{\alpha}{r + \delta + \sigma_F \max \{\mu(r), \mu_F'(r)\}} - \frac{D_F}{Y_F} \right] \cdot \left( \frac{r + \delta + \sigma_H \max \{\mu(r), \mu_H'(r)\}}{r + \delta + \sigma_F \max \{\mu(r), \mu_F'(r)\}} \right)^{\frac{\alpha}{1 - \alpha}}.
\]

We can use this equation to determine the equilibrium steady state interest rate, ascertain the deficit/surplus that a government must run to maintain that steady state, and hence establish the limits of fiscal space available to governments. This allows us to trace out a locus of deficits for any given level of public debt, which is similar in spirit to what Brunnermeier et al. (2021a) call the “Debt Laffer Curve”, a name which I borrow, although which in their paper refers to a relationship between deficits and issuance. Figure 8 shows the Debt Laffer Curve for a country under Autarky for three levels of idiosyncratic risk. With low risk, precautionary saving is relatively weak, and the equilibrium interest rate is positive except at very low levels of public debt, hence there is no bubble. With higher risk, precautionary saving is stronger, equilibrium interest rates are lower and there is a large bubble, affording the government considerable fiscal space. One useful notion of fiscal space is the level of debt (relative to GDP) at which the interest rate is zero, and hence below which there is a bubble, allowing the government to run a primary deficit indefinitely. I call this the “Bubble Bound”:

**Definition 3.** The Bubble Bound, denoted \( D_i \), is the maximum sustainable level of public debt below which a government can run a perpetual deficit, i.e. where \( r = 0 \).

Another useful notion of fiscal space is the maximum primary deficit (relative to GDP) that a government can sustain. I call this the “Profligacy Peak”:
Definition 4. The Profligacy Peak, denoted \(-\overline{s}_i\), is the maximum sustainable primary deficit:

\[-\overline{s}_i = \max_{D_H} -rD_H \quad s.t. \quad B_H(r) + B_F(r) = D_H + D_F\]

Note that at the Profligacy Peak, the elasticity of the interest rate to public debt equals negative 1, \(\partial D_H \left[ -r(D_H)D_H \right] = 0 \iff \frac{d\mu}{dD_H} \frac{D_H}{r} = -1.\)

In Figure 8 I also highlight the Bubble Bound and Profligacy Peak for the different levels of risk.

Figure 8: The Debt Laffer Curve under Autarky

Without Credit Constraints

We can examine the case without credit constraints more closely and achieve some useful analytical benchmark results. The single equilibrium equation 11 becomes:

\[
\frac{D_H}{Y_H} = \left( \frac{\gamma \sigma_H - 1}{\mu(r)} \right) \frac{\alpha}{r + \delta + \sigma_H \mu(r)} + \left[ \left( \frac{\gamma \sigma_F - 1}{\mu(r)} \right) \frac{\alpha}{r + \delta + \sigma_F \mu(r)} - \frac{D_F}{Y_F} \right] \left( r + \delta + \sigma_H \mu(r) \right)^{1-\alpha}
\]

Lemma 2. Under Autarky, the Bubble Bound of debt/GDP, below which the government can
sustain positive primary deficits, is given by:

\[ \frac{D_i^A}{Y_i} = \left( \frac{\gamma \sigma_i}{\mu_0} - 1 \right) \frac{\alpha}{\delta + \sigma_i \mu_0} \]

where \( \mu_0 \equiv \mu(0) = \sqrt{\frac{2\rho}{1+\gamma}}. \)

Lemma 3. Under Financial Integration and without credit constraints, the Bubble Bound of debt/GDP is given by:

\[
\frac{D_{H,Y}^{1,UC}}{Y_H} = \left( \frac{\gamma \sigma_H}{\mu_0} - 1 \right) \frac{\alpha}{\delta + \sigma_H \mu_0} + \left[ \left( \frac{\gamma \sigma_F}{\mu_0} - 1 \right) \frac{\alpha}{\delta + \sigma_F \mu_0} - \frac{D_F}{Y_F} \left( \frac{\delta + \sigma_H \mu_0}{Y_F} \right) \right] \frac{\delta + \sigma_H \mu_0}{Y_F} \]

Proof. These Lemmas follow immediately from equation 11 after setting \( r = 0 \) and rearranging.

One immediate consequence of this is that the Bubble Bound is higher under Integration than under Autarky when Foreign risk is sufficiently high and when Foreign public debt is sufficiently low:

Theorem 3. The Bubble Bound is higher under Integration than under Autarky when Foreign public debt is below its own Autarkic Bubble Bound. When Foreign governments have no debt, this holds when overseas investors are net lenders, i.e. when \( \sigma_F^2 \geq \frac{2\rho}{\gamma(1+\gamma)}. \)

Proof. The first part follows immediately from Lemma 3. When \( \frac{D_F}{Y_F} = 0, \frac{D_{H,Y}^{1,UC}}{Y_H} \geq \frac{D_i^A}{Y_i} \) if \( \frac{\gamma \sigma_F}{\mu_0} - 1 > 0, \) or \( \sigma_F^2 \geq \frac{2\rho}{\gamma(1+\gamma)}. \)

However, even if this result holds and the Bubble Bound is higher under Integration, the maximum sustainable deficit may still be lower, as shown in Figure 9 below. For the maximum deficit to be higher under Integration than under Autarky, Foreign risk must be higher still than the condition above.
A close corollary of this result is that the Bubble Bound is increasing in both foreign and domestic risk.

**Corollary 2. The Bubble Bound is typically increasing in Foreign risk, \( \sigma_F \).** A sufficient condition under which this is true is that Foreign entrepreneurs are net borrowers, i.e. when \( \sigma_F^2 \leq \frac{2\rho}{\gamma(1+\gamma)} \). A necessary condition is that \( \sigma_F \leq \left( \frac{\delta(1-\alpha)}{\mu_0} + \frac{\mu_0}{\gamma} \right) \frac{1}{\alpha} \), which holds for all reasonable parameter values, even if Foreign investors are net lenders.

**Proof.** The derivative of the bubble condition with respect to the degree of Foreign risk is:

\[
\frac{\partial}{\partial \sigma_F} \left( \frac{D_H}{Y_H} \right) = \frac{\gamma \alpha}{\mu_0} \left( \delta + \sigma_H \mu_0 \right)^{\frac{1}{1-\alpha}} \left( \delta + \sigma_F \mu_0 \right)^{\frac{1}{\alpha-1}}
\]

\[
+ \left( \frac{\gamma \sigma_F}{\mu_0} - 1 \right) \frac{\alpha}{\alpha - 1} \left( \delta + \sigma_H \mu_0 \right)^{\frac{1}{1-\alpha}} \left( \delta + \sigma_F \mu_0 \right)^{\frac{1}{\alpha-1}-1} \mu_0 > 0
\]

\[
\iff \frac{\gamma}{\mu_0} + \left( \frac{\gamma \sigma_F}{\mu_0} - 1 \right) \frac{1}{\alpha - 1} \left( \delta + \sigma_F \mu_0 \right)^{-1} \mu_0 > 0
\]

which is positive when either of the conditions hold.

**Corollary 3. Under Autarky, the Bubble Bound is always increasing in domestic risk, \( \sigma_H \);**
under Integration, a sufficient condition for this is when domestic investors are net borrowers, i.e. when $\sigma^2_H \leq \frac{2\rho}{\gamma(1+\gamma)}$, and Foreign public debt is below its Autarkic Bubble Bound.

Proof. For the first part:

$$\frac{\partial}{\partial \sigma_H} \left( \frac{D_H}{Y_H} \right) = \frac{\gamma}{\mu_0} \alpha (\delta + \sigma_H \mu_0)^{-1} + \left( \frac{\gamma \sigma_H}{\mu_0} - 1 \right) \alpha (\delta + \sigma_H \mu_0)^{-2} \mu_0 > 0$$

$$\iff \frac{\gamma}{\mu_0} + \left( \frac{\gamma \sigma_H}{\mu_0} - 1 \right) (\delta + \sigma_H \mu_0)^{-1} \mu_0 > 0 \iff \frac{\delta}{\mu_0} + \frac{\mu_0}{\gamma} > 0$$

which always holds. For the second part:

$$\frac{\partial}{\partial \sigma_H} \left( \frac{D_H \gamma^{1,UC}}{Y_H} \right) = \frac{\gamma}{\mu_0} \alpha (\delta + \sigma_H \mu_0)^{-1} + \left( \frac{\gamma \sigma_H}{\mu_0} - 1 \right) \alpha (\delta + \sigma_H \mu_0)^{-2} \mu_0$$

$$+ \left[ \frac{\gamma \sigma_F}{\mu_0} - 1 \right] \frac{\alpha}{\delta + \sigma_F \mu_0} - \frac{D_F}{Y_F} \right] \frac{\alpha \mu_0}{1 - \alpha} \left( \frac{\delta + \sigma_H \mu_0}{\delta + \sigma_F \mu_0} \right)^{\frac{2\alpha - 1}{1 - \alpha}} > 0$$

which is certainly positive when $\frac{\gamma \sigma_H}{\mu_0} \geq 1$, or $\sigma^2_H \leq \frac{2\rho}{\gamma(1+\gamma)}$, and $\frac{D_F}{Y_F} \leq \left( \frac{\gamma \sigma_F}{\mu_0} - 1 \right) \frac{\alpha}{\delta + \sigma_F \mu_0}$, although this is certainly not a necessary condition and the result will likely hold even when domestic investors are not net borrowers. □
The Profligacy Peak is the second key notion of fiscal space. As I have shown above quantitatively, whether or not the maximum sustainable deficit rises or falls as a result of integration is more nuanced, and the concept escapes formal theorems as far as I can tell. However, it is clear quantitatively that the Profligacy Peak is higher under Integration when Foreign risk is sufficiently high, as shown in Figure 11 below, and moreover is increasing in both Home and Foreign risk.
With Credit Constraints

As discussed above, given that entrepreneurs in a given country are ex-ante identical, and because of the linearity of their policy functions, all entrepreneurs behave as scaled-up or down versions of one another. Consequently, entrepreneurs in a given country will be either all borrowers or all lenders, depending on the risk they face and the risk faced by entrepreneurs in the other country. As such, credit constraints will bind on entrepreneurs in at most one country at a time, and if they do bind they will bind for all entrepreneurs in the country. The condition for credit constraints to be binding is

$$\frac{\alpha K_{H}^{\gamma_{H}}}{\gamma_{H}} - \delta - r \geq \lambda_{i}, \text{ or } r < \rho - \frac{\gamma(1+\gamma)}{2} (\sigma_{i} \lambda_{i})^{2},$$

which note certainly puts the equilibrium in the downward-sloping part of the capital locus.

With credit constraints, the expression for the Bubble Bound is as follows:

**Lemma 4.** Under Financial Integration and with credit constraints, the debt/GDP Bubble Bound is given by:

$$\frac{D_{H}}{Y_{H}}^{I, CC} = \frac{\alpha \max \left\{ \frac{\gamma_{H}}{\mu_{0}} - 1, -\theta_{H} \right\} + \left[ \alpha \max \left\{ \frac{\gamma_{F}}{\mu_{0}} - 1, -\theta_{F} \right\} \right]}{\delta + \sigma_{F} \max \left\{ \mu_{0}, \mu_{F_{0}} \right\}} - \frac{D_{F}}{Y_{F}} \left( \frac{\delta + \sigma_{H} \max \left\{ \mu_{0}, \mu_{H_{0}} \right\}}{\delta + \sigma_{F} \max \left\{ \mu_{0}, \mu_{F_{0}} \right\}} \right)^{\frac{1}{\alpha}}.$$
Proof. This again follows immediately from equation 11 after setting \( r = 0 \) and rearranging.

The Bubble Bound is now no longer generally increasing in Home or Foreign risk, but as the following theorem shows, when credit constraints bind the Bubble Bound is higher than in their absence.

**Theorem 4.** The Bubble Bound in the presence of credit constraints is at least as large as it is without credit constraints, provided Foreign public debt is below its Autarkic Bubble Bound.

**Proof.** Taking without loss of generality the case where constraints bind at Home, so Home entrepreneurs are net borrowers \((\gamma \sigma_H < \mu_0)\) and Foreign net lenders \((\gamma \sigma_F > \mu_0)\). Comparing the Bubble Bounds under Integration with and without credit constraints:

\[
\frac{D_{H/CC}}{Y_H} \geq \frac{D_{H/UC}}{Y_H} \iff \max \left\{ \frac{\gamma \sigma_H}{\mu_0} - 1, -\theta_H \right\} \frac{\alpha}{\delta + \sigma_H \max \{\mu_0, \mu_{H0}^c\}} + \left[ \left( \frac{\gamma \sigma_F}{\mu_0} - 1 \right) \frac{\alpha}{\delta + \sigma_F \mu_0} - \frac{D_F}{Y_F} \right] \left( \delta + \sigma_H \max \{\mu_0, \mu_{H0}^c\} \right)^{\frac{\alpha}{1-\alpha}} \geq \left( \frac{\gamma \sigma_H}{\mu_0} - 1 \right) \frac{\alpha}{\delta + \sigma_H \mu_0} \max \left\{ \frac{\gamma \sigma_H}{\mu_0} - 1, -\theta_H \right\}
\]

The second term on the left hand side on of the inequality is clearly at least as great as that on the right due to the presence of the max operator in the numerator (the \( \mu \)s are certainly positive, being the Sharpe ratios, as some light algebra demonstrates). The first term on the left hand side is also at least as great as that on the right, to see this note the following (and recall that \( \frac{\gamma \sigma_H}{\mu_0} - 1 < 0 \) by assumption):

\[
\max \left\{ \frac{\gamma \sigma_H}{\mu_0} - 1, -\theta_H \right\} \frac{\alpha}{\delta + \sigma_H \max \{\mu_0, \mu_{H0}^c\}} \geq \left( \frac{\gamma \sigma_H}{\mu_0} - 1 \right) \frac{\alpha}{\delta + \sigma_H \mu_0} \]

A similar logic proves the case when constraints bind for Foreign entrepreneurs.

\(\square\)
With credit constraints it is again difficult to derive explicit theorems regarding the size of the maximum sustainable deficit – the Profligacy Peak – but the quantitative evidence below suggests it is higher when constraints bind, as in the “wings” of Figure 13 below. When one country is very safe and the other very risky, credit constraints bind, push the equilibrium interest rate down and allow the government to run a perpetual deficit on the order of 1% of GDP, when a balanced budget would have been required in the absence of constraints.
The role of credit constraints. It is important to note that the mechanism at play here by which credit constraints affect the equilibrium is fundamentally different from that in Kiyotaki & Moore (1997). In that paper, a negative productivity shock lowers the price of land/capital, which in turn tightens the borrowing constraint and lowers productive capacity further, amplifying the contraction. Here, credit constraints on DM entrepreneurs limit the competition for safe assets; forcing savings to be redirected to government bonds rather than to private borrowers, lowering the interest rate and expanding the bubble.

The role of credit constraints here also differs slightly from Reis (2021). In his paper entrepreneurs have different qualities, with higher quality entrepreneurs having both higher marginal product and also lower (no) idiosyncratic risk. Without credit constraints, the high quality entrepreneurs will borrow from low quality ones and will ultimately dominate production, effectively completing markets. Thus the interest rate on bonds will equal the marginal product of capital, which exceeds the growth rate, eliminating the bubble in government debt and thus fiscal space – as measured above – disappears. Credit constraints therefore serve to stop the concentration of production among the high quality entrepreneurs.
What my paper distinctively shows is that even if borrowing cannot effectively complete markets, credit constraints still lower the equilibrium risk-free real interest rate – and expand the bubble – simply by restricting the supply of safe assets.

Credit constraints also limit the crowding out effects of public debt issuance in the country when they bind, and may exacerbate them elsewhere. Figure 14 below shows, for an example calibration when Home entrepreneurs are credit constrained, the equilibrium interest rates and capital stocks under financial integration before and after an issuance of public debt on the order of 80% of GDP. The equilibrium moves from clearly capital-constrained to just at the kink where constraints are no longer binding. This entails a loss of output on the order of 4% for the Home country but 8% for the Foreign country. Without credit constraints, the losses would be closer to 6% domestically and 4% overseas.

Figure 14: Issuance of public debt when entrepreneurs are credit constrained

The calibration required to generate the above figure is not necessarily realistic, with a maximum loan-to-value (LTV) ratio for entrepreneurs of only around $\theta_H = 0.2$. To this I have two responses:

1. Firstly the model is highly stylised, and a more realistic model featuring, for example,
stochastic idiosyncratic entrepreneurial productivity, similar to the work of Quadrini (2000) and Cagetti & De Nardi (2006), would feature binding credit constraints with realistic LTV ratios for sufficiently highly skilled entrepreneurs. This notwithstanding the critique that such collateral constraints may not be quantitative important from a macroeconomic perspective, e.g. Kocherlakota (2000) and Cordoba & Ripoll (2004), or that alternative credit constraints on profitability or cash flow seem to be more prevalent and binding more frequently in the data, as in the work of Drechsel (2021).

2. Secondly, the appropriate moment condition with which to calibrate θ is arguably not the maximum loan-to-value at which commercial banks are willing to offer loans to small businesses secured on physical capital. Instead the constraint reflects limits on banks’ ability to transform risky loans into risk-free securities; entrepreneurial debt is completely risk-free in the model and thus are a private source of supply of safe assets. Such limits arise from informational asymmetries and moral hazard and are expounded upon in BMS (2021b). When constraints bind at Home in the model, the total private supply of safe assets is \( \theta_H \frac{K_H}{Y_H} \); the total supply of safe assets by US financial institutions hovers around 60–75% of GDP (Gourinchas & Jeanne 2012), which with \( \frac{K_H}{Y_H} \simeq 3–4 \) implies \( \theta_H \simeq 0.2 \).

Note that issuance of public debt by the Foreign country lowers the Bubble Bound and Profligacy Peak for the Home country. Consider an example where both countries are identical and Home has a Bubble Bound of 2x GDP when there is no Foreign public debt. As Foreign public debt rises, however, the Bubble Bound falls. If both countries tried to exploit the bubble fully, conditional on the actions of the other, their Bubble Bound would be perhaps 1x GDP, the same as in Autarky. Thus when both countries are trying to exploit the bubble, Integration does not necessarily confer an advantage over Autarky; it allows one country to mine another country’s bubble, which is advantageous if their bubble is bigger than yours and they do not or cannot exploit it, but does not by itself increase the combined size of the two bubbles. However, this is only true when credit constraints don’t bind in equilibrium; if they
do, the Integrated bubble is larger than the two Autarky bubbles combined. One immediate implication of Theorem 4 is that the total amount of debt that can be sustained with zero interest rates under Integration (a global Bubble Bound, if you like) is higher when credit constraints are present and bind than when they would bind but are not present, and hence higher than that which can be sustained by the two countries combined under Autarky. The example in Figure 15 shows such a scenario.

Figure 15: Foreign debt issuance limits the ability to mine foreign bubbles

5 Welfare and the Optimal Quantity of Public Debt

I now turn to welfare and the optimal quantity of government debt. I consider only steady state welfare, with constant debt policies, rather than allowing for arbitrary time-varying debt policies, for ease of comparison with the above sections and with the prior literature. Optimal debt/GDP will in general be much higher than the Bubble Bound or Profligacy Peak, because these hinge on exploiting the bubble in public debt, but the welfare calculations trade off the taxation of workers – which has limited costs here due to taxes being non-distortionary – against the income and self-insurance benefits of public debt for entrepreneurs, which are
considerable.

One challenge of assessing the welfare at the aggregate level in this model is that the stochastic process that governs entrepreneurs’ wealth is a geometric Brownian motion – a random walk in logs – which does not have a stationary distribution. The distribution of wealth among entrepreneurs is thus lognormal and increasing in variance over time. This is not a problem for analysing macroeconomic aggregates, since the linearity of entrepreneurs’ policy functions implies the distribution of wealth does not matter for aggregates, but it does matter for welfare.

The typical trick in these circumstances is to add a stabilising force to limit this divergence, for example Poisson death shocks, which I use here, with newborns inheriting some lower level of wealth, e.g. the average. This gives rise to a stationary double-Pareto distribution in wealth, as discussed in for example Gabaix (2009) and Gabaix et al. (2016). I do not microfound this inheritance process, but one could explicitly model a bequest motive and inheritance taxes to generate such a process, as in Benhabib et al. (2016).

Note that the parameter $\rho$ above should now be interpreted as a composite $\rho = \tilde{\rho} + p$, where $\tilde{\rho}$ is the actual subjective time discount rate, and $p$ is the Poisson death rate of entrepreneurs, which I calibrate to $1/50$, for an average 50-year lifespan. In addition, I set $\gamma = 1$ to ensure both that a steady state exists and that the average welfare of entrepreneurs is finite. Consequently, total social welfare in steady state is given by the following Proposition:

**Proposition 3.** Steady state welfare in country $i$ is described by a Utilitarian welfare function, with weight $\varpi$ on the welfare of the representative worker and $1 - \varpi$ on the welfare of entrepreneurs:

$$W_i = \varpi V^w_i + (1 - \varpi) \int_0^\infty V^e_i(a) g_i(a) da$$

$$= \frac{1}{\rho} \left[ \ln \left( (1 - \alpha) K_i(r)^\alpha - r D_i \right) - \frac{1}{1 + \varphi} \right] + (1 - \varpi) \frac{m_i(r)^{\gamma - 1}}{1 - \gamma} \left( K_i(r) + B_i(r) \right)^{1 - \gamma} Z(r)$$

where $g_i(a)$ is the density of the stationary wealth distribution, $Z(r) = \frac{\zeta_i + \zeta_i - (1 - \gamma - \zeta_i) - (1 - \gamma - \zeta_i)}{(1 - \gamma - \zeta_i) - (1 - \gamma - \zeta_i)}$ and
where $\zeta_i = \frac{1}{2} \pm \sqrt{\frac{1}{2} + \frac{2p}{(\sigma_i \omega_i)^2}}$ are the Pareto tail parameters, and where I have suppressed notation denoting the dependence of $r$ on $D_i$.

**Proof.** See Appendix.

I use this measure of welfare to find the optimal quantity of public debt. I assume that the entire deficit is used to pay proportional transfers to workers, $-\tau = -\frac{rD}{(1-\alpha)K^\alpha}$ (or taxes if the level of debt necessitates surpluses). If instead we were to assume deficits paid for government spending, which is not valued by agents, there would be less welfare benefit to issuing debt. However, given that much of the persistent deficits highlighted by fiscal authorities of advanced nations over the coming decades will arise due to transfers, like pensions, I believe this is a reasonable benchmark. There are thus several countervailing direct and indirect effects of debt issuance on welfare, corresponding to how $D$ enters into the above directly, and also how it affects $r$ and hence $K$, $B$, $m$ and $Z$.

Issuing more debt has the following direct effect:

- Raises worker consumption through transfers if the deficit grows, $\partial_D [-r(D)D] = -r'(D)D - r > 0$, i.e. $D/Y$ is below the Profligacy Peak; or lowers consumption by raising raxes if not

Provided bondholdings are well-behaved, issuing more debt raises interest rates, which has the following indirect effects:

- Crowds out capital, reducing income for entrepreneurs and wages for workers, unless debt is sufficiently high that the equilibrium is in the upward-sloping part of the capital locus

- Provides more liquidity/risk-free income to entrepreneurs if they are net savers, or reduces their borrowing otherwise, both of which reduce the volatility of their net worth and hence consumption
If the Elasticity of Intertemporal Substitution is not equal to one, $1/\gamma \neq 1$, then issuing more debt also has the following effects:

- Raises the marginal propensity to consume of entrepreneurs, $m$, if $1/\gamma < 1$, and vice versa

- Changes the distribution of wealth among entrepreneurs by fattening the right tail, $\zeta'_+(r) < 0$, and thinning the left tail, $\zeta'_-(r) > 0$, so $Z'(r) > 0$ if $1/\gamma < 1$, and vice versa

**Breaking Ricardian equivalence.** Before quantifying the optimal level of debt I want to first address a curiosity of this model created by the separation of workers and entrepreneurs: Ricardian equivalence does not even without uninsurable risk. To illuminate the matter, consider the special case of autarky under complete markets ($\sigma = 0$), with log utility for both workers and entrepreneurs. The capital stock in steady state becomes $K = (\frac{\alpha}{\rho + \delta})^{\frac{1}{1-\alpha}}$. Consumption of workers becomes $C^w = w(1 - \tau) = (1 - \alpha)K^\alpha(1 - \tau)$ where the tax rate is determined in equilibrium by $D = \frac{\tau w}{r} \rightarrow \tau = \frac{\rho D}{1-\alpha Y}$, taking $\frac{D}{Y}$ as exogenous, chosen by the government; yielding $C^w = (1 - \alpha - \rho\frac{D}{Y})K^\alpha$. I treat worker labour supply as ex-ante inelastic here, for ease of comparison; this is without loss of generality since disutility from labour supply simply acts as a shifter on welfare and does not interact with debt issuance. Because markets are complete, the returns on both capital and bonds must be equal, $R = r = \rho$, so entrepreneurs’ consumption becomes $C^e = RK + rD = \alpha K^\alpha - \delta K + \rho D = \rho(K + D)$. Another way to see this is that $m = \rho$ when $\sigma = 0$.

Welfare in this complete markets special case for this two-agent model is given by:

$$W^{CM,TA} = \varpi \frac{1}{\rho} \ln \left( \left( 1 - \alpha - \rho \frac{D}{Y} \right) K^\alpha \right) + (1 - \varpi) \frac{1}{\rho} \ln \left( \rho \left( K + \frac{D}{Y} K^\alpha \right) \right)$$

One can treat this as the government’s objective function and take first order conditions to
find optimal debt issuance in steady state, and plug in a reasonable calibration to give:

\[ \frac{D^*}{Y} = \frac{(1 - \varpi)1 - \alpha}{\rho} - \varpi \frac{\alpha}{\rho + \delta} \]

\[ = 0.5 \cdot \frac{0.64}{0.06} + 0.5 \cdot 3 \simeq 3.8 \]

However, in the more standard representative agent Ramsey model where workers and entrepreneurs are one and the same, and where proportional labour taxes are non-distortionary now because labour supply is inelastic (lump-sum taxes with elastic labour supply of course gives a similar result), welfare is instead given by:

\[ W_{CM,RA}^{CM,RA} = \frac{1}{\rho} \ln ((1 - \alpha)K^\alpha + \rho K) \]

wherein the two terms relating to public debt cancel, so welfare is independent of public debt issuance, giving the standard Ricardian equivalence result.

It is thus worthwhile noting that Ricardian equivalence does not hold in this two-agent model, even with no distortions and complete markets, simply because of the separation of workers and entrepreneurs. Intuitively, Ricardian equivalence holds in the standard model because agents anticipate that they will have to repay the debt through future taxes, exactly offsetting the income received from holding bonds in present value terms. When the agents receiving the interest income are different from the agents being taxed, the equivalence falls apart.

**What is the welfare-maximising level of debt?** A range of calibrations suggest that under both Autarky and Integration, welfare is maximised at debt levels on the order of several multiples of GDP (4-6x) – far above current levels and far above the levels associated with the Bubble Bound or the Profligacy Peak notions of fiscal space discussed above. This is also substantially higher than optimal debt levels found in previous studies featuring incomplete markets, such as Aiyagari & McGrattan (1998) or Desbonnet & Kankanamge (2016), partly
because these studies feature highly distortionary labour and capital taxation and government spending, while I purposefully neutralise these effects. Figure 16 below shows welfare as a function of the debt/GDP ratio in my model, for given levels of risk, and moreover shows that the welfare-maximising debt level is increasing in both Home and Foreign risk.

An important exception occurs when credit constraints bind. The presence of binding credit constraints increases welfare, since the leverage of entrepreneurs is constrained, and hence so is the volatility of their wealth and thus consumption. As the panel on the right of Figure 16 suggests, when Foreign investment is sufficiently risky, credit constraints bind and raise Home welfare. Public debt issuance in this case may be immediately detrimental to welfare, so zero debt/GDP may be a (local) maximum for welfare.

Of course my paper abstracts from whether or not the government has the ability or will to fulfill such large debt obligations by running large enough surpluses, or market participants’ perceptions thereof. Nevertheless, the model does highlight that public debt itself is usually not harmful to welfare, indeed it is beneficial, except at enormously high debt levels. In other words, the crowding-out effects of public debt are very weak, even though Ricardian equivalence does not hold.

Figure 16: The welfare-maximising level of public debt may be several times GDP
Is debt issuance Pareto-improving? Much of the literature so far has, in the tradition of overlapping generations models, focused on whether intergenerational transfers (via public debt) are Pareto-improving, like Blanchard (2019) or Aguiar et al. (2021). This is related to dynamic inefficiency, which occurs when \( r < g \) in simple models without risk, like Diamond (1965). However, as Reis (2021) has shown, in the presence of risk, \( r < g \) does not suffice to generate dynamic inefficiency, which instead requires that the marginal product of capital be below the growth rate, \( R < g \), which is a much more onerous condition.

In my model however, neither of these conditions are relevant for determining whether debt is Pareto-improving, because of the separation of entrepreneurs and workers. Since workers’ welfare depends only on their consumption (their labour supply being inelastic), and since because they cannot save, their consumption equals their post-tax/transfer wage income; issuing more debt, by crowding out capital and lowering the wage, necessarily makes workers worse off, unless \( r \) is sufficiently negative that the government can run a persistent deficit and use that deficit to fund transfers to workers that are large enough to offset the decline in the wage. As alluded to above, debt must be below the level associated with the Profligacy Peak if transfers are to increase when issuing more debt. It must be further below still and with \( r \) sufficiently negative if the increase in these transfers is to be large enough to offset the decline in the wage from the crowding out of capital.

Figure 17 shows the Pareto frontiers traced out by different equilibria with different levels of government debt for two separate calibrations of the model: one realistic calibration in blue, with \( \sigma_H = 0.3 \) and \( \sigma_F = 0.5 \); and one in red with \( \sigma_H = 0.5 \) and \( \sigma_F = 0.5 \). Only the red line bends sufficiently that the government can increase worker welfare by issuing more debt; these equilibria feature \( r \) in the negative double-digits, though with \( R > 0 \) still. The realistic calibration, which with debt/GDP around the current level implies a real interest rate close to current market rates, features a Pareto frontier that is always decreasing in worker welfare: issuing government debt is not Pareto-improving.
One could alternatively consider another use of government deficits: subsidies for capital accumulation. Provided $r$ is low enough and hence the deficit large enough, the subsidies for capital may be substantial enough to offset the disincentive to holding capital from a higher $r$ that greater debt issuance brings about, potentially raising wages for workers. This would however come at the expense of greater risk-taking by entrepreneurs, dampening the aggregate welfare gains.

Regardless, my model gives a very different flavour to the discussions of the welfare benefits of debt issuance. Much of the discussion so far has focused on the potential transfers to the old or the poor, or the utility gains from the insurance benefits from expanding public services. But in my model most of the welfare gains actually come from providing a liquid, safe asset that entrepreneurs can use to partially insure themselves against the risks they alone face.

**Strategic interactions.** The presence of two governments simultaneously trying to issue the optimal quantity of public debt gives rise to some limited strategic interaction. Again considering only the steady state, suppose the government of country $i$ maximises the welfare of country $i$, given above, subject to bond market equilibrium and conditional on the public...
debt of country $j$. The optimality condition gives a best-response function of steady-state debt as a function of the debt of the other country, $D_i^*(D_j)$, which as Figure 18 below suggests, appears to be linear, although I have as yet found no expression for it. The Nash equilibrium is $(D_H^*(D_F^*), D_F^*(D_H^*))$, and is shown below for $\sigma_H = 0.3$ and $\sigma_F = 0.5$, and features debt between 3-5x GDP, clearly far greater than levels observed today, but also far beyond the point where $r < g$.

Figure 18: Nash equilibrium debt when both countries maximise welfare is >3x GDP

6 Dynamics Following a Government Spending Shock

I now turn to analysing the dynamics of the model following an unexpected, one-time persistent but transitory debt-funded increase in government spending above the steady state level, $G$, such as that many governments have undertaken in response to the Covid-19 pandemic. Government spending follows: $G_t = G + e^{-\kappa t}(G_0 - G)$, where $G_0 > G$. Since the stock of public debt is the only government choice that affects the equilibrium, the government could neutralise the effects of government spending shocks by paying for them by raising taxes and balancing the budget. But then, in this model government spending has no value, so adhering
rigidly to the logic of the model then of course the government should not increase spending in the first place. But let us consider the effects of debt-funded spending shocks nonetheless, if only because we observe them in practice, regardless of whether they are optimal or desirable responses to business cycles or other shocks.

Thus far I have not made explicit the dynamic process governing taxation, as it is not relevant for the steady state. However, if government spending is simply mean-reverting and tax rates are constant, then government debt dynamics may be unstable. To stabilise the system, I add a component to tax revenues that responds to the level of debt in excess of the target level – but not directly to government spending – to ensure tax revenues rise in response to excessive debt: \( T_t = rD + G + \chi (D_t - D) \).

The equations governing the dynamics of the model are laid out in full in Proposition 1, but we can simplify them considerably. Assuming log utility simplifies matters by an order of magnitude since the MPC simplifies to \( \rho \) and is hence exogenous; aggregate (entrepreneurial) consumption therefore only depends on total wealth, \( A_t = K_t + B_t \). In Autarky, since bond holdings must equal government debt, \( B_t = D_t \), which is also a state variable, consumption is predetermined. Log utility also serves to guarantee existence of a steady state equilibrium, as discussed above.

The Autarky case under log utility is particularly illuminating, because the dynamics can be expressed by a system of just two ODEs, and hence can be visualised by a phase diagram, with two initial conditions. In Appendix 2, I explore the case with \( \gamma \neq 1 \), wherein the dynamics are much the same as the log utility case, but I also allow for tax revenues to be a more general function of the state of the domestic economy, which can give rise to more exotic dynamics and potentially even real indeterminacy of equilibrium.

**Proposition 4.** *The aggregate dynamics of the model with log utility under Autarky are given*
by the following system of ordinary differential equations in \((K_t, D_t)\):

\[
\begin{align*}
\dot{K}_t &= \alpha K_t^{\alpha} - \delta K_t - \rho(K_t + D_t) + T(D_t) - G_t \\
\dot{D}_t &= r^A(K_t, D_t) D_t - (T(D_t) - G_t)
\end{align*}
\]

where \(r^A(K_t, D_t) = \alpha K_t^{\alpha - 1} - \delta - \frac{K_t}{K_t + D_t} \gamma \sigma^2 \) and \(T(D_t)\) is given above.

Proof. See Appendix.

\(\chi > 0\) is generally needed to ensure local stability of the steady state; turning it from an (un)stable saddle-path to a sink. Note however that \(\chi > 0\) is not sufficient to ensure global stability of the economy. \(\chi\) must be sufficiently large, the shock to government spending, \(G_0 - G\), sufficiently small and transitory, and the initial condition sufficiently close to the steady state. The role of \(\chi\) is reminiscent of the parameter governing the responsiveness of surpluses to excessive debt in Lorenzoni & Werning (2019), although there it pins down a unique equilibrium, whereas here it stabilises the dynamics around the steady state, and perhaps rules out multiple steady states, while the equilibrium itself is unique (i.e. there is a unique mapping from states to prices \(r(K_H, K_F, D)\)). In practice, \(\chi = 0.1\) seems sufficient to ensure stability for a government spending shock even on the order of magnitude of 25\% of GDP. Figure 16 shows the phase diagram for this economy; if taxes are sufficiently reactive to excessive debt, i.e. \(\chi\) is large enough, then the vector field is locally stable in a large enough neighbourhood around the steady state that the economy returns to the steady state as \(t \to \infty\). If not, the equilibrium is unstable and the economy diverges, which in a monetary model would entail hyperinflation and the real value of the nominal debt claim going to zero.
**Proposition 5.** The aggregate dynamics of the model with log utility under financial Integration are determined by the following system of ODEs in \((K_{Ht}, K_{Ft}, D_t)\):

\[
\begin{align*}
\dot{A}_{it} &= (\alpha K_{it}^{\alpha - 1} - \delta - r(K_{Ht}, K_{Ft}, D_t)) K_{it} + (r(K_{Ht}, K_{Ft}, D_t) - \rho) A_{it} \quad i = \{H, F\} \\
\dot{D}_t &= (r(K_{Ht}, K_{Ft}, D_t) - r) D_t - \chi(D_t - D) + G_{Ht} - G_H + G_{Ft} - G_F
\end{align*}
\]

where \(A_{it} = \min \left\{ \frac{\alpha K_{it}^{\alpha - 1} - \delta - r(K_{Ht}, K_{Ft}, D_t)}{\sigma_i^2}, \lambda_i \right\}^{-1} K_{it} \) for each \(i = \{H, F\}\) and where \(r(K_{Ht}, K_{Ft}, D_t)\) is determined implicitly by the bond market equilibrium equation in Theorem 1:

\[
D_t = \max \left\{ \frac{\gamma \sigma_H^2}{\alpha K_{Ht}^{\alpha - 1} - \delta - r(K_{Ht}, K_{Ft}, D_t)} - 1, -\theta_H \right\} K_{Ht}
\]

\[
+ \max \left\{ \frac{\gamma \sigma_F^2}{\alpha K_{Ft}^{\alpha - 1} - \delta - r(K_{Ht}, K_{Ft}, D_t)} - 1, -\theta_F \right\} K_{Ft}
\]

with \(r\) simply being the interest rate in steady state.

*Proof.* See Appendix. \qed

I omit the phase diagram here, since it is easier to simply imagine three planes intersecting
one another, with the vector fields all pointing towards the intersection (the steady state) when $\chi_H = \chi_F = \chi$ is sufficiently large. Under Integration, the response of output to the government spending shock is dampened when the spender is financially integrated with a riskier country. The response is further dampened when the private supply of safe assets is materially constrained, i.e. when the credit constraints bind.

Figure 20: Integration and credit constraints dampen the negative effects of spending shocks

Stability. It is fruitful to examine formally the stability of the steady state in response to shocks. As alluded to above, the responsiveness of taxes to excessive debt is critical to ensuring locally stability of the steady state. We can examine locally stability through the Jacobian of the dynamic system(s) above. Since everything in this section and the next relates to both the Autarkic and Integrated economies, I simplify the discussion by focusing on the case of Autarky. Local dynamics around the steady state are given by the following system of linear differential equations:

**Proposition 6.** The dynamics of the economy around the steady state under Autarky are
given by the following system of linear differential equations in \((K_t, D_t)\):

\[
\begin{pmatrix}
\dot{K}_t \\
\dot{D}_t
\end{pmatrix} =
\begin{bmatrix}
\alpha^2 K^\alpha - 1 - \delta - \rho & -\rho + \chi \\
r_K D & r + \frac{D K \sigma^2}{(K+D)^2} - \chi
\end{bmatrix}
\begin{pmatrix}
K_t - K \\
D_t - D
\end{pmatrix}
\]

Proof. Calculate the Jacobian matrix of the system of ODEs in Proposition 4. Note \(r = \alpha K^\alpha - 1 - \delta - \frac{K \sigma^2}{K+D}\) and \(r_K = \alpha (\alpha - 1) K^{\alpha - 2} - \frac{\sigma^2}{K+D} + \frac{K \sigma^2}{(K+D)^2}\).

The eigenvalues of the Jacobian matrix indicate whether or not the steady state is locally stable. For \(\chi\) sufficiently large the dynamics are stable (both eigenvalues negative); if \(r \ll 0\), \(\chi = 0\) may be sufficient for local stability. If \(\chi\) is too low, one eigenvalue may be positive and the system becomes a saddle path, which is unstable here since both \(K_t\) and \(D_t\) are state variables. Note that the trace of the matrix is very likely to be negative if \(r < 0\), and the determinant is also likely to be negative, unless \(\chi\) is sufficiently large. Negative trace and positive determinant entails that both eigenvalues are negative, and hence stability, but a negative trace and negative determinant entails one positive and one negative eigenvalue, and hence instability.

**Multiple steady states and debt traps.** Note that even when the steady state is unique when the tax and spending functions are undefined, when these functions are defined in terms of endogenous variables like \(D_t\), then these additional dynamics can introduce a second steady state, usually featuring lower capital and higher debt (a “debt trap”) and which is typically unstable. However, under certain conditions the existence of multiple steady states can give rise to interesting dynamics in the face of large shocks. With local stability, small shocks will result in the economy tending back to the main steady state as \(t \to \infty\); however, large shocks can push the economy out of the stable zone (in green below) and into the unstable zone (in red). This may result in the economy rapidly diverging away from steady state (the hyperinflation scenario); but if \(r \ll g\) and taxes do not respond, or respond only modestly, to excess debt, the dynamics around the second steady state may be sufficiently slow that
the economy may get “stuck” there for an extremely long time. In this scenario the economy faces a debt trap, which features higher debt and lower output than the standard steady state. It is hard to get out of the debt trap naturally; it is a knife-edge case as to whether the economy returns to the normal steady state or explodes, and in any case this may take decades or even centuries. It may be necessary to engage in austerity in order to return to the normal steady state within a reasonable timeframe.

Figure 21: Large shocks may push the economy into a Debt Trap

7 Conclusion

In evaluating the fiscal space of governments today, we must confront two questions: why has the interest rate on public debt fallen so much, in spite of rising debt and high deficits; and why has this fall occurred primarily in the last two decades? The answer I believe lies in nascent global imbalances: greater demand for safe assets, precipitated by the rise of China and other emerging market economies, and the limited if not declining private supply of safe assets, as exemplified by the collapse of the asset-backed securities market during the financial crisis. In this paper I have provided a highly tractable environment in which to model these
two features and drawn simple but policy-relevant analytical and quantitative conclusions about the limits to the fiscal space available to governments.

The key take-away for policymakers is that precautionary saving by EM investors and the limited private supply of safe assets can be powerful forces generating significant fiscal space, but this fiscal space is limited, and great care is needed to design fiscal policies that are both sustainable in the long-run and stabilising in the short-run. Governments should target the long-run level of debt, not the deficit. This debt level can be large and still sustainable and consistent with deficits, provided \( r < g \), but governments must be willing to adapt their long-run spending and taxation plans to accommodate changing conditions; in other words, adjust deficits to hit the debt target. Governments can borrow substantially in response to economic crises, but must be sufficiently aggressive in reducing their borrowing subsequently. They can return to running (primary) deficits in the long run, but in the aftermath of a spending shock they must run smaller deficits, perhaps even surpluses, to ensure stability and avoid inflationary spirals or debt traps, even if the prevailing interest rate is below the growth rate.

A complementary if not competing hypothesis to the global imbalances thesis for why \( r < g \) today is the demographic thesis, which states that compared to 50 years ago, the world population today is on average older and thus saves more. What then do these hypotheses imply for the fiscal space of governments in the future? Goodhart & Pradhan (2020) assert that continued ageing of the population implies a decline in global demand for safe assets as older workers start to run down their savings in retirement, while Auclert et al. (2021) take the opposing view. Whatever the case, in my view the global imbalances hypothesis implies if anything more fiscal space in future. Emerging market economies are likely to continue to grow in importance, becoming a larger share of world output and holding a larger share of world wealth. However, despite this the riskiness of investment in these economies may not diminish greatly in the near future, in part because so much of the risk involved stems from underlying structural and political drivers, like risks of political upheaval or
politically-motivated confiscation or redistribution of assets that don’t necessarily diminish with economic growth.

Nevertheless, the link between economic development and the riskiness of investment is not clear, and it may be the case that financial development occurring alongside economic development allows for greater risk-sharing among investors, thus reducing the extent or bite of idiosyncratic risk. On the other hand, the safe-asset status of US or other developed-country government bonds is not guaranteed, and it is notable that capital flows from emerging markets to the developed world have to an extent receded since the financial crisis, and in the last year or so it appears EM holdings of DM government bonds have also declined somewhat. Either way, idiosyncratic risks to investors and the resulting global imbalances, as well as limits on the ability of the private sector to create substitute safe assets, are likely to remain important determinants of the fiscal space available to governments, and thus warrant further research.
References


Ball, L. & Mankiw, N. G. (2021), Market power in neoclassical growth models.


Global Imbalances & The Limits of Fiscal Space

Lee Tyrrell-Hendry


Corneli, F. (2017), Medium and long term implications of financial integration without financial development, Bank of Italy working papers.


Drechsel, T. (2021), Earnings-based borrowing constraints and macroeconomic fluctuations.


Holston, K., Laubach, T. & Williams, J. C. (2020), Adapting the laubach and williams and holston, laubach, and williams models to the covid-19 pandemic.


Kocherlakota, N. R. (2021), Public debt bubbles in heterogeneous agent models with tail risk.


Le Grand, F. & Ragot, X. (2022), Should we increase or decrease public debt? optimal fiscal policy with heterogeneous agents.


Reis, R. (2021), The constraint on public debt when $r < g$ but $g < m$.


Appendix 1: Extended proofs

Proof of Lemma 1:

Proof. The entrepreneur’s value function can be expressed recursively by taking the derivative of the sequence form with respect to time, or similar variational argument:

\[ V(t, a) = \max_{\{c_s, k_s\}_{s \in [t, \infty)}} \mathbb{E}_t \left[ \int_t^\infty e^{-\rho(s-t)} u(c_s) ds \mid a_t = a \right] \]

\[ \rho V(t, a) = \max_{c, k \leq \lambda a} u(c) + \lim_{dt \to 0} \frac{1}{dt} \mathbb{E}_t [dV(t, a)] \]

This yields Hamilton-Jacobi-Bellman equation, wherein the conditional expectation can be expanded using Ito’s Lemma:

\[ \rho V(t, a) = \max_{c, k \leq \lambda a} u(c) + \partial_a V(t, a) (R_t - r_t) k + r_t a - c + \frac{1}{2} \partial_{aa} V(t, a) \sigma^2 k^2 + \dot{V}(t, a) \]

For the remainder of the proof it is convenient to separate the two maximisation problems:

\[ \rho V(t, a) = \max_c \{u(c) - \partial_a V(t, a)c\} + \max_{k \leq \lambda a} \left\{ \partial_a V(t, a) (R_t - r_t) k + \frac{1}{2} \partial_{aa} V(t, a) \sigma^2 k^2 \right\} \]

\[ + \partial_a V(t, a) r_t a + \dot{V}(t, a) \]

With CRRA utility, \( u(c) = \frac{c^{1-\gamma}}{1-\gamma} \), the first-order conditions are given by the following. In particular, the optimal investment decision is myopic; the constraint only plays a role when it binds, as shown by Vila & Zariphopoulou (1997) in their Proposition 4.2:

\[ c = \partial_a V(t, a)^{-\frac{1}{\gamma}} \quad k = \begin{cases} -\frac{\partial_a V(t, a)(R_t - r_t)}{\partial_{aa} V(t, a) \sigma^2} & \text{if unconstrained} \\ \lambda a & \text{if constrained} \end{cases} \]
When $\gamma \neq 1$, I guess that the entrepreneur’s value function has the form: $V(t, a) = \frac{m_t^\gamma a^{1-\gamma}}{1-\gamma}$, which implies linear policy functions:

$$c(t, a) = m_t a \quad k(t, a) = \omega_t a = \begin{cases} R_t - r_t a & \text{if } \frac{R_t - r_t}{\gamma \sigma^2} < \lambda \\ \lambda a & \text{if } \frac{R_t - r_t}{\gamma \sigma^2} \geq \lambda \end{cases}$$

$$b(t, a) = (1 - \omega_t) a$$

One can verify this guess by plugging it into the HJB equation and verifying that the equation still holds with equality, which it does when the marginal propensity to consume, $m_t$, is determined by the ODE given in the lemma; this ODE can be derived by dividing both sides of the HJB equation by $\frac{\gamma}{1-\gamma} m_t^\gamma a^{1-\gamma}$ and rearranging:

$$\frac{\dot{m}_t}{m_t} = m_t - r_t + \frac{r_t - \rho}{\gamma} + \frac{1 - \gamma}{\gamma} (R_t - r_t) \omega_t + \frac{\sigma^2 (\gamma - 1)}{2} \omega_t^2$$

When $\gamma = 1$, I guess the value function has the form $V(t, a) = E_t + \frac{1}{\rho} \ln(a)$, which by the same process yields the ODE:

$$\dot{E}_t = \rho E_t + 1 - \ln(\rho) - \frac{1}{\rho} \left( (R_t - r_t) \omega_t + r_t - \frac{1}{2} \sigma^2 \omega_t^2 \right)$$

where $\omega_t$ is the same as above. The entrepreneur’s Euler equation can be found by differentiating their policy function with respect to time, $\frac{dc_t}{c_t} = \frac{\dot{m}_t}{m_t} + \frac{d\omega_t}{\omega_t}$, and plugging in the above ODE for $m_t$ and the budget constraint:

$$\frac{dc_t}{c_t} = \left( \frac{r_t - \rho}{\gamma} + \frac{R_t - r_t}{\gamma \omega_t} + \frac{\sigma^2 (\gamma - 1)}{2 \omega_t^2} \right) dt + \min \left\{ \frac{R_t - r_t}{\gamma \sigma^2}, \lambda \right\} \sigma dW$$

With logarithmic utility, the entrepreneurs’ choices and paths for wealth trivially satisfy the transversality condition: $\lim_{t \to \infty} E_{t_0} \left[ e^{-\rho(t-t_0)} \frac{1}{\rho a_t} a_t \right] = 0$. For CRRA utility more generally, the condition still holds provided the economy is at or tends to the steady state in the long-run, and provided the marginal propensity to consume is positive, $m > 0$ (for more discussion of which see below), and provided $\rho \geq -\frac{1}{2} \gamma (1 - \gamma) \sigma^2 \omega^2$, for when these hold:
\[
\lim_{t \to \infty} E_{t_0} \left[ e^{-\rho(t-t_0)} m^{-\gamma} a_t^{1-\gamma} \right] = \lim_{t \to \infty} \exp \left\{ (1 - \gamma)s - \frac{1}{2} \gamma(1 - \gamma)\sigma^2 \omega^2 - \rho \right\} (t - t_0) = 0, \tag*{where \( s \) is the drift of savings, and which equals zero in the limit by virtue of the economy tending towards steady state.}
\]

\textbf{Proof of Theorem 2:}

\textit{Proof.} To guarantee a steady state equilibrium exists, we need that total bondholdings are below the total stock of public debt for some sufficiently low interest rate, and above for some sufficiently high interest rate, i.e. \( B_H(r^-) + B_F(r^-) < D_H + D_F \) and \( B_H(r^+) + B_F(r^+) > D_H + D_F \). Recall that bond holdings in a given country comprise a product of a weight on bondholdings (relative to capital) and the level of the capital stock: \( B_i(r) = \frac{1-\omega_i(r)}{\omega_i(r)} \cdot K_i(r) \). It is easy to see that as \( r \to \rho \) from below, bondholdings for either country become arbitrarily large. The bond weight, \( \frac{1-\omega_i(r)}{\omega_i(r)} = \frac{\gamma \sigma_i}{\mu_i(r)} - 1 \), is clearly positive and in fact tends to positive infinity since \( \lim_{r \uparrow \rho} \mu_i(r) \to 0 \) (credit constraints will not bind here so we can ignore the max operator), and the capital stock will approach its complete markets level in this limit: \( \lim_{r \uparrow \rho} K(r) = \left( \frac{\alpha}{\rho+\delta} \right)^{\frac{1}{1+\gamma}} \). Thus bondholdings \( \forall i \lim_{r \uparrow \rho} B_i(r) \to \infty \).

It remains to show that for interest rates sufficiently low that bond holdings for at least one country tend to negative infinity; if this holds in one country, total bond holdings will also tend to negative infinity, since bondholdings in the other country will be finite. Consequently, we will have shown (since all these functions are continuous), that bond market equilibrium will hold for some \( r^* \) for any finite amount of total public debt.

In the absence of credit constraints, there exists an \( \bar{r} < 0 < \rho \) such that when approached from above the capital stock tends to infinity; this \( \bar{r} \) is the interest rate such that the expression in the denominator of the capital stock equals zero, i.e. \( \bar{r} + \delta + \sigma \sqrt{\frac{2 \gamma (\rho - \bar{r})}{1+\gamma}} = 0 \), or \( \bar{r} = -\delta - \frac{\gamma \sigma^2}{1+\gamma} - \sqrt{\left( \delta + \frac{\gamma \sigma^2}{1+\gamma} \right)^2 + \frac{2 \gamma (\rho - \bar{r})}{1+\gamma} \sigma^2 - \delta^2} \). Thus \( \forall i \lim_{r \downarrow \bar{r}} K_i(r) \to \infty \). For interest rates below this level, the capital stock is undefined. Moreover, since the weight on bonds is monotonically increasing in \( r \) (\( \frac{\partial}{\partial r} \left[ \frac{1-\omega_i(r)}{\omega_i(r)} \right] = -\gamma \sigma_i \mu_i(r)^{-2} \mu'(r) > 0 \)), there is always an \( \hat{r} < \rho \) sufficiently low such that the weight on bonds is negative. This occurs when \( \frac{\gamma \sigma_i}{\mu_i(r)} - 1 = 0 \), i.e.
\[ \hat{r} \equiv \rho - \frac{(1+\gamma)}{2}\sigma^2_i. \]

Hence, in order to guarantee a steady state equilibrium exists, all we need is that \( \hat{r} > r \). If so, there must exist an \( r^* \) such that total bond holdings equal any given finite level of total public debt. If not, the weight on bonds will become negative at an interest rate where the capital stock is undefined, or rather bond holdings will always remain positive and will tend to positive infinity as \( r \downarrow r \), so a steady state equilibrium may not exist for arbitrary levels of public debt, though it will exist for sufficiently high levels of public debt.

The condition \( \hat{r} > r \) implies

\[
\rho - \frac{(1+\gamma)}{2}\sigma^2_i > -\delta - \frac{\gamma\sigma^2_i}{1+\gamma} - \sqrt{\left(\delta + \frac{\gamma\sigma^2_i}{1+\gamma}\right)^2 + \frac{2\gamma\rho}{1+\gamma}\sigma^2_i - \delta^2}. 
\]

This simplifies through some tedious algebra to \( \rho - \delta > \frac{\gamma\sigma^2_i}{2}\gamma(\gamma - 1) \), which is the condition stated. Note that if \( \gamma \leq 1 \), the condition always holds, so a steady state equilibrium always exists. The right hand side is also clearly increasing in \( \gamma \) and \( \sigma \), so higher levels of risk aversion or risk make this condition harder to satisfy.

Note also that entrepreneurial consumption must also obviously be positive in order for the equilibrium to exist. When \( \gamma \leq 1 \) this is always satisfied, and when \( \gamma > 1 \) it is satisfied when Foreign risk is not too great, since (ignoring credit constraints for simplicity) the marginal propensity to consume in the steady state is given by:

\[
m = r + \frac{\rho - r}{\gamma} + \frac{1}{2} \frac{1}{\gamma} \frac{(R(r) - r)^2}{\gamma\sigma^2} > 0
\]

\[
m = r + \frac{2}{1+\gamma}(\rho - r) > 0
\]

\[
\Rightarrow r^* \begin{cases} < \frac{2\rho}{1-\gamma} & \text{if } \gamma < 1 \\
< \frac{2\rho}{\gamma-1} & \text{if } \gamma > 1
\end{cases} \iff r^* < \rho < \frac{2\rho}{1-\gamma}
\]

i.e. the condition \( m > 0 \) requires the equilibrium interest rate, \( r^* \), to be below some value that is always greater than \( \rho \), but in equilibrium the \( r^* < \rho \), as is typical for models with incomplete markets. When \( \gamma = 1 \), the MPC is equal to \( \rho \), so the condition is trivially satisfied. In order for the condition to hold when \( \gamma > 1 \), the equilibrium interest rate must be above some negative value, \( r^* > -\frac{2\rho}{\gamma-1} \). Using from above that \( r^* > \hat{r} \), a sufficient condition to ensure \( m > 0 \) when \( \gamma > 1 \) is thus \( \rho > \frac{\sigma^2_i}{2}\gamma(\gamma - 1) \).
The potential for a steady state to not exist contrasts with Angeletos & Panousi (2011), where the presence of labour income for entrepreneurs ensures that bondholdings tend to negative infinity as interest rates approach zero from above, ensuring an equilibrium interest rate always lies between 0 and $\rho$. This is because in their set-up bondholdings depend negatively on human capital, which since entrepreneurial labour income is riskless is discounted at the risk-free rate and hence is proportional to the inverse of the interest rate in the steady state. Obviously ensuring the interest rate is always positive defeats the purpose of my analysis here, so this is one reason I dispense with entrepreneurial labour income, as outlined above.

**Proof of Corollary 1:**

*Proof.* Let us consider the case where Home entrepreneurs are net borrowers and Foreign entrepreneurs are net lenders, i.e. $\frac{\gamma \sigma_H}{\mu(r)} - 1 < 0 < \frac{\gamma \sigma_F}{\mu(r)} - 1$, which requires that $\sigma_H < \sigma_F$.

Global bond holdings are given by:

$$B_H(r) + B_F(r) = \left(\frac{\gamma \sigma_H}{\mu(r)} - 1\right) \left(\frac{\alpha}{r + \delta + \sigma_H \mu(r)}\right)^\frac{1}{\alpha - 1} + \left(\frac{\gamma \sigma_F}{\mu(r)} - 1\right) \left(\frac{\alpha}{r + \delta + \sigma_F \mu(r)}\right)^\frac{1}{\alpha - 1} = D_H + D_F$$

We will proceed by proving that in this scenario global bond holdings are strictly increasing in the interest rate, $r$. Differentiating with respect to $r$:

$$\frac{\partial}{\partial r} [B_H(r) + B_F(r)] = -\gamma \sigma_H \mu(r)^{-2} \mu'(r) \left(\frac{\alpha}{r + \delta + \sigma_H \mu(r)}\right)^\frac{1}{\alpha - 1} \frac{1 + \sigma_H \mu'(r)}{\alpha}$$

$$+ \frac{1}{\alpha - 1} \left(\frac{\gamma \sigma_H}{\mu(r)} - 1\right) \left(\frac{r + \delta + \sigma_H \mu(r)}{\alpha}\right)^\frac{1}{\alpha - 1} - \gamma \sigma_F \mu(r)^{-2} \mu'(r) \left(\frac{\alpha}{r + \delta + \sigma_F \mu(r)}\right)^\frac{1}{\alpha - 1} \frac{1 + \sigma_F \mu'(r)}{\alpha}$$

$$+ \frac{1}{\alpha - 1} \left(\frac{\gamma \sigma_F}{\mu(r)} - 1\right) \left(\frac{r + \delta + \sigma_F \mu(r)}{\alpha}\right)^\frac{1}{\alpha - 1}$$

One can easily note that the first and third terms are positive, since $\mu'(r) = -\frac{1}{2} \sqrt{\frac{2 \gamma}{1+\gamma}}(\rho - \rho)^{-\frac{1}{2}}$.
Moreover, the second term is also positive, since by assumption \( \frac{\gamma \sigma}{\mu(r)} - 1 < 0 \), which implies \( r < \rho - \frac{\gamma^2}{2} \gamma (1 + \gamma) \). This in turn implies that \( r < \rho - \frac{\gamma^2}{2} \frac{\gamma}{1+\gamma} \), or \( 1 + \sigma_H \mu'(r) > 0 \) (and of course \( \alpha < 1 \)). Finally, the forth term is also positive, since by assumption \( \frac{\gamma \sigma}{\mu(r)} - 1 > 0 \), and \( r < \rho - \frac{\gamma^2}{2} \frac{\gamma}{1+\gamma} \) implies that \( r > \rho - \frac{\gamma^2}{2} \frac{\gamma}{1+\gamma} \) since \( \sigma_F > \sigma_H \), which in turn implies \( 1 + \sigma_F \mu'(r) < 0 \).

A similar argument holds for the case where Foreign entrepreneurs are net borrowers and Home net lenders. Clearly however it is not necessary for this condition to hold for bondholdings to be monotonic, nor is it necessary for bondholdings to be monotonic for equilibrium to be unique, but the more restrictive conditions under which the equilibrium is unique are not particularly revealing.

In the Autarky case, there are to my knowledge no simple and revealing conditions that suffice to guarantee uniqueness. In this case, the equilibrium condition simplifies to:

\[
B(r) = \left( \frac{\gamma \sigma}{\mu(r)} - 1 \right) \left( \frac{\alpha}{r + \delta + \sigma \mu(r)} \right)^{\frac{1}{1-\alpha}} = D
\]

which yields a unique \( r^* \) when its derivative is positive:

\[
\frac{\partial}{\partial r} [B(r)] = - \frac{\gamma \sigma}{\mu(r)^2} \mu'(r) \left( \frac{\alpha}{r + \delta + \sigma \mu(r)} \right)^{\frac{1}{1-\alpha}} + \frac{1}{\alpha - 1} \left( \frac{\gamma \sigma}{\mu(r)} - 1 \right) \left( \frac{r + \delta + \sigma \mu(r)}{\alpha} \right)^{\frac{1-\alpha}{\alpha}} \frac{1 + \sigma \mu'(r)}{\alpha} > 0
\]

Assuming public debt is positive, it must be the case that \( \frac{\gamma \sigma}{\mu(r)} - 1 > 0 \), so to guarantee monotonicity without dissecting each term more thoroughly requires \( 1 + \sigma \mu'(r) < 0 \), or \( r^* > \rho - \frac{\gamma^2}{2} \frac{\gamma}{1+\gamma} \), or that the equilibrium lies in the upward-sloping part of the capital locus, which in general I have shown above does not hold, except for very high levels of public debt. Nevertheless, bondholdings are very likely to be monotonically increasing in the interest rate if \( \gamma \leq 1 \), and when risk is not too high if \( \gamma > 1 \).
Proof of Proposition 3:

Proof. First, we can find an expression for the stationary distribution of wealth over entrepreneurs from the Kolmogorov Forward equation (KFE), using the assumption discussed in the text that entrepreneurs die at rate $p$ and are reborn with the average level of wealth, $A$:

$$\dot{g}(t,a) = -\partial_a [s(a)g(a)] + \frac{1}{2}\sigma^2\partial_{aa} [\omega(a)^2g(a)] - pg(a) = 0$$

where, $s(a) = (R-r)\omega a + ra - c$ denotes average saving (i.e. ignoring shocks to wealth, which is captured by the second term). Since both consumption, $c = ma$, and capital holdings, $k = \min \left\{ \frac{R-r}{\gamma\sigma^2}, \lambda \right\} a$ (denoting $\omega \equiv \min \left\{ \frac{R-r}{\gamma\sigma^2}, \lambda \right\}$), are proportional to wealth, we have that $s(a) = sa$.

The first thing to note is that the first term in the KFE drops out, because in the steady state $s$ (like aggregate savings) equals zero, $s = (R-r)\omega + r - m = 0$, to see this note that in steady state, the ODE determining the MPC simplifies to:

$$m = r - \frac{r - \rho}{\gamma} + \frac{\gamma - 1}{\gamma}(R-r)\omega - \frac{\sigma^2(\gamma - 1)}{2}\omega^2$$

Thus:

$$s = \frac{r - \rho}{\gamma} + \frac{1}{\gamma}(R-r)\omega + \frac{\sigma^2(\gamma - 1)}{2}\omega^2 = 0$$

which is exactly the condition satisfied by the steady state Euler equation, $\dot{C} = 0$.

It is well-established (e.g. see Gabaix (2009)) that a double-Pareto distribution satisfies the KFE equation. To see this, plug in the guess $g(a) = \kappa a^{-\zeta - 1}$:

$$0 = \frac{1}{2}\sigma^2\partial_{aa} [\omega^2\kappa a^{1-\zeta}] - p\kappa a^{-\zeta - 1}$$

$$0 = \zeta^2 - \zeta - \frac{2p}{(\sigma\omega)^2}$$
This polynomial has two solutions, \( \zeta_\pm = \frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{2p}{(c_0)^2}} \), which represent the two tail parameters for each branch of a double-Pareto distribution around the re-entry point, \( A \).

We can now find the constant \( \kappa \) by using the condition that the distribution must integrate to 1, the unit mass of entrepreneurs:

\[
1 = \int_0^A \kappa \left( \frac{a}{A} \right)^{-\zeta_- - 1} da + \int_A^\infty \kappa \left( \frac{a}{A} \right)^{-\zeta_+ - 1} da
\]

so that, \( \kappa = \frac{\zeta_+ - \zeta_-}{(\zeta_- - \zeta_+)A} \)

The remainder of the proof is as follows:

\[
W_i = \varpi V_{it}^w + (1 - \chi) \int_0^\infty V_i^e(a) g_i(a) da
\]

\[
= \varpi \frac{1}{\rho} \left[ \ln(C^w) - \frac{L^{1+\varphi}}{1+\varphi} \right] + (1 - \varpi) \int_0^\infty \frac{m_i^{-\gamma}}{1-\gamma} a^{1-\gamma} g_i(a) da
\]

\[
= \varpi \frac{1}{\rho} \left[ \ln \left( (1 - \tau_i)(1 - \alpha)K_i^{\alpha} \right) - \frac{1}{1+\varphi} \right] + (1 - \varpi) \frac{m_i^{-\gamma}}{1-\gamma} \int_0^\infty a^{1-\gamma} \kappa \left( \frac{a}{A_i} \right)^{-\zeta_+ - 1} da
\]

\[
= \varpi \frac{1}{\rho} \left[ \ln \left( \left(1 - \alpha - r D_i Y_i \right)K_i^{\alpha} \right) - \frac{1}{1+\varphi} \right]
\]

\[
+ (1 - \varpi) \frac{m_i^{-\gamma}}{1-\gamma} (K_i + B_i)^{1-\gamma} \frac{\zeta_i + \zeta_i - (1-\gamma - \zeta_i-)(1-\gamma - \zeta_i+)}{(1-\gamma - \zeta_i-)(1-\gamma - \zeta_i+)}
\]

which is the equation given in the text, where I have used in evaluating the integral that new entrepreneurs are born with wealth equal to the average among all entrepreneurs, i.e. \( A_i \). \( \square \)

**Proof of Proposition 4:**

**Proof.** Take equation 2 and substitute in equations 4 and entrepreneurs’ aggregated policy function for consumption, \( C_{it} = \rho (K_{it} + B_{it}) \), while also using bond market clearing, 7:

\[
\dot{K}_{it} + \dot{D}_{it} = \alpha K_{it}^\alpha - \delta K_{it} + r_{it} D_{it} - \rho (K_{it} + D_{it})
\]
Furthermore, subtract equation 6 from both sides, yielding:

\[
\dot{K}_{it} = \alpha K_{it}^\alpha - \delta K_{it} - \rho (K_{it} + D_{it}) + T(D_{it}) - G_{it}
\]

which is the expression given. One can arrive at the same equation by manipulating the Euler equation (equation 3) in a similar fashion. The second differential equation is obviously immediate from equation 6, substituting in the equilibrium interest rate \( r_{i}^A(K_{i}, D_{i}) = \alpha K_{i}^{\alpha - 1} - \delta - \frac{K_{i}}{K_{i} + D_{i}} \gamma \sigma_{i}^2 \).

**Proof of Proposition 5:**

**Proof.** Again starting with 2, this time defining \( A_{it} \equiv K_{it} + B_{it} \), and again substituting in equations 4 and \( C_{it} = \rho (K_{it} + B_{it}) \) and this time also equation 5, one arrives at the first ODE of Proposition 5:

\[
\dot{A}_{it} = \left( \alpha K_{it}^{\alpha - 1} - \delta - r_{t} \right) \min \left\{ \frac{\alpha K_{it}^{\alpha - 1} - \delta - r_{t}}{\sigma_{i}^2}, \lambda_{i} \right\} A_{it} + (r_{t} - \rho) A_{it} \quad i = \{H, F\}
\]

Adding together the government debt flow equations 6 for both countries, and using the definitions of \( T(D_{it}) \) – where for sake of simplicity I am assuming \( \chi_{H} = \chi_{F} = \chi \) – one arrives at the second ODE:

\[
\dot{D}_{t} = (r_{t} - r)D_{t} - \chi(D_{t} - D) + G_{Ht} - G_{H} + G_{Ft} - G_{F}
\]

The two static equations can be derived immediately from equations 5 and 8 by substituting in equation 4 and using that \( B_{it} = (1 - \omega_{it}) A_{it} \), yielding:

\[
K_{it} = \min \left\{ \frac{\alpha K_{it}^{\alpha - 1} - \delta - r_{t}}{\sigma_{i}^2}, \lambda_{i} \right\} A_{it} \quad i = \{H, F\}
\]

\[
D_{t} = \left( 1 - \min \left\{ \frac{\alpha K_{Ht}^{\alpha - 1} - \delta - r_{t}}{\sigma_{Ht}^2}, \lambda_{H} \right\} \right) A_{Ht} + \left( 1 - \min \left\{ \frac{\alpha K_{Ft}^{\alpha - 1} - \delta - r_{t}}{\sigma_{Ft}^2}, \lambda_{F} \right\} \right) A_{Ft}
\]
Appendix 2: Extensions

Dynamics when $\gamma \neq 1$ and $T_{it} = T(K_{it}, D_{it})$:

Now I explore the dynamics of the equilibrium, allowing for non-log CRRA utility and for tax revenues to be a more general function of the state of the domestic economy. As I will show, this can result in the steady state being unconditionally stable, with all negative eigenvalues, implying indeterminacy of equilibrium; many functions $C(t)$ converge to the steady state in the long run and hence are consistent with the transversality condition.

The dynamics of equilibrium under Autarky now become:

$$
\dot{K}_t = \alpha K_t^\alpha - \delta K_t - C_t + T(D_t, K_t) - G_t
$$

$$
\dot{D}_t = r^A(K_t, D_t) D_t - (T(D_t, K_t) - G_t)
$$

$$
\dot{C}_t/C_t = \frac{r^A(K_t, D_t) - \rho}{\gamma} + \frac{1 + \gamma}{2} \left( \frac{\alpha K_t^{\alpha - 1} - \delta - r^A(K_t, D_t)}{\gamma^2} \right)^2
$$

where again $r^A(K_t, D_t) = \alpha K_t^{\alpha - 1} - \delta - \frac{K_t}{K_t + D_t} \gamma \sigma^2$ and $T(D_t, K_t)$ is a continuously differentiable function with $T_D \geq 0$.

The Jacobian of this system evaluated at the steady state is as follows:

$$
\begin{pmatrix}
\dot{K}_t \\
\dot{D}_t \\
\dot{C}_t
\end{pmatrix} =
\begin{bmatrix}
\alpha^2 K_t^{\alpha - 1} - \delta + T_K & T_D & -1 \\
\gamma r_K D - T_K & r_D D + r - T_D & 0 \\
\frac{r_K}{\gamma} + \frac{1 + \gamma}{\gamma} \frac{(R_r - r)}{\gamma^2} (R_K - r_K) & C & \left[ \frac{r_D}{\gamma} - \frac{1 + \gamma}{\gamma} \frac{(R_r - r)}{\gamma^2} r_D \right] C & 0
\end{bmatrix}
\begin{pmatrix}
K_t - K \\
D_t - D \\
C_t - C
\end{pmatrix}
$$

Consider for simplicity calibrations where the steady state is unique. Since we have two state variables, $K$ and $D$, and one jump variable, $C$, the system is conditionally stable when two of its eigenvalues are negative and one is positive; the transversality condition then suffices to rule out explosive solutions and pin down a unique equilibrium path for consumption, capital and debt. If fewer than two eigenvalues are negative, the system is locally unstable and perturbations from the steady state – for example due to a government spending shock –
can put the economy on an unstable path, as discussed in the main part of the text. There is however a new possibility of indeterminacy of equilibrium, where consumption is now not pinned down uniquely as a function of capital and debt; if all eigenvalues are negative then the system is unconditionally stable, meaning within some range any choice of consumption is consistent with the transversality condition and will asymptote to the steady state.

It should be noted that this indeterminacy is distinct from the purely nominal indeterminacy encountered frequently in monetary economies with flexible prices operating under an interest rate rule (Sargent & Wallace 1975); here the indeterminacy is real, not nominal. But the cause of this real indeterminacy is also distinct from that which arises in New Keynesian models with nominal rigidities and passive interest rate policy (Clarida et al. 1999). Since indeterminacy here arises due to the specification of the government’s fiscal rule, one might term this fiscal indeterminacy.

There are to my knowledge no simple conditions under which indeterminacy or instability arises; it is more fruitful simply to calculate the eigenvalues numerically to ascertain their signs, especially in the equilibrium with integrated global capital markets. However, both the indeterminacy case and the instability case arise when the determinant of the Jacobian matrix above is negative, which implies that either all eigenvalues are negative or only one is, since the determinant of a matrix is the product of its eigenvalues $|J| = \lambda_1 \lambda_2 \lambda_3$. Loosely speaking, the determinant is negative when debt growth responds only weakly to higher debt but relatively strongly to higher capital, while at the same time consumption growth responds strongly to higher debt but weakly to higher capital:

$$|J| = - \left[ \partial_K (\dot{D}) \partial_D (\dot{C}) - \partial_D (\dot{D}) \partial_K (\dot{C}) \right]$$

Contrarily, when the determinant is positive, we are likely to encounter the case of conditional stability, with one positive eigenvalue and two negative. It is possible but unlikely that all eigenvalues are positive, since the trace of the Jacobian – which is the sum of the eigenvalues, $tr(J) = \lambda_1 + \lambda_2 + \lambda_3$ – is likely to be negative (implying at least one eigenvalue is negative),
and will be provided tax revenues respond sufficiently strongly to excessive debt and are not too-rapidly increasing in capital:

\[ tr(J) = \partial_K \left( \dot{K} \right) + \partial_D \left( \dot{D} \right) \]