Education Choice under Uncertainty and Public Policy*

Vincent Hogan†
Department of Economics
University College Dublin
and
University of Warwick

Ian Walker
Department of Economics
University of Warwick

This version: August 19, 2002

Abstract

We analyse how progressive taxation and education subsidies affects schooling decisions when the returns to education are stochastic. We use the theory of real options to solve the problem of education choice in a dynamic, life-cycle consistent, stochastic model. We show that education attainment will be an increasing function of the risk associated with education. Furthermore, this result holds whether or not agents can borrow in order to pay for education and regardless of the degree of risk aversion. We also examine the link between consumption over the life-cycle and education choice to show that higher initial wealth will usually – but not always – have a positive effect on eduction attainment. Finally we show that progressive taxes will tend to reduce education attainment for the poor but increase it for the rich.

JEL Classification: J24, C61, D81.

Keywords: Education Choice; Dynamic Optimization, Optimal Stopping, Uncertainty.

---

*We thank seminar participants at University College Dublin and the University of Warwick, Institute of Fiscal Studies, IZA-SOLE Annual Meetings 2002, North American Summer Meetings of the Econometric Society, 2002. Thanks are also due to Robert Marquez and Colm Harmon. This research has be supported by a Marie Curie Fellowship of the European Union programme "Improving Human Resource Potential and the Socio-Economic Knowledge Base" under contract HPMF-CT-2000-00879 and also by the Department of Education and Skills, through the Centre for the Economics of Education. Responsibility for the content of the paper is solely the authors.

†Corresponding author: vincent.hogan@ucd.ie; Tel.:+353-1-716-8300; Fax.: +353-1-283-0068; Mail: Dept. of Economics, UCD, Belfield, Dublin 4, Ireland.

1 Introduction

The study of how individuals choose their education levels and the measurement of the economic returns to education has been of great interest since Mincer (1974). The literature has typically viewed education choice as an investment in human capital; to be thought of in much the same way as we think of investment in physical or financial capital (see Card, 1999, for a comprehensive survey). However, the concept of risk – routinely included in theoretical and empirical discussions of other investment – is largely absent from discussions of individual schooling choice.[1]

This is a curious omission as the risk associated with education choices will surely be an important determinant of how individuals arrive at those choices. Dominitz and Manski (1996) show that individuals believe that education carries substantial risk and suggest that this influences their education decisons. We can imagine that a high level of risk (whether due to the prospect of failure to graduate, unemployment or the variability of wages conditional on graduation) might dissuade individuals from continuing with education. Alternatively, individuals might stay in education, as a form of insurance, if risk declines with higher education. In either case, analysing precisely how individuals react to risk will be of crucial interest to policy makers seeking to influence education attainment in general, and to discourage early school leaving in particular.

In this paper we analyse the effect of risk on schooling choice using techniques similar to those used in the description of financial and physical investment under uncertainty.[2] Our basic approach is to view education choice as an option problem. We think of an individual in school as possessing an option to leave at any time and take up work at a wage related (stochastically) to the time spent in school. Thus, the decision to leave school is a special case of the classic "tree cutting" or "optimal stopping" problem. Once that option is exercised, the individual cannot return to school.

The option approach is a close approximation to reality. Most individuals stay in education full time until they judge it optimal to leave, and after leaving, they do not return. Empirically, in the UK, over 90% of college students have come directly from school, while in the OECD as a whole, only 6.4% of those aged 25-29 years are still in education (full or part-time).[3]

[1]For examples from financial investment see Merton (1971) and Campbell et.al (1997). For physical capital see Caballero and Engle (1999) and Dixit and Pindyck (1994).

[2]The application of these methods to various economic problems have been analysed in Malliaris and Brock (1982), Kamien and Schwartz (1991) and Dixit (1993). Their application to irreversible physical investment (so called "real options") is surveyed in Dixit and Pindyck (1994).

[3]See www.hesa.ac.uk and and Table E3.1 of OECD (2001) a summary of which is available from www.oecd.org

We embed our model of education choice in a life-cycle model of consumption smoothing, so that we can allow individuals to subsidise education by borrowing or running down assets. This enables us to analyse the impact of wealth on education decisions. Furthermore we use our model to analyse the effects of tax and education policy controlling for the link between education choice, uncertainty and consumption. We show that risk interacts with the consumption and education decision in some counter intuitive ways. Firstly, higher risk encourages individuals to accumulate more human capital. Secondly, we can also show that consumption will jump upon graduation even if individuals are allowed borrow against future earnings while in education.

The previous literature on risk in education is not large. The most relevant theoretical work is Williams (1979) who adapted the optimal portfolio choice model of Merton (1971) to allow for investment in human capital. The model predicted that higher risk would induce individuals to accumulate less human capital. Groot and Oosterbeek (1992) used an unrestrictive (non-parametric) definition of risk, but at the cost of assuming risk neutrality. By using the techniques of real options, we can easily accommodate risk aversion but limit ourselves to certain stochastic processes for returns to education.

Keane and Wolpin (1997) implemented an empirical version of a dynamic model of education choice. Belzil and Hansen (2002) estimated a similar model in order to examine the correlation between the return to schooling and unobserved ability. However, in order for these models to be empirically tractable, both abstracted from the issue of consumption smoothing and focused on stochastic income rather than stochastic returns to education.

Chen (2001) estimated the parameters of a static model of college choice when the returns to a college education are uncertain. In order to be able to implement the model empirically, she also had to exclude the possibility of consumption smoothing. She found that the annualised return to a four year college education is 6.5% and that the associated risk (standard deviation of returns) is 7.5%.

We use our model to simulate the response of individuals to a variety of policy measures (fee reductions, tax increases, reduction in progressivity of the tax system etc.). In doing so, we build upon a large literature including Trostel (1993) and Heckman et. al. (1998) who examine the effects of tax policy in dynamic general equilibrium models under certainty and Eaton and Rosen (1980) and Altonji (1993) who examine policy effects in a stochastic two period models. As far as we know, we are the first to analyse policy questions using a model that allows for both uncertain returns to education and intertemporal optimisation. In contrast to many of the previous studies, we find that tax increases can actually increase education attainment for all but the poorest individuals faced with progressive taxation.

The paper proceeds a follows. Section 2 presents an overview of the problem and clarifies exactly how we model stochastic returns. We also solve a simple model of education choice with uncertain returns, where there is no borrowing or lending. Section 3 models the joint education and consumption decisions, enabling us to examine the effect of wealth on education choice. We also consider the policy implications of the model. Section 4 discusses some extensions and section 5 concludes.

2 Education Choice

We start with a model of education choice similar in structure to Card (2001). An individual chooses the number of years schooling (S) in order to maximise his or her expected discounted life time utility (1) subject to a budget constraint (2).

$$V = E\left\{\int_0^S e^{-\rho t}\left\{u(c_t) + \phi_t\right\} dt + \int_S^T e^{-\rho t} u(c_t) dt\right\} \tag{1}$$

$$\int_0^T e^{-rt} c_t dt = A_0 + \int_S^T e^{-rt} Y_t(s) dt - \int_0^S e^{-rt} F_t dt - e^{-rT} A_T \tag{2}$$

Assuming that the minimum school leaving age is normalised to $t = 0$, lifetime utility is provided by consumption (c) throughout life (i.e. both during and after school) via u, the instantaneous utility function and also by the direct (dis)utility of education, ϕ, where both u and ϕ are increasing concave functions, ρ is the constant rate of time preference, F are school fees, A is the stock of financial assets, and E is the expectations operator. Education choice is an optimal stopping problem, because the individual faces a once and for all decision to leave school (i.e. choose S) and he or she cannot return at a later date. The model differs from that of Card (2001) is so far as we allow for returns to be stochastic and maximise expected utility.

Note that as the problem is literally an option problem, it is best suited to analysing education choice after the end of compulsory education. We interpret S as being the time spent in *post-compulsory* education. We ignore the case of those who leave school early in violation of the law. Although, as we show in section 3.4, the model does offer some insight into the interaction of school leaving laws and the other parameters of the model, especially individual financial wealth.

We assume that the income process is time separable so that the return measured in terms of lifetime income is the same as the return measured in terms of income over any shorter interval. Formally we have

$$Y_t(s) = W(s)h(t-s) \tag{3}$$

We can think of W as the starting wage after leaving education with S years completed and h as being the factor by which the wage grows in each period as experience and seniority are accumulated ($h(0) = 1$). In essence (3) is a continuous time version of the standard Mincer (1974) equation and so is consistent both with the view that education represents the accumulation of human capital. Following our interpretation of S as post-compulsory schooling, we interpret $Y_t(0)$ as being the income profile of an individual with only the minimum education required by law and not as income of those with absolutely no education.

This specification of earnings (3) while reflecting standard empirical analysis, includes two probably unrealistic simplifications. Firstly, Heckman et. al. (2001) have cast doubt on the empirical relevance of the time separability assumption, providing evidence that in the US at least, earnings growth after leaving school is a function of the education level.

Secondly, by making S the choice variable we identify education with time spent in school and college and not necessarily with the accumulation of formal credentials. Of course the two are closely related, and there is empirical evidence of so-called "sheep-skin" effects i.e. non-linearities in education choice and earnings associated with school and college completion dates.[4] We ignore both issues here as their inclusion would complicate the analysis without shedding much light on the role of risk.

2.1 A Simple Example with Certainty

The solution to the dynamic programme depends crucially on the nature of the budget constraint (2) and subsequent sections of the paper we make it more realistic. For the moment, in order to provide a benchmark for comparison, we solve a simple example under certainty. To be specific, we assume that $u(c) = c$, and that there is no borrowing or lending, so that $c_t = 0 \ \forall \ t < S$ and $c_t = Y_t(s) \ \forall \ t \geq S$. We also assume that there are constant returns (g) to education and that $\rho > g$ (otherwise the agent would never leave school). In order to avoid the other corner solution (leave school immediately) we need to assume that ϕ is constant through time and positive, so that education is valued for its own sake. Finally, we assume that individuals are infinitely lived ($T = \infty$) and that $F_t = 0$, so that

[4]See Denny and Harmon (2001). Altonji (1993) presents a three period model of college attendence with stochastic returns (via uncertain graduation) and sheepskin effects but without consumption smoothing.

neither time nor F are state variables. Thus individuals choose S to maximise

$$V_0 = E\left\{\int_0^S e^{-\rho t}\phi dt + \int_S^\infty e^{-\rho t} Y_t(s) dt\right\} \tag{4}$$

In the absence of uncertainty we have $Y_t(s) = Y_0 \exp(gs)$, where the experience factor $h(t-s)$ has been set equal to one (so that earnings are constant after leaving school) and Y_0 represents earnings with only the minimum schooling.

In what follows it turns out to be more convenient to specify the school leaving decision in terms of $\bar{Y}$, the threshold level of income. The idea is intuitive. While in school the individual keeps an eye on the shadow wage i.e. the wage that he would get were he to leave school immediately. When it reaches a certain critical level, the individual will leave school. Consequently we can state the following well known result:

Proposition 1 *When returns to education are certain, the optimal level of education is given by*

$$S^* = \frac{1}{g}\ln\left(\frac{\rho}{\rho - g}\frac{\phi}{Y_0}\right)$$

and the associated threshold level of the shadow wage is given by

$$\bar{Y} = \frac{\rho\phi}{\rho - g}$$

where g is the return to education, ρ is the discount rate and ϕ is the intrinsic utility of education

Proof. By direct differentiation of (4) ■

2.2 Risky Education

Before proceeding to analyse the case of uncertain returns, we need to clarify what exactly we mean by risk. In this context we mean that two otherwise identical individuals may end up with different lifetime income profiles, just because of a different draw from the distribution of returns to education. Specifically we model the return to education as being drawn from a normal distribution. To keep things simple and to avoid time becoming a state variable, we continue to assume time separability and that $h(t-s) = 1$.

Consider staying on in school for κ more periods. The return to this extra schooling, $r(\kappa)$, will equal

$$r(\kappa) = \frac{Y(s+\kappa) - Y(s)}{Y(s)} \quad \backsim \quad N(g\kappa, \kappa\sigma^2)$$

which is distributed as a normal random variable with mean g and standard deviation σ when $\kappa = 1$. By taking limits, we can show that, in continuous time, the return to a infinitesimally small extra period in school ($r \equiv dY/Y$) will be distributed as $N(gds, \sigma^2 ds)$ implying that Y follows a geometric Brownian motion[5]

$$\frac{r-g}{\sigma} \backsim N(0, ds)$$

or in more usual notation

$$\frac{dY}{Y} = gds + \sigma dz \tag{5}$$

where dz represents the increments of a standard Weiner process i.e. where each increment is drawn from $N(0, ds)$. Note that in the absence of uncertainty ($\sigma^2 = 0$) the income process (5) reduces to (3).

Equation (5) states that for each instant that the individual remains in school her shadow wage trends up at rate g. In addition at each instant the shadow wage is subject to a (proportionate) shock that has zero mean and variance equal to σ^2. Therefore even if individuals start with the same (deterministic) Y_0 they will end up with different Y_s.[6]

It will sometimes be useful to work in terms of the distribution of Y. If returns (log wage) are normally distributed then the wage itself will have a log-normal distribution conditional on the initial value.[7]

$$\ln Y_s - \ln Y_0 \backsim N(\mu s, \sigma^2 s)$$
$$\mu = g - \frac{\sigma^2}{2}$$

Geometric Brownian motions are analytically convenient and have been used widely in economics and operations research to model asset prices.[8] A Brownian motion is a continuous time generalisation of a random walk (see Dixit, 1993) and is therefore a reasonable continuous time representation of many economic time series. However, it does have certain specific characteristics that we need to be aware of. Firstly, like a random walk, it is a Markov process implying that the probability distribution of Y at any time in the future is conditional on its current value and only on its current value. Knowing past values confers no extra information.

Secondly the increments of a Brownian motion process that occur over any two non-

[5]See Dixit (1993) for a derivation of a Brownian motion as the limit of a random walk.

[6]We treat Y_0 at deterministic as it will be known to the agent by the time she comes to make her education decision.

[7]Because of Ito's lemma $dx/x \neq d\ln x$. See Dixit (1993) for a discussion.

[8]See Malliaris and Brock (1982) and Dixit and Pindyck (1994) for surveys and references.

overlapping intervals are independent. Thirdly these increments are normally distributed with both the mean and variance of a growing linearly[9] with schooling i.e. $var(r) = \sigma^2 ds \neq (\sigma ds)^2$ – implying that the coefficient of variation of the shadow wage is constant. Fourthly, the sample path of a Brownian motion is jagged (so no time derivative exists) but continuous (so that no two paths will ever be exactly the same). In fact it can be shown that the Brownian motion is the only process that satisfies these conditions (see Dixit, 1993).

Finally note that an implication of (5) is that the shadow wage is non-stationary i.e. we are excluding the possibility of there being diminishing returns to education. This is probably unrealistic, but it simplifies the analysis considerably and does not change any of the fundamental results of the paper. For those interested, an appendix discussing the case of diminishing returns is available upon request.[**Included here as Appendix B**]

2.3 Uncertainty: A Simple Case

In this section, we solve simple model of education choice with uncertain returns. We make the same simplifying assumptions as for the certainty case (i.e. $u(c_t) = Y_t$, $F_t = 0$, $T = \infty$) so that we maximise (4) as before. The only difference is that the shadow wage now follows the stochastic process (5). We defer to the next section, all consideration of the impact on education of the possibility of borrowing and lending.

We solve the problem using the real option approach similar to that of Dixit and Pindyck (1994). Where returns are certain, this approach is unnecessarily complicated but it turns out to be the only practical method when returns are stochastic. The optimal time in school (S^*) will be a stochastic variable, so it is easier to express the control variable in terms of the level of the shadow wage at which it will be optimal to leave school. This variable, which we denote Y^* to distinguish from the threshold level under certainty, $\bar{Y}$, will be deterministic.

The intuition of the option approach is straight-forward. At any point in time, while the individual is still in school, she has the option of leaving school. This option itself has value. If she exercises this option she will loose the value of the option (because he cannot return to school in the future) and will receive a life time income that is a function of accumulated schooling. If she chooses not to exercise the option, she will receive whatever in-school income/utility she has and will wait until next period when she will have the chance to exercise the option again. By this time the value of the option will have changed in a manner related to the underlying process for the shadow wage given by (5). The resulting capital gain or loss is uncertain when viewed from the previous period. So exercising (or

[9] Judd (1998) models education risk explicitly as an implication of moral hazard. In this formulation, risk acts like a fixed cost of entry to the initial level of education and does not impact on the marginal effect of education above that level i.e. in our notation $\ln Y_s - \ln Y_o \backsim N(\mu s, \sigma^2)$

not) the option involves taking a gamble.

More formally, V_t in (4) can be thought of as the value of the option to leave school and start earning income at time t. Assuming that we don't exercise the option (i.e. for $t \in [0..S]$) then we can write equation (6) to describe how V will change over time.

$$pV = \phi + \frac{1}{dt}E\{dV\} \tag{6}$$

This Bellman equation (6) can best be understood as an arbitrage equation.[10] The right hand side is the return from staying in school (i.e. holding the option) for length of time dt. It consists of the dividend received over the period (which in our case is the constant utility derived from education) and the expected capital gain or loss in the value of the option over the period. Along the optimal path, this return must be equal to the return from the alternative investment strategy of selling the asset and investing the proceeds at the discount rate.

Because Y follows a Brownian motion so does V and using Ito's lemma[11] we can write the stochastic differential for V as

$$dV = \{gYV_Y + \frac{1}{2}\sigma^2 Y^2 V_{YY}\}dt + \sigma Y V_Y dz$$

Note that $E[dV]$ contains a term in the variance of Y. This has important implications for the effect of risk on decisions. On average shocks have no effect on Y i.e. $E[dY] = Yg$. However if $V_{YY} > 0$ they will have a positive effect on the change in the value of the option because the effect of a negative shock will be smaller in absolute terms than will the effect of positive shocks. The results is that V will trend up (down) over time due to repeated shocks to Y, if V_{YY} is positive (negative).

We can substitute dV into the Bellman equation, use the fact that $E[dz] = 0$ and divide by dt to get

$$pV = \phi + gYV_Y + \frac{1}{2}\sigma^2 Y^2 V_{YY} \tag{7}$$

The equation is a second order non-homogenous ordinary differential equation. It has a free boundary given by Y^*, the threshold level of the shadow wage at which the agent will choose to leave school. We can verify by substitution that the general solution will be

$$V = B_1 Y^{\theta_1} + B_2 Y^{\theta_2} + \phi/\rho \tag{8}$$

[10] We can also derive (6) from (4) rigorously using Bellman's Principle of Optimality (see Kamien and Schwartz, 1991, pp. 259-262 , for details).

[11] See Malliaris and Brock (1982) for a detailed discussion and Dixit (1993) for a slightly less formal explanation.

where θ_1 is the positive and θ_2 the negative root of the fundamental quadratic Q.

$$Q = \frac{1}{2}\sigma^2\theta^2 + (g - \frac{1}{2}\sigma^2)\theta - p \tag{9}$$

Economic theory provides three conditions (10) that determine the two constants of integration and the free boundary.

$$\begin{aligned} \lim_{Y \to 0} V(Y) &= \frac{\phi}{\rho} \\ V(Y^*) &= \frac{Y^*}{\rho} \\ V_Y(Y^*) &= \frac{1}{\rho} \end{aligned} \tag{10}$$

The first states that as the shadow wage tends to zero the individual will never leave education and so the value of being in school will simply equal the present value of the direct utility of perpetual education (ϕ/p). This implies that the negative root, θ_2, should have no influence on V, as Y tends to zero. If it did then the value of the option to leave school would tend to infinity. The only way of ensuring this is if $B_2 = 0$.

The second part of (10) is the "value matching" condition. When income reaches a certain threshold level (Y^*) the option is exercised, the individual leaves school and receives that income for life. The present value of this perpetual income stream is Y^*/ρ. Thus at time $t = S$, when the option is about to be exercised, its value will equal Y^*/ρ.

The third condition, the "smooth pasting" condition, states that for the threshold level of income to be chosen optimally, the net gain to any small changes in Y^* must have only second order effects. If we stay in school now while the market wage is Y, then we can leave school sometime in the future and earn (possibly) an even higher wage. The value of this option to leave, when the current shadow wage is Y, is given by $V(Y)$. When we leave school we gain Y/ρ but loose $V(Y)$. The net gain from leaving school when the (shadow) wage is Y is therefore $Y/\rho - V(Y)$, so the optimal choice of Y^* implies the smooth pasting condition.[12]

Using the value matching and smooth pasting conditions we can solve for Y^* generating the result which we state as Proposition 2 below.[13]

[12]This justification of the smooth pasting condition is intuitive but simplistic. A more complete, yet accessible, treatment can be found in Dixit and Pindyck (1994).

[13]For completeness we note that $B_1 = \frac{\phi}{\rho(\theta_1 - 1)}\left(\frac{\theta_1}{\theta_1 - 1}\phi\right)^{-\theta_1}$

Proposition 2 *If returns to education are normally distributed with mean g and variance σ^2, and if there is no borrowing or lending, it is optimal for the lifetime income maximiser to cease education when the shadow wage (Y) reaches a threshold level given by*

$$Y^* = \frac{\theta_1}{\theta_1 - 1}\phi$$

where θ_1 is the positive *root of Q in (9), ρ is the discount rate and ϕ is intrinsic utility of education. Furthermore we have $\frac{\partial Y^*}{\partial g} > 0$, $\frac{\partial Y^*}{\partial \rho} < 0$, $\frac{\partial Y^*}{\partial \sigma} > 0$ and $\lim_{\sigma \to 0} Y^* = \bar{Y}$ and $\lim_{\sigma \to \infty} Y^* = \infty$.*

Proof. The expression for Y^* follows directly from solving for Y^* from (10) given (8) and (9). The derivatives follow by application of the implicit function theorem to (9). The limits follow when L'Hôpital's Rule is applied to the expression for θ_1. ■

As in the case of certainty, sufficient conditions for $Y^* > 0$ are that $\phi > 0$ and $\rho > g$. If the latter were not the case, school would always provide a better return (on average) and it would be optimal to stay in school for ever. As we would expect, Y^* is an increasing function of g and a decreasing function of ρ. Thus high returns to education will cause individuals to stay in school longer whereas a high discount rate will induce them to leave earlier.

The threshold level of the shadow wage (Y^*) is also an increasing function of risk, so the threshold is higher than under certainty $(Y^* > \bar{Y})$. Furthermore Y^* reduces to $\bar{Y}$ as $\sigma^2 \to 0$. We can also show that Y^* becomes infinite as $\sigma^2 \to \infty$, implying that the agent will never leave school.

The fact that risk increases the amount of schooling is, perhaps, surprising. Using the investment analogy, one might have expected less investment in human capital as the risk associated with that investment rose. Our result is due to the fact that leaving school is an irreversible decision. Risk creates a value to waiting because if we stay in school we have the option to leave next period in order to take advantage of a good draw from the distribution of returns or to remain in education so as to avoid a bad draw.[14] Uncertainty has an asymmetric effect, increasing the potential upside payoff from the option, but, because we will stay in school if the market wage turns out to be low, the downside payoff is unchanged. This effect becomes stronger as the riskiness of education increases. Indeed when risk becomes infinite, the agent will never want to exercise the option to leave.

This result is in line with what we would expect from financial option theory. Increased risk in the underlying security tends to increase the value of the option because increased

[14] We see the current draw before deciding to leave or not.

variability implies that the option is more likely to be "in the money" at some point in the future.

Note also that risk has an effect on the education decision even though the agent is apparently risk neutral i.e. $u(c) = c$. Again the reason is that risk in the presence of an irreversibility creates a value to waiting — even for the risk neutral investor. What is at issue is not the avoidance of risk, but the trade-off between current and future risk. Another way of seeing this is to note that while instantaneous utility is linear, lifetime utility, V, has $V_{YY} > 0$. In fact the coefficient of relative risk aversion for lifetime utility is equal to $-\theta_1 < -1$. It is as if the irreversibility has changed a risk neutral agent into a risk lover.

All this is intuitive, but note that it has the implication that the individual will accumulate more human capital when the risk associated with that investment is higher. This prediction contrasts with that of the portfolio model of Williams (1979). In his model, an increase in the risk of human capital (or any other asset), would cause the individual to accumulate less of it, other things being equal. The reason for this difference is the nature of the choice facing the agent. His approach treats education as occurring continuously and at the same time as work. There is no irreversibility, the agent can come in and out of education as she pleases for zero cost (other than forgone wages). Because there is no irreversibility there is no value to waiting.

Figure 1 illustrates the solution of the model. The graph shows the function $V(Y)$, the value of the option to wait and the function $\Omega(Y) = Y/\rho$, the value of leaving education when the market wage is Y. At the optimal point, V and Ω are equal and meet as tangents. For shadow wages less than the optimal ($Y < Y^*$), the value of the option to wait (V) is greater than the life-time utility from leaving now (Ω), so the individual remains in education. When the shadow wage is zero, the optimal decision would be to say in school for ever, generating a life-time utility of ϕ/ρ. As the shadow wage increases, Ω, the gain from leaving also increases. But so does the cost of leaving i.e. the value of the option to leave at some point in the future. At the optimal threshold the two are equal. Note it may appear from the diagram that it is optimal to remain in school if $Y > Y^*$. This is not true. Because of the value matching condition, the value of lifetime income is given by $\Omega(Y)$ when $Y > Y^*$, so that the full function V is given by $[abd]$. The line segment $[bc]$ is irrelevant.

2.4 Risk Aversion

The previous sections assumed that the agent was risk neutral. In this sub-section we allow for individuals to have preferences over risk. We continue to assume that there is no borrowing or lending, so $c_t = 0 \;\; \forall \;\; t < S$ and $c_t = Y^* \;\; \forall \;\; t > S$.

The solution is more or less the same as in the previous sub-sections. The Bellman equation is given by (6) as before and so will have the same general solution as before. The form of the utility function only affects utility after leaving school as we have precluded the possibility that the agent may borrow against future income in order to subsidize consumption before graduation.

In fact the only difference between this formulation and the previous section is the boundary conditions. When the individual exercises his option to leave school he will receive lifetime utility equal to Ω. We can calculate this by direct integration assuming that income is constant at Y^*after graduation and assuming that $u(c)$ is CRRA with $u(c) = c^{1-\gamma}/(1-\gamma)$, $\gamma > 0$.

$$\Omega(Y^*) = \int_S^\infty e^{-\rho(t-S)} \frac{c^{1-\gamma}}{1-\gamma} dt = \frac{[Y^*]^{1-\gamma}}{(1-\gamma)\,\rho}$$

As before we assume $V(0) = \phi/\rho$. The value matching and smooth pasting conditions become, $V(Y^*) = \Omega(Y^*)$, and $V_Y(Y^*) = \Omega_Y(Y^*)$ respectively and have the same interpretation as in the last section. The result is qualitatively the same as before. All the derivatives of Y^* have the same sign as before; γ just acts as a scaling factor.

Proposition 3 *When (i) there is no borrowing or lending; (ii) returns to education are normally distributed; (iii) preferences are $u(c) = c^{1-\gamma}/(1-\gamma)$, the threshold level of the shadow wage at which it is optimal to cease education is given by*

$$Y^* = \left[\frac{\phi\theta_1(1-\gamma)}{\theta_1 - (1-\gamma)}\right]^{\frac{1}{1-\gamma}}$$

where θ_1 is the positive *root of Q in (9). As before we have* $\frac{\partial Y^*}{\partial g} > 0$, $\frac{\partial Y^*}{\partial \rho} < 0$, $\frac{\partial Y^*}{\partial \sigma} > 0$

Proof. The expression for Y^* follows directly from solving the Value Matching and Smooth Pasting conditions for Y^* given (8) and (9). The derivatives follow by application of the implicit function theorem to (9). ■

3 Consumption Smoothing and Education Choice

In this section we allow the individual to borrow against future income in order to subsidize consumption while in full time education. The absence of liquidity constraints raises the possibility that an individual will stay in education longer, borrowing to fund consumption during the school years and paying back the debt from higher future earnings.

We assume that the individual maximises lifetime utility V from (1) with the added assumptions that ϕ and F are constant and that $T = \infty$ to give (11).

$$V = E\left\{\int_0^S e^{-\rho t}\left\{u(c_t)+\phi\right\}dt + \int_S^\infty e^{-\rho t}u(c_t)dt\right\} \tag{11}$$

Utility is maximised subject to the budget constraint (2) and the stochastic returns to education (5) which we rewrite in differential form as (12).

$$\begin{aligned}
dA_t &= (rA_t - F - c_t)dt \quad \forall \quad t\in[0..S] \\
dA_t &= (rA_t + Y_t - c_t)dt \quad \forall \quad t\in[S..\infty] \\
dY_t &= gY_tdt + \sigma Y dz \quad \forall \quad t\in[0..S] \\
dY_t &= \alpha Y_t dt \quad \forall \quad t\in[S..\infty]
\end{aligned} \tag{12}$$

The first equation in (12) states that an individual in school must finance consumption and (constant) school fees, F, by running down asset balances. The second equation states that after graduation asset balances can be rebuilt using earned income. The third equation in (12) shows the evolution of the shadow wage while the individual is in school. As before we assume that the shadow wage evolves according to a geometric Brownian motion so that returns to education are normally distributed and the level of the wage upon graduation is lognormally distributed. The final equation in (12) states that earned income will grow at rate α after leaving school. In order to keep things simple and facilitate an analytical solution we assume that this growth rate is deterministic. We avoid a corner solution by assuming $\phi > 0$ and ensure convergence of the integral by assuming that ρ is greater than g, r and α.[15] Finally, it is worth noting that this formulation precludes insurance or any hedging of labour income uncertainty. The only other asset in the model has returns that are not correlated with the returns to education.[16]

The Bellman equation associated with (11)-(12) is given by (13) where subscripts indicate partial derivatives. Note that the value function, V, is now a function of two state

[15]This is a sufficient rather than necessary condition.

[16]Judd (1998) provides a justification for this assumption by modeling education risk as being due to moral hazard.

variables, the shadow wage (as before) and also the level of net financial assets.

$$\rho V = \max_c \left\{ u(c) + \phi + V_Y gY + V_{YY}\frac{\sigma^2}{2}Y^2 + V_A(rA - F - c) \right\} \tag{13}$$

As before we think of school attendance as being equivalent to possessing an option to leave school and earn a salary. The value of this option, V, evolves according to (13). There are some differences with the Bellman equations of previous sections. Firstly, the per period payoff ("dividend") of being in school is now expressed in terms of utility $u(c) + \phi$, where the first term represents the utility of consumption while in school and the second represents the intrinsic utility (or disutility) from being in school. Secondly, school fees (F) must be deducted from the cash available for consumption. The third, and most important difference, is that the individual is able to subsidize consumption while in school by running down asset balances. To this end the individual can choose the level of consumption while in school to maximise lifetime utility or equivalently to maximise the value of the option to quit school.[17]

Assuming that the individual will always choose consumption optimally given assets and the wage (i.e. education) then we have the standard first order condition for intertemporal consumption smoothing $u_c(c) = V_A$. If we assume that utility is CRRA, $u(c) = c^{1-\gamma}/(1-\gamma)$, and substitute the first order condition into the Bellman equation, we get equation (14) that describes the stochastic evolution of the option to quit school, conditional on assets and the wage.

$$\rho V = \frac{\gamma V_A^{\frac{\gamma-1}{\gamma}}}{(1-\gamma)} + \phi + gYV_Y + V_{YY}\frac{\sigma^2}{2}Y^2 + V_A(rA - F) \tag{14}$$

We can verify by substitution that the solution to (14) is given by (15) where θ_1 is the positive root of Q in (9) and we have eliminated the negative root in order to impose finite value on the option.

$$\begin{aligned} V(A,Y) &= \frac{B_0}{(1-\gamma)}(A - F/r)^{1-\gamma} + B_1 Y^{\theta_1} + \phi/\rho \\ B_0 &= \left(\frac{\gamma}{r\gamma - r + \rho}\right)^{\gamma} \end{aligned} \tag{15}$$

The last two terms are the same as (8), the value function for the simplest case. (Nonetheless the value assigned to the option will be different as B_1 will be different). The first part

[17] The two are equivalent due to Bellman's Principle of Optimality. See Kamien and Schwartz, 1991, op. cit.

of (15) represents the life-time utility derived from consumption out of financial assets. In effect the introduction of financial assets creates a lower bound for life-time utility. The worst case for the individual is that she never leaves school. In this case she would consume out of assets for ever and enjoy the direct utility of schooling generating a life-time utility of $V = B_0(1-\gamma)^{-1}(A - F/r)^{1-\gamma} + \phi/\rho$. Only in the case where the the option to leave school has positive value, will she exercise it at some point, leave school and achieve a life-time utility strictly greater than the lower bound. Note that this suggests that the fundamental structure of the problem is not altered by the elimination of liquidity constraints nor is it dependent on the precise specification of preferences (see appendix).

When the individual exercises her option and leaves school she will receive a certain salary which will generate a certain lifetime utility, Ω (i.e. the second integral in (11)). The exact value of of post school life-time utility, $\Omega(Y, A)$, depends on how wages evolve after leaving school. Using the usual argument we can construct (16), a Bellman equation for Ω.

$$\rho\Omega = \max_c \{u(c) + \Omega_Y \alpha Y + \Omega_A(rA + Y - c)\} \tag{16}$$

This equation is similar to (13) but different in interpretation. The individual once again chooses consumption so as to maximise the value of life time utility conditional on assets and the process of income. Here, however, the wage is actually received by the individual as she is working, whereas for equation (14), the Y was the shadow wage i.e. the wage the individual would get the moment he left school. As the individual has left education at this stage, there is no optimal stopping problem and there are no value matching or smooth pasting conditions. The necessary boundary conditions are provided by the assumption that the integral in (11) converges i.e. life time utility is finite.

If we assume that consumption is optimally chosen after leaving school and that utility is CRRA, then (16) has the familiar solution (17)

$$\Omega(A, Y) = \frac{B_0}{1-\gamma}\left(A + \frac{Y}{r-\alpha}\right)^{1-\gamma} \tag{17}$$

which allows us to state Proposition 4.

Proposition 4 *Consumption will jump up upon graduation.*

Proof. For CRRA preferences $c^* = V_A^{-1/\gamma}$ before graduation and $c^* = \Omega_A^{-1/\gamma}$ after graduation. From (17) and (15) we have $V_A(A, Y^*) > \Omega_A(A, Y^*)$ ■

Proposition 4 works because while (17) has the same form as the first term of (15), generated by the common consumption smoothing structure to both problems, there is a crucial difference between the two. We can view (17) as stating that life-time utility is a function of total wealth, which is equal to the sum of financial wealth, A, and human capital $Y/(r-\alpha)$. This follows from the assumption that the optimizing individual will borrow against future income in order to smooth consumption.

The situation is different before graduation, however. The human capital term is absent from the first term in (15). The reason is that, strictly speaking, the individual has no marketable human capital, before graduation. What she does have is the option to acquire marketable human capital (by leaving school) at some date in the future. The value of this option appears additively in the value function and not within the parentheses in the same manner as A. This is because we assume that the option to leave school is an asset which, while it may have value, nevertheless cannot be traded or used as a collateral for a loan i.e. the value of the option is absent from the budget constraint (12). In that sense there is a liquidity constraint in this problem albeit one that is entirely realistic – but not apparent at first glance. Essentially the reason for the jump in consumption is that graduation converts the option (which cannot be traded) into human capital (which can), so consumable wealth jumps.

Finally, note also that V is only defined when $rA > F$ i.e. when assets are greater than the present value of future school fees. If this condition is violated then the nature of the problem is fundamentally altered. The reason is that the individual must be able pay her way in school or else she will forced to leave school. The problem is no longer one of optimal stopping as there is no longer a free choice of when to exercise the option. Furthermore, while it may appear from the requirement that $rA > F$ that there is some restriction on borrowing against future income, this is not so. As can be seen from the integral version of the budget constraint (2), the agent is free to borrow and lend unlimited amounts subject to life-time budget balance. The only liquidity constraint is that the individual cannot sell the option itself in an attempt to boost consumption.

3.1 Solving for Y^*

Now we are in a position to characterise the threshold level of the shadow wage and to show how it is affected by the other parameters of the model. We impose the value matching and smooth pasting conditions (18) both of which have the same interpretation as before.

$$\begin{aligned} V(Y^*, A) &= \Omega(Y^*, A) \\ V_Y(Y^*, A) &= \Omega_Y(Y^*, A) \end{aligned} \tag{18}$$

The result is a system of two non-linear simultaneous equations that jointly determine B_1 and Y^* conditional on A and F and the parameters of the model.

Proposition 5 *When borrowing is possible and consumption is chosen optimally, an individual will leave education when the shadow wage is Y^*, where Y^* is increasing in the mean return and risk of education, increasing in education fees and ambiguously affected by wealth*

$$Y^* = Y^* \begin{pmatrix} g, & \sigma, & A, & F \\ + & + & +/- & + \end{pmatrix}$$

.

Proof. See Appendix ■

In order to illustrate the model we present a numerical solution of (18) and simulate the effects of changes in various parameters. Table 1 presents the baseline values of the parameters used in the simulation. All are plausible, if conservative, values. For simplicity we simulate the model assuming $\gamma = 1$ i.e. log utility.[18] We assume that the expected rate of return on education (g) is 7% per annum which is in line with OLS estimates but less than most IV estimates (see Card, 2001). The estimate of risk (σ) at 2% seems reasonable given our choice of g. It is also in line with estimates provided by Harmon et. al. (2001) and Conneely and Uusitalo (1999) but less than the 7.5% estimated by Chen (2001). The discount rate (ρ) and interest rate (r) are set equal so that consumption is constant through time (apart from the once-off jump at graduation). This is convenient because it ensures that financial asset balances will be constant, enabling us to ignore the distinction between initial balances and balances at graduation (see below). We choose two baseline values for financial assets, the first ("rich") ensures that asset income is 2.5 times fees (F), whereas the second ("poor") sets asset income to be 1.5 times fees.

Together Y_0 and ϕ act as numeraires for the problem. The parameter Y_0 can be thought of as representing the income received by an individual who leaves school immediately after the end of compulsory education. Without loss of generality, Y_0 is set equal to unity so that Y is expressed in terms of a multiple of the wage associated with minimum education. We set the intrinsic utility of school so as to ensure that in the absence of uncertainty, an individual would optimally choose to leave after exactly 2 years of post compulsory schooling.

Figure 2 shows the value matching and smooth pasting conditions evaluated at the parameter values in Table 1. The two functions V and Ω are equal and meet as tangents

[18] In fact we assume that $u(c) = \ln c - k$, where k is a constant set so that lifetime utility is positive for all possible parameter values.

at the optimal point. For shadow wages less than the optimal ($Y < Y^*$), $V > \Omega$, so the individual remains in education until $Y = Y^*$. Note given our parameters, $Y^* = 1.15$ implying an average of just over two years of post compulsory schooling.

Figure 3 shows the effect of education choice on the time path of consumption. Before graduation, the individual lives off asset income, pays school fees and consumes the remainder ($c = rA - F$). Asset balances are constant so this strategy is sustainable even if the individual never leaves school.[19] Following graduation, fees no longer have to be paid and income comes on stream, so consumption jumps to a new higher level ($c = rA + Y^*$) and remains there forever.

3.2 The Effect of Risk and Return

Proposition 5 states that the effect of increases in the expected rate of return to education is unambiguously positive. Figure 4 illustrates the point numerically for both rich and poor individuals. Note that for all but the highest expected returns, poor individuals would optimally decide to leave school before the end of compulsory education were that possible. The imposition of the compulsory schooling law constrains $Y = 1 > Y^*$, reducing their life time utility. It is worth reemphasising that the rich and poor are identical in all respects other than in their initial endowment of wealth. The poor individuals have such low wealth that they can boost lifetime consumption by leaving school and working even at a low wage. We discuss this further below.

The effect of risk on education is also positive for essentially the same reason as before: an irreversible decision in the presence of risk creates an incentive to wait. Allowing for consumption smoothing does not change this fundamental result (although it will change the magnitude of the effect). The result is also independent of the structure of preferences. As shown in the appendix, a sufficient condition for the result to hold is that $\Omega_{YY} < 0$. This is certainly true for CRRA preferences and will likely be true for all "well behaved" preferences.

Figure 5a illustrates the effect numerically. The effect is positive, but is much smaller than the effect on Y^* of the difference in assets between rich and poor. For the poor, risk must be very high before it provides sufficient incentive to stay in school. When you are poor the value to waiting is very low when compared to the cost of waiting in terms of income forgone. Figure 5b presents the effect for the rich only in order to more clearly illustrate the positive non-linear effect of increasing risk.

[19]If $\rho > r$ the sustainable strategy is for consumption and assets to fall continuously reaching zero only at infinite time.

Because the evolution of income is stochastic, there is no expression for S corresponding to Proposition 1. When returns are stochastic, S^* will be a random variable and the best we can do is to describe its distribution. An analytical description of the full distribution is complicated (although see Malliaris and Brock (1982) for an example). Instead we describe it numerically by simulating the system.[20]

Figure 6 and Table 2 present the results of this simulation for various different levels of risk in the case of rich individuals. As can be seen, increasing risk leads to an increase in S^*. This is to be expected given that Y^*, the target level of the shadow wage will have increased. It is also clear that the variance of S^* will rise. Again this is intuitive: as the process for the shadow wage gets more uncertain, the time it takes for that process to reach any given level becomes more uncertain. What is more surprising is that the distribution of S^* becomes increasingly skewed at higher levels of risk. The reason is that direct effect of higher risk on the mean and variance of S^* makes higher values of S^* relatively more likely than lower values. This coupled with the fact that S^* is bounded at zero results in a skewed distribution.

3.3 The Effect of Wealth and Fees

Proposition 5 tells us that an increase in fees reduces Y^* and causes individuals to leave education earlier on average. This is exactly what we would expect. Similarly, we would expect that an increase in initial financial wealth would lead to longer stays in full time education. It turns out that while this is true for most plausible values of the parameters of the model, the effect of wealth on education is, nevertheless, theoretically ambiguous. The ambiguity stems directly from the irreversible nature of the school leaving decision.

There are two opposing forces at work when wealth increases. The direct effect is to relieve the budget constraint, enabling the student to consume more before graduation, facilitating a longer spell in education. This is the mirror image of the effect of a change in fees. There is, however, an additional effect of an increase in wealth. Consumption smoothing implies that some of the extra wealth will be used to finance consumption after graduation. This will reduce the marginal utility of labour income after graduation i.e. Ω_Y falls. However, the smooth pasting condition requires that $V_Y = \Omega_Y$. The only way this can be achieved given that $V_{YY} > 0$ is for Y^* to fall. In essence the increase in assets reduces

[20] If the individual starts with income Y_0 how long will it take for income to reach the threshold value (Y^*) when it evolves according to (5)? The probability that an individual will still be in school at time t (so that S^* greater than t) is equal to the probability that the income process will not have reached the trigger level at time t (so $Y_t < Y^*$). This implies that $P(S^* \leq t) = 1 - P(Y_t < Y^*) = 1 - \Phi(Z_t^*)$ where $Z_t^* = (\ln Y^* - \ln Y_0 - \mu t) \,/\, (\sigma\sqrt{t})$ and Φ is the c.d.f. of a standard normal random variable.

the marginal value of waiting, inducing the individual to leave earlier. As we show in the appendix, the second effect is dominated by the direct effect for most parameter values, generating a positive effect of wealth on education.

Note also that the asset balance referred to in Proposition 5 is the balance at graduation i.e. A_t where $t = S^*$. As we do not know S^* with certainty, we cannot in general say much about asset balances at graduation. It is not difficult to show however that balances at graduation are monotonic in initial balances i.e. $\partial A_{S^*}/\partial A_0 > 0$. Furthermore, when the discount and interest rates are equal ($\rho = r$) then asset balances are constant through all time so $A_{S^*} = A_0$.

Figure 7 illustrates the effect of financial wealth on the threshold level of the shadow wage. The effect is clearly positive and almost linear. The figure is parameterised in terms of ratio of asset income to fees (rA/F). If this ratio is less than one then the education choice model is not relevant as the individual cannot afford any education (beyond the compulsory minimum level). The figure shows that for low wealth ($rA/F < 2.33$), the individual would optimally choose to leave education before the completion of compulsory schooling, if this were possible. These individuals can afford to continue education (i.e. $rA > F$), but would prefer not to. The reason is simply that, with such low wealth, they need labour income to provide consumption.

This desire is a direct consequence of being able to smooth consumption, while being unable to use the value of the option itself to subsidise consumption before graduation. The desire to leave school early would also be mitigated if the irreversibility were eliminated. In that case individuals could leave school early in life in order to build up asset balances and then return later to add to human capital.

3.4 Some Policy Implications

In the section of the paper we use the model to examine the impact on individuals' education decisions of some simple stylised government policies. We can model the direct effect of an education subsidy as a reduction in F ("fees" in the model), perhaps even becoming negative. We already know from Proposition 5 that the effect of a reduction in fees is to increase the threshold shadow wages, and thus lead to an increase in schooling.

If the subsidy is financed from general taxation, then there will be no other effects on the individuals education choice. Many real world tuition finance programmes, however, require the student to pay back some of the tuition after graduation. We think of three broad types of tuition payment plans. Firstly, tuition could be paid back in fixed installments as with a standard loan repayment. In the context of the model, this would equivalent to levying a lump-sum tax on earnings after graduation. Alternatively, the repayments could be fixed as proportion of income. In the model this is the equivalent of a proportional tax on income. Alternatively, a tuition payment plan could combine proportionate and lump sum elements equivalent to a progressive (or even regressive) wage income tax.

It turns out that we can easily accommodate the three different taxes in the model of section three. The state variable is still Y, but now we interpret it now as being the (shadow) wage gross of taxes/repayments. We define a new variable $x = f(Y)$ which is the net wage received upon graduation.

$$x = f(Y) = Y - \tau Y^{\varepsilon} \tag{19}$$

The function $f(Y)$ summarizes the relevant parameters of the tax system. The parameter ε is equal to the ratio of the marginal tax rate to the average tax rate. It represents the extent to which the tax system is progressive or regressive. For lump-sum taxes $\varepsilon = 0$ (i.e. perfectly regressive) and we interpret τ as the amount of the lump sum tax. For proportional taxes, $\varepsilon = 1$ and we interpret τ as the proportionate tax rate. For regressive taxes, $\varepsilon \in [0, 1)$, the marginal tax rate is less than the average tax rate for all income. For a progressive tax system, $\varepsilon > 1$, marginal tax rates are higher than average tax rates for all incomes.[21]

The variable x directly effects the problem only through Ω, the utility after graduation. The structure of the option is unaffected as is the form of the function V which must still solve the Bellman equation (13). The value matching and smooth pasting conditions will change to $V(A, Y) = \Omega(A, f(Y))$ and $V_Y(A, Y) = \Omega_x(A, f(Y))f_Y(Y)$ respectively.[22] This modification to the model allows us to state Proposition 6.

Proposition 6 *The imposition of either a lump sum or a proportional tax will lead to an increase in Y^*. An increase in the degree of progressivity of the tax system could lead to an increase or decrease in schooling depending on the degree of risk aversion and the degree of progressivity.*

Proof. See Appendix ■

At first glance this may seem a curious result. We might have expected a increase in a proportional tax to reduce the benefit of schooling and so lead to less education. In fact the tax does reduce the benefit of schooling, so that the value of the option to wait falls. But the value of leaving school, Ω, falls by more. The net result is that school becomes

[21]Note for simplicity we assume that capital gains and net interest payments are not taxable income.

[22]Note that the actual value of the option will be affected via the smooth pasting and value matching conditions, leading to a different value for the constant B_1.

relatively more attractive, and the individual stays for longer. This is illustrated in Figure 8. Following the imposition of a tax, the individual seeks to maintain living standards by boosting gross wage. The only way to to this is to stay in school longer. In essence, we have an income effect without any associated substitution effect. There is no counteracting substitution effect because both a lump-sum and a proportionate tax will not change the risk and return associated with continuing to the next level of education. In fact, it is straight forward to show that if the tax revenue is returned to the individuals in a lump-sum, the income effect will be nullified, thus compensated changes in proportional taxes will have no effect on education attainment.

The situation can be different when taxes are progressive (or regressive). In that case, the after-tax risk and return to education will be different for different levels of education. For example, a progressive tax will levy a higher proportional charge on higher incomes, so that the risk and return associated with proceeding from a lower to a higher level of education will both be reduced. This in turn, will reduce the value of the option to wait. If large enough, this substitution effect can overcome the income effect and lead to fall in education. As we show in the appendix, a necessary (but not sufficient) condition for this to occur is that $(\theta - \varepsilon)\ln Y^* < 1$. This condition illustrates how risk (via θ) interacts with the degree of progressivity of the tax system to determine the strength of the substitution effect. When higher risk (lower θ) is combined with higher progressivity and also with lower education choice (e.g. due to poverty), the condition will hold and progressivity can have a negative (uncompensated) impact on education choice.

We illustrate this in Figure 9, where the baseline parameters are from Table 1. The horizontal axis represents the parameter ε which goes from zero (representing a perfectly regressive lump sum tax) through to unity (representing a proportional tax) and beyond (representing progressive taxation).

For individuals rich enough to already be in education beyond the compulsory minimum, increasing the progressivity of the tax system will lead to increased education i.e. income effects dominate. For individuals, whose lack of financial wealth would induce them to leave before the end of compulsory education, if that were possible, increasing the progressivity of the tax exacerbates the problem. For them the substitution effect is dominant. The fact that continuing education will be taxed at an ever increasing rate induces them to leave earlier. If they were wealthier they would react to the declining net return to education by staying school longer in order to boost income. Lack of financial wealth makes that strategy undesirable because consumption while in school is so low.

It is worth comparing our results with the rest of the literature. Trostel (1993) calibrates a dynamic general equilibrium model of human capital accumulation (without uncertainty) to show that a proportional (compensated) wage tax can have a negative impact on human capital accumulation. A crucial assumption for this result is that labour supply is elastic. The imposition of the tax reduces labour supply, and thus the effective return to human capital.

Lin (1998) shows that in a non-stochastic OLG model, an uncompensated increase in a (proportional) wage tax can reduce human capital accumulation. This result depends crucially on a capital market channel that is absent in our model. An increase in wage taxes can reduce savings, leading to a lower stock of physical capital. This in turn leads to higher interest rates which makes investment in human capital less attractive at the margin. The negative effect disappears if tax revenue is redistributed to tax payers. In this case their income and saving remain the same so interest rates remain unchanged. Heckman et. al. (1998) use a similar model to examine the impact on human capital accumulation of a consumption tax.

Eaton and Rosen (1980) is one few papers to consider explicitly the effect of taxation in model of education choice with uncertainty. They show that in a two period model, the imposition of a proportional (uncompensated) wage tax will have an ambiguous effect on education. However, when preferences exhibit constant relative risk aversion and initial wealth is sufficiently high, they show that an uncompensated proportional wage tax has a positive effect on human capital accumulation.[23]

We noted above that taxes have a positive effect on education attainment of the rich, but may have a negative effect on the education of the poor. This suggests that we could boost education attainment by simply levying taxes. This raises the interesting prospect that the optimal policy mix aimed at increasing the level of education throughout the society would be to give education subsidies to the poor only, and tax the rich at higher rates (via high degree of progressivity). Both policies would independently boost education. Of course the utility of the rich would fall as a result of the imposition of the tax.

The policies most often contemplated, however, involve education subsidies that must be paid back after graduation – even by poor individuals. We analyse the effect of such policies on poor individuals in Figure 10. We assume that individual receives a per period subsidy equal to the cost of education for as long as she is in school. After graduation she pays a constant lump-sum repayment throughout her life, so that the (expected) present value of the payments are equal to the value of the subsidy.[24] As before, the baseline parameters

[23] In their model uncertainty is multiplicative in income, so the marginal product of human capital is stochastic but the rate of return is deterministic i.e. $Y(s) = \varepsilon W(s)$ where ε is stochastic (mean one) and W is deterministic.

[24] There is a technical difficulty as S^* is stochastic. We assume that payments are set in advance of graduation and are known by the individual to be equal to the present value of fees when $S = E(S^*)$ with

are from Table 1.

We can see from the figure that when the subsidy reaches about 35% of fees, the poor individuals are induced to stay in education beyond the minimum. When the subsidy reaches 46%, these individuals will choose the same level of education as the rich individuals would in the absence of the tax and subsidy. Of course these results are extremely sensitive to the parameters, in particular, the precise definition of "rich" and "poor". Nevertheless, they do illustrate how the effects of a lack of wealth can be overcome by a policy that imposes no cost on the government over the lifetime of the individual.

4 Discussion and Extensions

The model we have presented was structured so that it would yield analytical solutions. In this section we argue that the results of section 3 are quite robust and that most (but not all) of the extensions that we might contemplate would complicate but not elucidate.

The most obvious extension to the model would be to account for finite life and education opportunities i.e. $S \leq T < \infty$. It turns out that it is very easy to accommodate this change. If we do not insist on a deterministic length of life, we can allow death/retirement to arrive according to a Poisson process with parameter λ. As is well known, this is equivalent to increasing the discount rate from ρ to $\rho + \lambda$ and keeping $T = \infty$. So the qualitative results will be exactly the same.

For a deterministic death/retirement date, time becomes the third state variable of the problem. A term involving V_t will appear in the Bellman equation (13) and the value matching and smooth pasting conditions will be $V(A, Y, t) = \Omega(A, Y, t)$ and $V_Y(A, Y, t) = \Omega_Y(A, Y, t)$ respectively. This free boundary problem will have to be solved numerically as there will be no closed form solutions for V or Ω. However, the basic results of Proposition 5 will not be affected as the structure of the problem is unchanged. There is a still an option. Its value still increases in uncertainty. It is this value of waiting that drives all the main results of the model. All that has changed is that the option now has a finite expiry date. In fact the problem is now very close to the Black-Scholes analysis of a financial call option.

A variant on this would be to place an upper bound on schooling i.e. $S < T$. In fact this is not necessary. The fact that individuals do not stay in education for their entire lifetimes is not because it is physically impossible, but because it is not optimal to do so in the face of diminishing returns to education. We discuss this further below.

zero subsidy.

Another extension is to include post schooling risk i.e. that the income process after graduation should be stochastic. A consumption smoothing problem with stochastic labour income would require numerical solution. But again, the overall structure of the option problem would not change. With irreversibility, there would still be a value to waiting due to the uncertainty regarding the shadow wage (i.e. uncertainty before graduation). The value matching and smooth pasting conditions that determine the value of the option would still be defined in terms of the same function V. The introduction of uncertainty post graduation leaves the structure of the problem unchanged. The only difference would be that the function $\Omega(A, Y)$ would not have an analytical representation and its value would be affected (negatively) by the variance of the wage process after graduation.

Comparison of our model with that of Williams (1979) suggests an interesting extension. Because he treats education as occurring continuously and at the same time as work, his model more accurately reflects the structure of informal and "on-the-job" training whereas ours reflects the structure of more formal education.[25] Inspite of this difference, one can view his model as being similar to ours except that he assumes that an individual can return to school (part-time) at zero cost (other than forgone wages). By contrast, in our model, return is impossible. Furthemore we argued that this difference in structure explained the difference in results of the two models. This suggests that we could devise a model in which returning to full time education from the labour market was possible, but at a cost. We suspect that such a model would embed our model and that of Williams (1979) as special cases.

The most interesting extensions to the model relate to the stochastic process for the shadow wage. We choose the geometric Brownian motion because its simplicity facilitates solution of the model, but nevertheless is consistent with empirical education research, being a continuous time version of a Mincer equation with random coefficients. However, we could improve upon this specification along several dimensions. Most obviously we could introduce diminishing returns to education. If we model returns as diminishing in time spent in school, then we will make time a state variable, necessitating a numerical solution. Alternatively we could specify returns to be diminishing in higher levels of Y. This results in an Ornstein-Uhlenbeck process.[26] Unfortunately, it is difficult to relate the parameters of this process to real world education decisions. Furthermore, the qualitative solution is not affected by using an Ornstein-Uhlenbeck process in place of a geometric

[25] In fact in Williams (1979) the agent is obliged to accummulate at least some human capital in every period in order to avoid a corner solution.

[26] An example: $\frac{dY}{Y} = (\alpha_0 - \alpha_1 Y)ds + \sigma dz$. See Malliaris and Brock (1982) p. 270 for a review of these processes.

Brownian motion. The value of the option to wait still increases in uncertainty – and all our qualitative results hold.[27]

A more fundamental change would be to allow σ to vary with S. The model of section 3 assumed that the distribution of the returns to education was the same for all levels of education (so that the mean and variance of the shadow wage rose linearly with education). This makes the problem tractable and, given the relative paucity of information on this issue, seems reasonable. But it is at least possible that risk could increase or decrease with education. If further education decreased risk, we might expect individuals to choose further education as a form of insurance. However, to be set against this is the fact that lower risk would decrease the value of the option to stay in school suggesting that it would be optimal to leave earlier. Analysing how these two effects interact would make for a interesting extension to our model.

Another useful extension could be made by explicitly considering "sheep-skin" effects i.e. the possibility that the mean and variance of education returns may be function not of time in school, but of qualifications attained. A related issue is to recognise that continuing in education is not automatic, but requires passing exams. How exactly these considerations would affect education choice is unclear, but again we suspect that our basic qualitative results would still hold for the same reason as before – irreversibility in the presence of uncertainty creates a value to waiting. It is this value that generates our results.

5 Conclusions

In this paper we apply the techniques of option theory to the study the education decisions of individuals when the returns to education are uncertain. We view an individual in school as possessing an option to leave at any time and take up work at a wage related (stochastically) to the time spent in school. Once that option is exercised, the individual cannot return to school.

We show that high returns to education will cause individuals to stay in school longer whereas a high discount rate will induce them to leave earlier. Furthermore we also show that increasing risk will cause an individual to delay leaving school. This result is not dependent on the risk preferences of agents as it holds for risk neutral agents also. On the face of it, this is curious result, we would expect that higher risk would lead to less investment in human capital. The result stems from treating education as an option. Once the agent leaves school, he can never return. Higher uncertainty, therefore provides an incentive to delay leaving so as to see if uncertainty may resolve itself.

[27] The solution to this problem is provided in an appendix that is available from the authors upon request.

We also showed that introducing the ability to borrow did not change the fundamental structure of the model. We showed that education attainment usually, but not always, increases in initial wealth. We also showed that increased labour income taxation would induce individuals to stay in school longer, unless those taxes were highly progressive and/or individual is highly risk averse.

A Proof of Propositions 5 and 6

We make use of the matrix version of the implicit function theorem. Let $G_1(Y^*, A, F, r, \theta_1, \rho)$ and $G_2(Y^*, A, F, r, \theta_1, \rho)$ be the system of implicit equations that jointly determine Y^* and the constant of integration B_1.

$$\begin{aligned} G_1 &= \Omega_Y(Y^*, A) - V_Y(Y^*, A) = 0 \\ \\ G_2 &= \Omega(Y^*, A) - V(Y^*, A) = 0 \end{aligned} \tag{20}$$

Before proceeding note that we can sign the derivatives (21) independently of the specification of preferences using only the fact that education choice is modeled as an optimal stopping problem. To see this note that there are no "cross terms" (i.e. no YV_A or AV_Y terms) in (13). Therefore its solution will always be of the form $V(A,Y) = V^1(A) + B_1Y^{\theta_1} + \phi/\rho$ where $V^1(A)$ satisfies $\rho V^1(A) = u(c^*) + (rA - F - c^*)V_A^1$ and $u_c(c^*) = V_A^1$. Only the form of $V^1(A)$ will be affected by the particular parameterisation of preferences that we choose. Therefore we have:

$$\begin{array}{ll} V_{B_1} = Y^{\theta_1} > 0 & \Omega_{B_1} = 0 \\ V_{YB_1} = \theta_1 Y^{\theta_1 - 1} > 0 & \Omega_{YB_1} = 0 \\ V_F < 0 & V_{YF} = 0 \\ V_{YA} = 0 & V_{Y\theta_1} = B_1 Y^{\theta_1 - 1}(\ln\theta_1 + 1) \\ V_{\theta_1} = B_1 Y^{\theta_1}\ln\theta_1 & V_{YY} = B_1\theta_1(\theta_1 - 1)Y^{\theta_1 - 2} > 0 \end{array} \tag{21}$$

We can sign some more derivatives on the assumption that preferences are CRRA and $\gamma > 0$. Note that these derivatives would probably hold for any "well behaved" preferences i.e. $u_{cc}/u_c < 0$.

$$\begin{aligned} \Omega_{YY} &< 0 \\ \Omega_{YA} &< 0 \\ V_A &> \Omega_A > 0 \end{aligned}$$

The Jacobian of the system (20) is given by J. Its determinant, $|J| \neq 0$, so the implicit function theorem applies.

$$|J| = -V_{B_1}\left[\Omega_{YY} - V_{YY}\right] > 0$$

Define J_x to be the matrix

$$J_x = \begin{bmatrix} -\frac{\partial G_1}{\partial x} & -V_{YB_1}(Y^*, A) \\ \\ -\frac{\partial G_2}{\partial x} & -V_{B_1}(Y^*, A) \end{bmatrix}$$

where all derivatives are evaluated in the neighbourhood of the optimum. Using the implicit function theorem we can state the following derivatives hold in the neighbourhood of Y^*:

$$\frac{\partial Y^*}{\partial F} = \frac{|J_F|}{|J|} = \frac{V_{YB_1}V_F}{|J|} < 0$$

$$\frac{\partial Y^*}{\partial \theta_1} = \frac{|J_{\theta_1}|}{|J|} = \frac{-V_{B_1}V_{Y\theta_1} + V_{YB_1}V_{\theta_1}}{|J|} < 0$$

Because g and σ affect Y^* only through θ_1, itself determined implicitly by Q in (9), we have

$$\frac{\partial Y^*}{\partial g} = \frac{\partial Y^*}{\partial \theta_1}\frac{\partial \theta_1}{\partial g} > 0$$

$$\frac{\partial Y^*}{\partial \sigma} = \frac{\partial Y^*}{\partial \theta_1}\frac{\partial \theta_1}{\partial \sigma} > 0$$

Finally we note that $\frac{\partial Y^*}{\partial A}$ is ambiguous.

$$\frac{\partial Y^*}{\partial A} = \frac{|J_A|}{|J|} = \frac{V_{B_1}\Omega_{YA} - V_{YB_1}(\Omega_A - V_A)}{|J|}$$

$$= \frac{Y\Omega_{YA} - \theta_1(\Omega_A - V_A)}{Y\left[V_{YY} - \Omega_{YY}\right]}$$

The second term in the numerator reflects the easing of the budget constraint brought about by an increase in wealth. Because wealth is relatively scarce before graduation, we have $\Omega_A < V_A$. So relieving this scarcity will facilitate an increase in schooling. The first term is negative for CRRA and most well behaved preferences. This reflects the fact that increasing the level of financial wealth will reduce the marginal value of labour income after graduation, Ω_Y. This in turn will reduce the marginal value of the option to wait, V_Y, via the smooth pasting condition. Thus it becomes more attractive to leave school early.

When utility is CRRA the second term will dominate for most parameter values. The first term in the numerator is equal to

$$-Y\gamma B_0\left(A + \frac{Y^*}{r-\alpha}\right)^{-\gamma-1}$$

while the second term is equal to

$$-\theta_1 B_0 \left[\left(A + \frac{Y^*}{r-\alpha} \right)^{-\gamma} - \left(A - \frac{F}{r} \right)^{-\gamma} \right]$$

For most values of the parameters $A - \frac{F}{r}$ will be very small relative to $A + \frac{Y^*}{r-\alpha}$. So raising both to a negative power will make the whole of the second term a large positive number. The first term will be small in magnitude as it is raised to a lower power. Thus, overall the numerator will be positive generating the positive derivative illustrated in Figure 7

Finally note that the derivative $\partial Y^*/\partial A$ states the effect on Y^* of the state variable namely assets at graduation i.e. A_{S^*}. It would be more useful to know the effect of initial assets, A_0. In general we cannot derive an explicit relationship between A_0 and A_{S^*} because S^* is not know with certainty. However when preferences are CRRA then we can solve for c and substitute into the budget constraint (12) to give

$$A_t = \left[A_0 - \tfrac{F}{r}\right] e^{\frac{(r-\rho)}{\gamma})t} + \tfrac{F}{r} \quad \forall \quad t \in [0..S^*]$$

$$A_t = \left[A_S + \tfrac{Y^*}{r}\right] e^{\frac{(r-\rho)}{\gamma})t} - \tfrac{Y^*}{r} \quad \forall \quad t \in [S^*..\infty)$$

The implication is that asset balances at any point in time (including graduation) will be monotonic is initial balances so that $\partial Y^*/\partial A_0$ has the same sign as $\partial Y^*/\partial A_{S^*}$.

Furthermore, we can show that regardless of preferences, $A_0 = A_{S^*}$, if $\rho = r$. To see this, recall that we showed that preferences will only affect V through the function $V^1(A)$, which satisfies $\rho V^1(A) = u(c^*) + (rA - F - c^*)V_A^1$ and $u_c(c^*) = V_A^1$. Use the envelope theorem to differentiate the differential equation (13) with regard to A in the neighbourhood of optimal consumption, to get $\rho V_A^1 - rV_A^1 = V_{AA}^1(rA - F - c)$. If $\rho = r$ then it must be the case that $c = rA - F$. Substituting this back into the budget constraint (12) gives the result.

We extend the model to account for taxes, by specifying x to be net income and ε to be parameter that models the progressivity of the tax system, as in (19). The function $f(Y)$ summarizes the three types of taxes. In general ε equals the ratio of marginal to average tax rates. For lump-sum taxes $\varepsilon = 0$ and we interpret τ as the amount of the lump sum tax. For proportional taxes, $\varepsilon = 1$ and we interpret τ as the proportionate tax rate. For a progressive tax system we have $\varepsilon > 1$.

The system (20) can be re-written to account of the taxes as

$$\begin{aligned} G_1 &= \Omega_x(f(Y^*), A) f_Y(Y) - V_Y(Y^*, A) = 0 \\ G_2 &= \Omega(f(Y^*), A) - V(Y^*, A) = 0 \end{aligned}$$

As before the Jacobian of the system is non-zero, so the implicit function theorem applies.

$$|J| = -V_{B_1}\left[\Omega_{xx} f_Y f_Y + f_{YY}\Omega_x - V_{YY}\right]$$

It is clear that Ω_{xx} has the same sign as Ω_{YY}. Furthermore $f_{YY} \leq 0$ for $\varepsilon = 0$ and $\varepsilon \geq 1$. Thus $|J| > 0$ and the distinction between net and gross income will not affect the sign of any of the derivatives in Proposition 5 when taxes are lump sum, proportional or progressive. Only in the case of a particular choice of parameters and for some particular values of ε that must be between zero and one, will the the derivatives change sign.

In order to prove Proposition 6 we calculate

$$\begin{aligned} |J_\tau| &= +V_{B_1} f_\tau f_Y \Omega_{xx} - V_{YB_1}\Omega_x f_\tau > 0 \\ |J_\varepsilon| &= +V_{B_1}\left[f_\varepsilon f_Y \Omega_{xx} + f_{Y\varepsilon}\Omega_x\right] - V_{YB_1} f_\varepsilon \Omega_x \end{aligned}$$

For $\varepsilon = 0$ or $\varepsilon = 1$, we have

$$\frac{\partial Y^*}{\partial \tau} = \frac{|J_\tau|}{|J|} > 0$$

where τ can be interpreted as either a lump-sum or proportional tax rate, depending on ε.

For progressive taxes, the effect of changes in the degree of progressivity are more complex. For most values of the parameters $|J_\varepsilon| > 0$ and so

$$\frac{\partial Y^*}{\partial \varepsilon} = \frac{|J_\varepsilon|}{|J|} > 0$$

To see this rewrite $|J_\varepsilon|$ as $V_{B_1} f_\varepsilon f_Y \Omega_{xx} + \Omega_x\left[V_{B_1} f_{Y\varepsilon} - V_{YB_1} f_\varepsilon\right]$. The first term is positive, so a sufficient condition for $|J_\varepsilon| > 0$ is that term in square brackets is also positive. Evaluating this term explicitly gives $(\theta - \varepsilon)\ln Y^* > 1$ as a sufficient condition. Note that there is no simple condition sufficient to ensure that $|J_\varepsilon| < 0$, but $(\theta - \varepsilon)\ln Y^* < 1$ would be necessary for $|J_\varepsilon| > 0$.

B Diminishing Returns [Not For Publication]

Section 2.3 introduced assumed constant returns and also that school provides positive utility directly (in order to avoid a corner solution). In this appendix we allow for returns to diminish as schooling increases. We now specify the shadow to follow a mean reverting process

$$\frac{dY}{Y} = (\alpha_0 - \alpha_1 Y)ds + \sigma dz$$

This process is similar to the geometric Brownian motion (5) and we can apply similar techniques. Note that we have specified the return to education to be a diminishing function of the shadow wage and not a of elapsed schooling time. We do this for analytical convenience so as to avoid getting a partial deferential equation with time as a state variable.

As before, the Bellman equation (6) describes the evolution of the value of the option to leave school over the period $[0..S]$. Using Ito's lemma we evaluate the stochastic differential dV.

$$dV = \left\{ (\alpha_0 - \alpha_1 Y)V_Y Y + \frac{\sigma^2}{2} Y^2 V_{YY} \right\} dt + \{\sigma V_Y Y\} dz$$

Replacing dV in the Bellman equation, dividing across by dt and using $E[dz] = 0$, we get a second order ordinary differential equation similar to (7) with the exception that we have a slightly more complicated expression in place of g.

$$pV = \phi + (\alpha_0 - \alpha_1 Y)V_Y Y + \frac{\sigma^2}{2} Y^2 V_{YY}$$

It can be verified by substitution that (22) is a general solution to a differential equation of this form where $H(.)$ is the series representation of the confluent hypergeometric function[28] and θ_1 and θ_2 are the positive and negative roots, respectively, of $\frac{\sigma^2}{2}\theta(\theta - 1) + \alpha_0\theta - p = 0$.

$$V(Y) = B_1 Y^{\theta_1} H(Y; \theta_1) + B_2 Y^{\theta_2} H(Y; \theta_2) + \phi/p \tag{22}$$

$$\begin{aligned} H &= 1 + \frac{\theta}{b}x + \frac{\theta(\theta+1)}{b(b+1)}\frac{x^2}{2!} + \frac{\theta(\theta+1)(\theta+2)}{b(b+1)(b+2)}\frac{x^3}{3!}\ldots \\ x &\equiv \frac{2\alpha_1 Y}{\sigma^2} \\ b &\equiv 2\theta + \frac{2\alpha_0}{\sigma^2} \end{aligned}$$

[28] See Dixit and Pindyk (1994) page 163 and the references cited therein. Note that H reduces to the exponential function when $b = \theta$.

As before we can use the fact that $V(Y) \to \phi/p$ as $Y \to 0$ to set $B_2 = 0$. The value matching and smooth pasting conditions have the same form as (10) and define Y^* and B_1 implicitly. If we solve for Y^* we get (23) which itself must be solved numerically as both H and H_Y are infinite series.

$$(Y^* - \phi)\left[\frac{\theta}{Y^*} + \frac{H_Y}{H}\right] = 1 \tag{23}$$

Note that the solution to this model incorporates the solution to the simpler model of section 2.3 as a special case. If we eliminate the diminishing returns and set $\alpha_1 = 0$ then $H_Y = 0$ and (23) reduces to Y^* from Proposition 2.

Table 3 shows values of Y^* for certain sample values of α_0, α_1 and σ calculated by numerical simulation of (23). For these simulations we normalize $Y_0 = 1$ and set $\phi = 0$ as with diminishing returns it is no longer needed to avoid a corner solution. We also assume that $\rho = 0.1$. Examination of the table confirms that Y^* is increasing in α_0 and decreasing in α_1. As before, higher returns to education provide an incentive to stay in school. Now we have the additional factor that the return to education is lower at higher levels of education. This provides an incentive to leave education earlier.

We can also see from Table 3 that the threshold level of income is an increasing function of uncertainty. Greater risk will cause the individual to delay leaving school. Again this effect occurs even though the agent is risk neutral, and for the same reason as before – irreversibility in the presence of uncertainty provides an incentive to delay the decision.

References

[1] Altonji, Joseph (1993), "The Demand for and Return to Education when Education Outcomes are Uncertain", *Journal of Labor Economics,* vol. 11, pp. 48-83.

[2] Belzil, Christian and Hansen, Jorgen (2002), "Unobserved Ability and the Return to Schooling", Discussion paper no. 508, IZA-Bonn, forthcoming *Econometrica*

[3] Caballero, Ricardo J, and Engel, Eduardo (1999), "Explaining Investment Dynamics in US Manufacturing: A Generalized (S,s) Approach", *Econometrica*, vol. 67, pp. 783-826.

[4] Campbell, John Y. and Lo, Andrew W. and MacKinlay, A. Craig (1997), *The Econometrics of Financial Markets*, Princeton University, Princeton, New Jersey.

[5] Card, David (1999), "The Causal Effect of Education on Earnings", in: Orley Ashenfelter and David Card, eds., *Handbook of Labour Economics*, vol. 3, North-Holland, Amsterdam.

[6] Card, David (2001), "Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems", *Econometrica*, vol. 69, no. 5, pp. 1127-1160.

[7] Carneiro, Pedro and Heckman, James J. and Vytlacil, Edward (2001), "Estimating the Return to Education When it Varies Among Individuals", Mimeo, Department of Economics, University of Chicago.

[8] Chen, Stacey (2001), "Is Investing in College Education Risky?", mimeo, Department of Economics, SUNY – Albany.

[9] Conneely, Karen and Uusitalo, Roope (1999), "Estimating Heterogeneous Treatment Effects in the Becker Schooling Model", mimeo, Department of Economics, University of Helsinki.

[10] Denny, Kevin and Harmon, Colm (2001), "Testing for Sheepskin Effects in Earnings Equations: Evidence for five Countries", *Applied Economic Letters*, vol. 8, pp. 635-637.

[11] Dixit, Avinash K. (1993), *The Art of Smooth Pasting*, Harwood Academic Publishers, Switzerland.

[12] Dixit, Avinash K. and Pindyck, Robert S. (1994), *Investment Under Uncertainty*, Princeton University, Princeton, New Jersey.

[13] Dominitz, Jeff and Manski, Charles (1996), "Eliciting Student Expectations of the Returns to Schooling", *Journal of Human Resources*, vol. 31(1), pp. 1-26.

[14] Eaton, Jonathan and Rosen, Harvey (1980), "Taxation, Human Capital and Uncertainty", *American Economic Review* , vol. 70, pp. 705-715.

[15] Harmon, Colm and Hogan, Vincent and Walker, Ian (2001), "Dispersion in the Economic Returns to Schooling", Working Paper 16/01, Department of Economics, University College Dublin.

[16] Heckman, James J., Lochner, Lance and Taber, Christopher (1998), "Tax Policy and Human Capital formation", *American Economic Review (Papers and Proceedings)*, vol. 88, pp. 293-297

[17] Heckman, James J., Lochner, Lance and Todd, Petra E. (2001), "Fifty Years of Mincer Earnings Regressions." Mimeo, University of Chicago.

[18] Judd, Kenneth, (1998), "Taxes, Uncertainty and Human Capital", *American Economic Review (Papers and Proceedings)*, vol. 88, pp. 289-292.

[19] Kamien, Morton I. and Schwartz, Nancy L. (1991), *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management*, Elsevier Science, Amsterdam.

[20] Keane, Michael and Wolpin, Kenneth (1997), "The Career Decisions of Young Men", *Journal of Political Economy,* vol. 105, no. 3, pp. 473-522.

[21] Lin, Shuanglin (1998), "Labour Income Taxation and Human Capital Accumulation", *Journal of Public Economics,* vol. 68, pp. 291-302.

[22] Malliaris, A. G. and Brock, W.A. (1982), *Stochastic Methods in Economics and Finance*, Elsevier Science, Amsterdam.

[23] Merton, Robert C. (1971), "Optimal Consumption and Portfolio Rules in a Continuous-Time Model", *Journal of Economic Theory*, vol. 3, pp. 373-413.

[24] Mincer, Jacob (1974), *Schooling, Experience and Earnings*, Columbia University Press, New York.

[25] OECD (2001), *Education at a Glance*, OECD, Paris.

[26] Trostel, Philip (1993), "The Effect of Taxation on Human Capital", *Journal of Political Economy,* vol. 101, pp. 327-350.

[27] Williams, Joseph (1979), "Uncertainty and the Accumulation of Human Capital over the Life Cycle", *Journal of Business*, vol. 52, no. 4, pp. 521-548.

Table 1: Key Parameters for Simulation

Parameter	Definition	Value
γ	CRRA	1.0
Y_0	Wage with min. Schooling	1.0
ρ	Discount rate	0.1
r	Return on Financial Assets	0.1
ϕ	Intrinsic utility from Education	1.659
A	Financial Assets: "Rich"	2.5
	Financial Assets: "Poor"	1.5
F	Fees	0.1
g	Mean Return to Education	0.07
σ	Stn. Dev. of Return to Education	0.02

Table 2: Optimal School Leaving

Education Risk σ	Threshold Income Y^*	Time in School (S^*) $E(S^*)$	$Stn(S^*)$	$Skew$
0	1.1506526	2.0	-	-
0.01	1.1509341	2.02	0.20	0.30
0.02	1.1517757	2.07	0.42	0.60
0.03	1.1531684	2.16	0.66	0.90
0.04	1.1550980	2.28	0.93	1.15
0.05	1.1575448	2.42	1.20	1.21
0.06	1.1604849	2.57	1.44	1.13
0.07	1.1638910	2.68	1.63	1.03
0.08	1.1677324	2.76	1.76	0.94
0.09	1.1719768	2.80	1.86	0.88

1. Based on 10,000 draws from distribution of S^*
2. Key parameters: $Y_0 = 1$; $\rho = r = 0.1$; $g = 0.07$;
$\phi = 1.659$; $A = 2.5 * F/r$

Table 3: Threshold Income with Diminishing Returns to Education

α_0	0.15	0.15	0.2	0.2
α_1	0.025	0.05	0.02	0.01
Risk (σ)				
0.00	2.00	1.00	5.00	10.00
0.02	2.01	1.01	5.01	10.02
0.04	2.02	1.01	5.04	10.08
0.06	2.04	1.02	5.09	10.18
0.08	2.07	1.04	5.16	10.31
0.10	2.1	1.05	5.24	10.47
0.12	2.14	1.07	5.33	10.65
0.14	2.18	1.09	5.43	10.85
0.16	2.23	1.12	5.54	11.08
0.18	2.28	1.14	5.66	11.32
0.20	2.33	1.17	5.79	11.57

1. Simulation of basic model as in equation (23)
2. Key parameters: $Y_0 = 1$; $\rho = 0.1$; $\phi = 0$
see text for discussion

Figure 1: The Threshold Shadow Wage

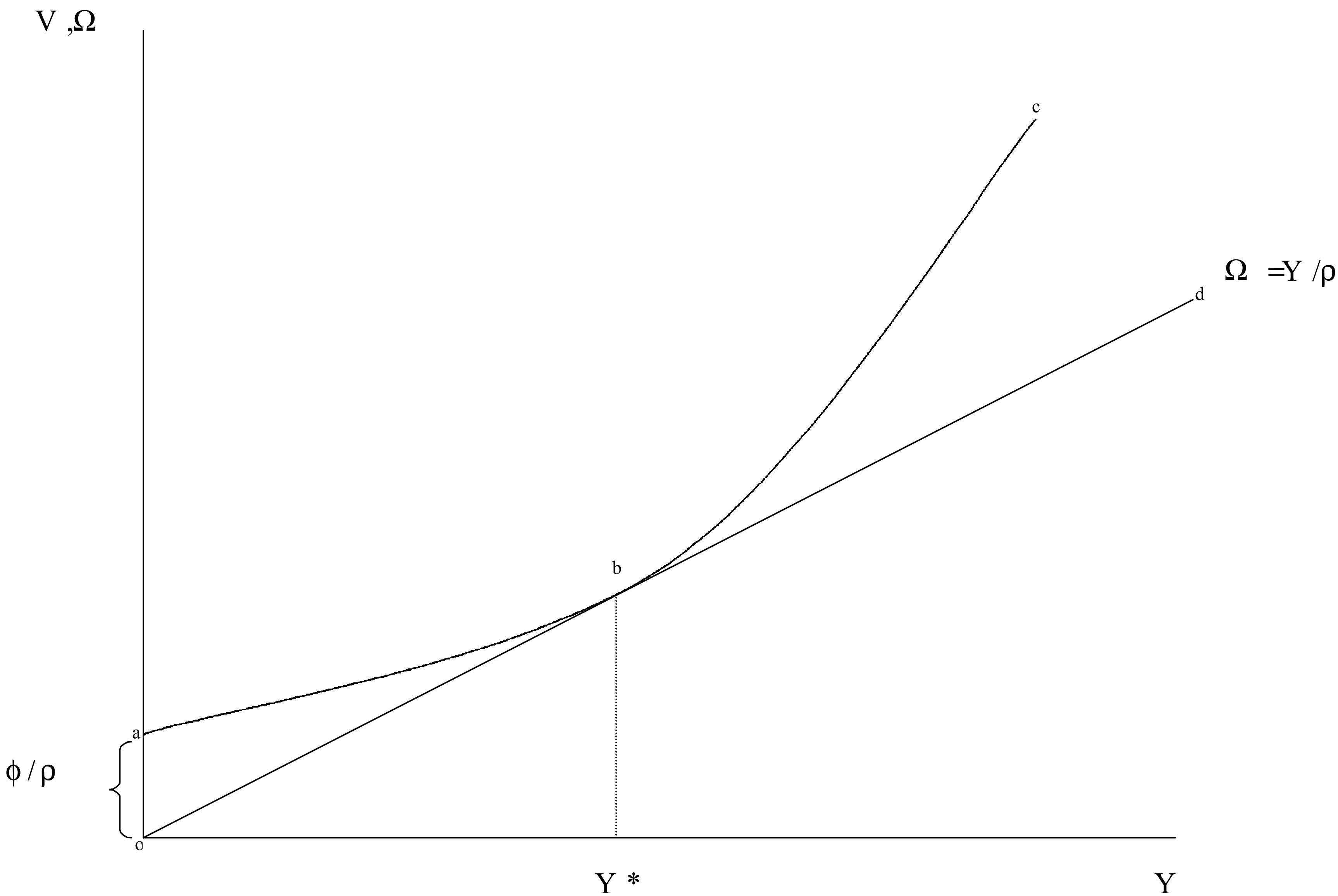

Figure 2: Value function

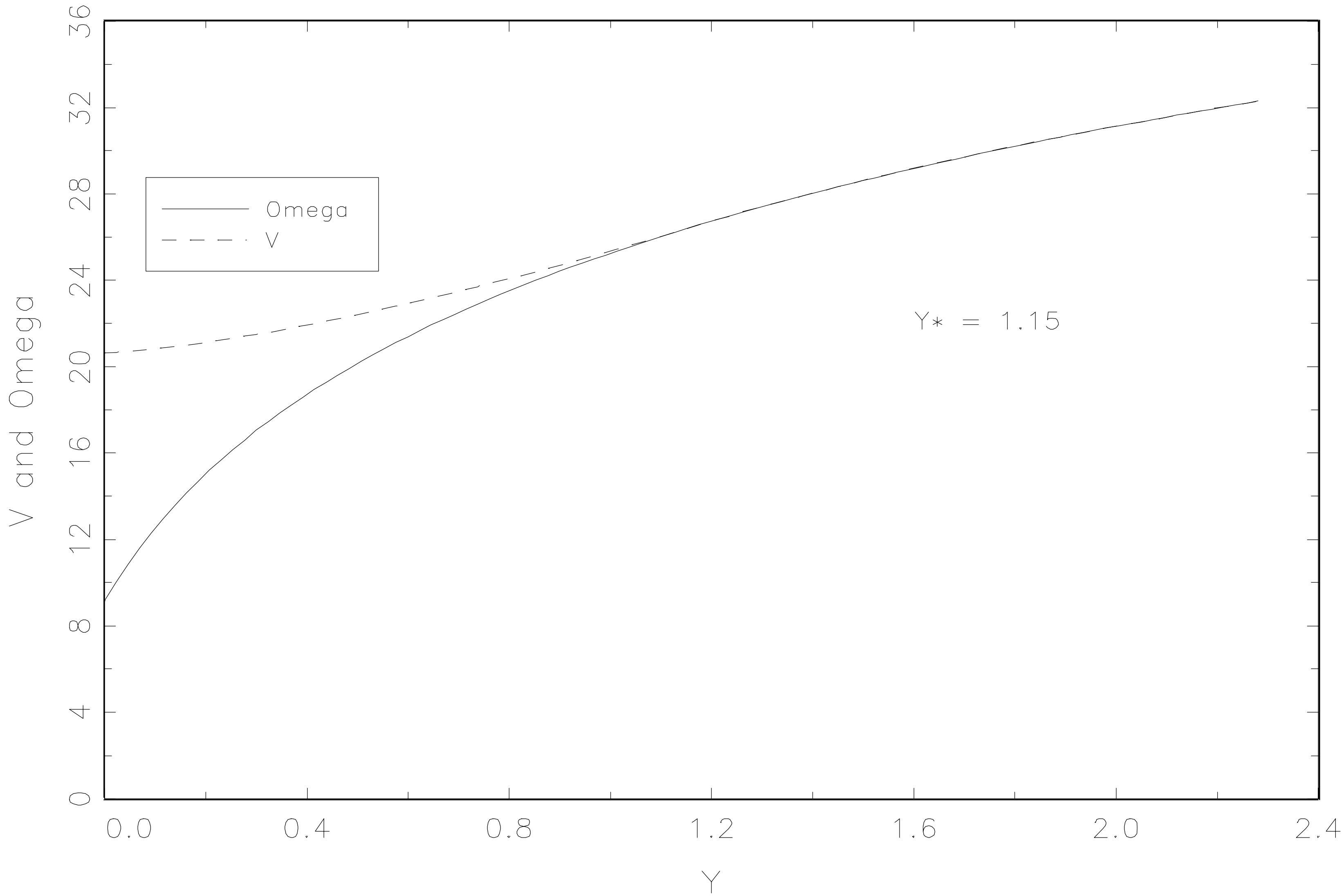

Figure 3: Consumption Path

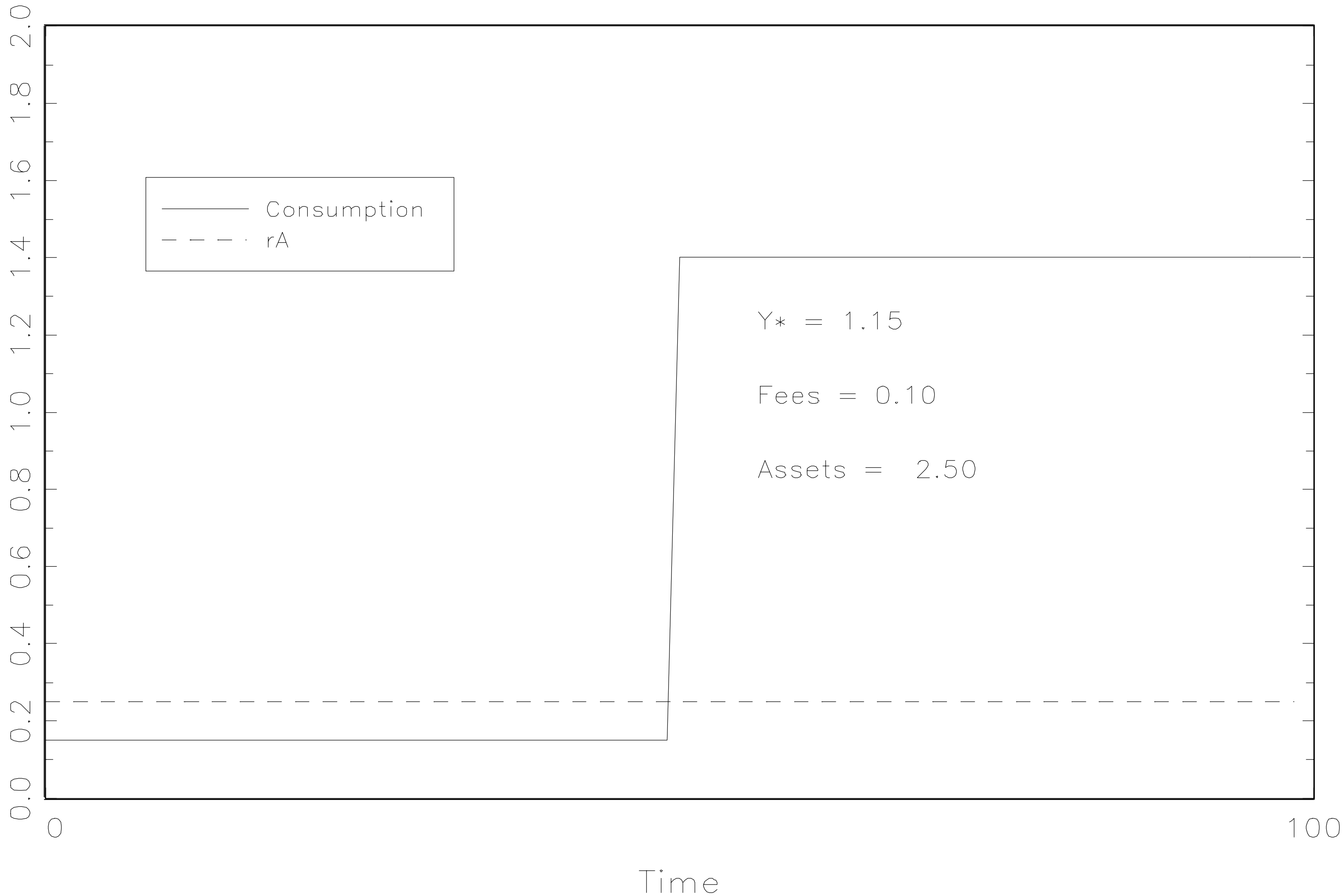

Figure 4: The Effect of Expected Returns (g) on Y*

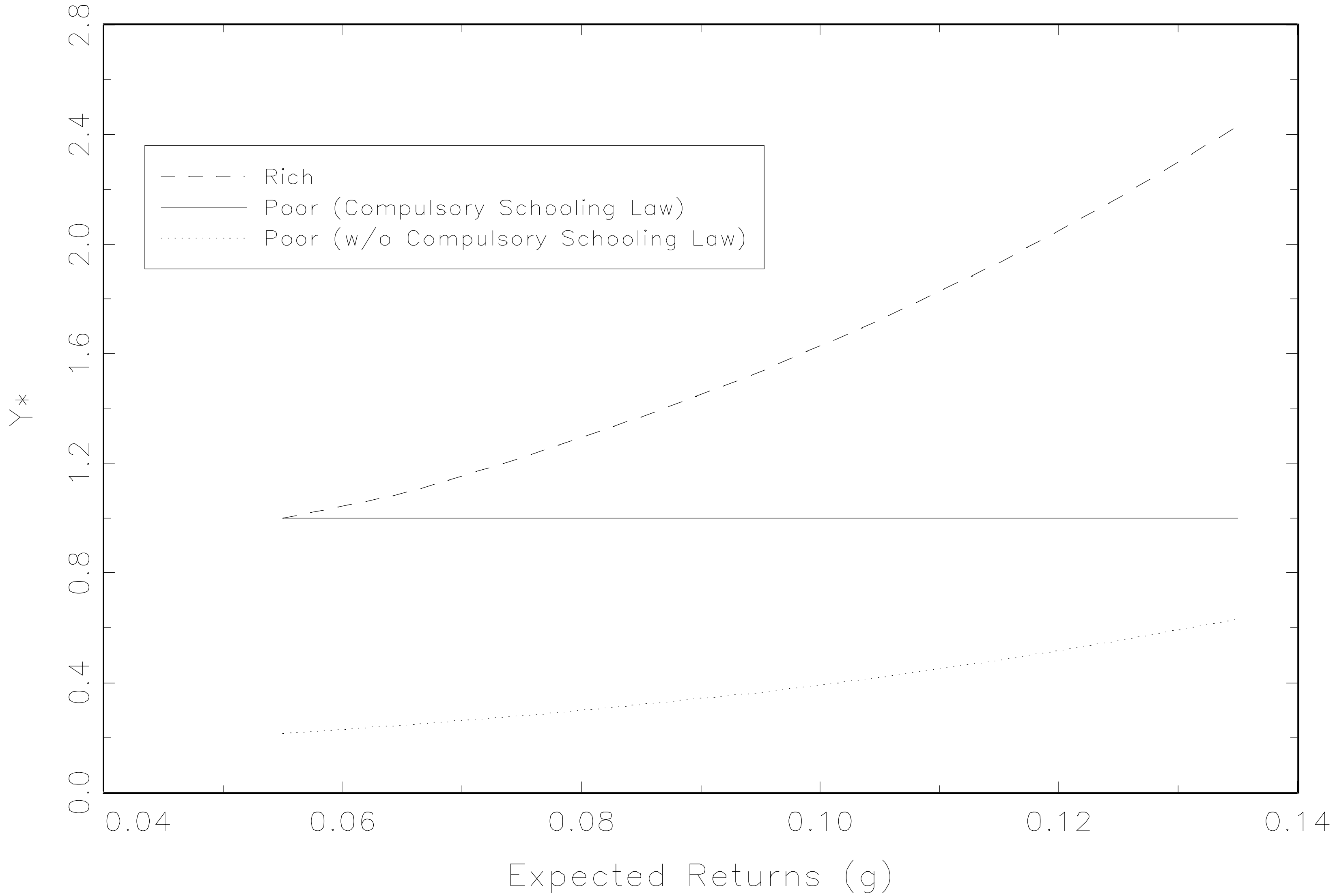

Figure 5a: The Effect of Risk (Sigma) on Y*

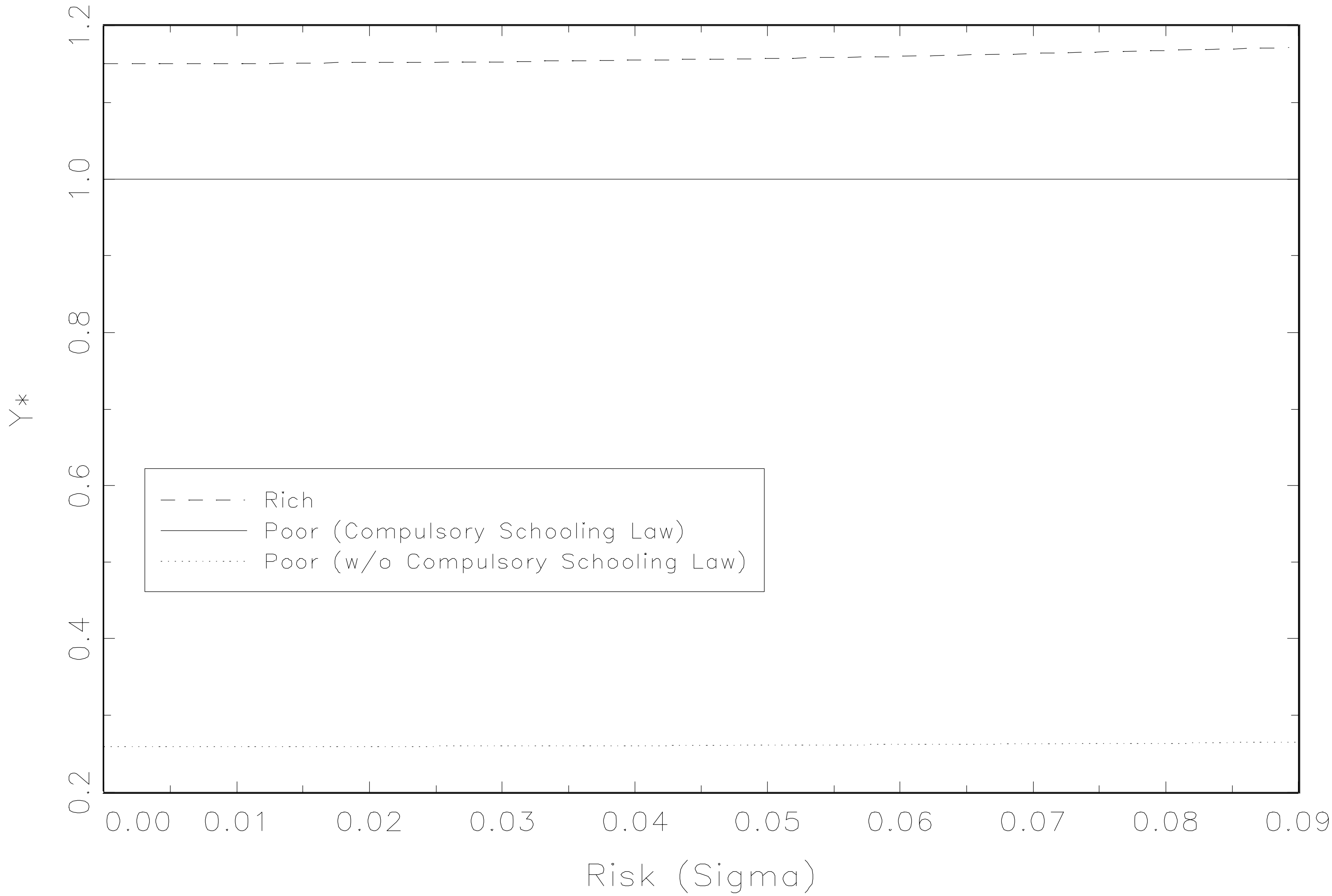

Figure 5b: The Effect of Risk on Y* for Rich

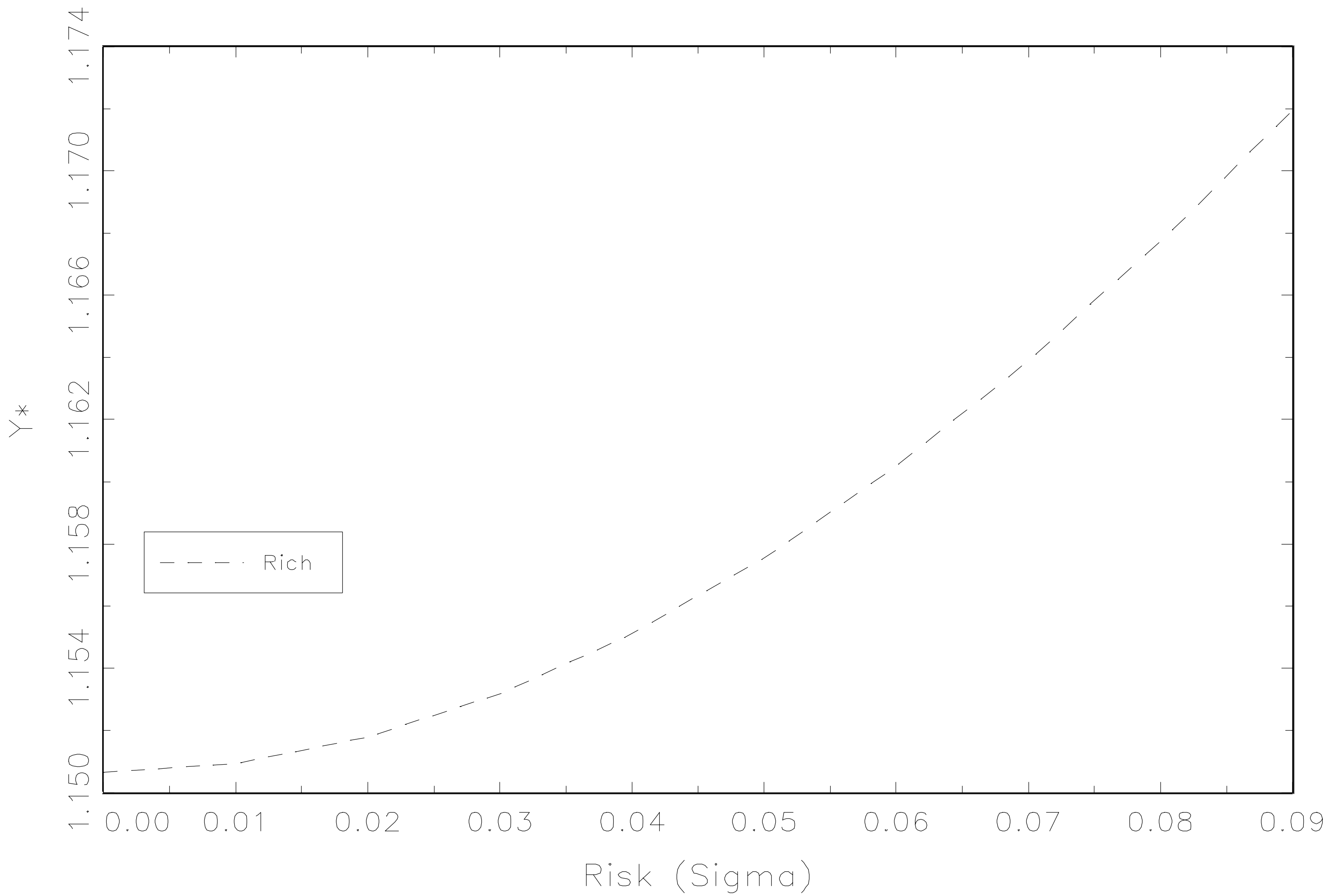

Figure 6: Density of S* for Risk Levels

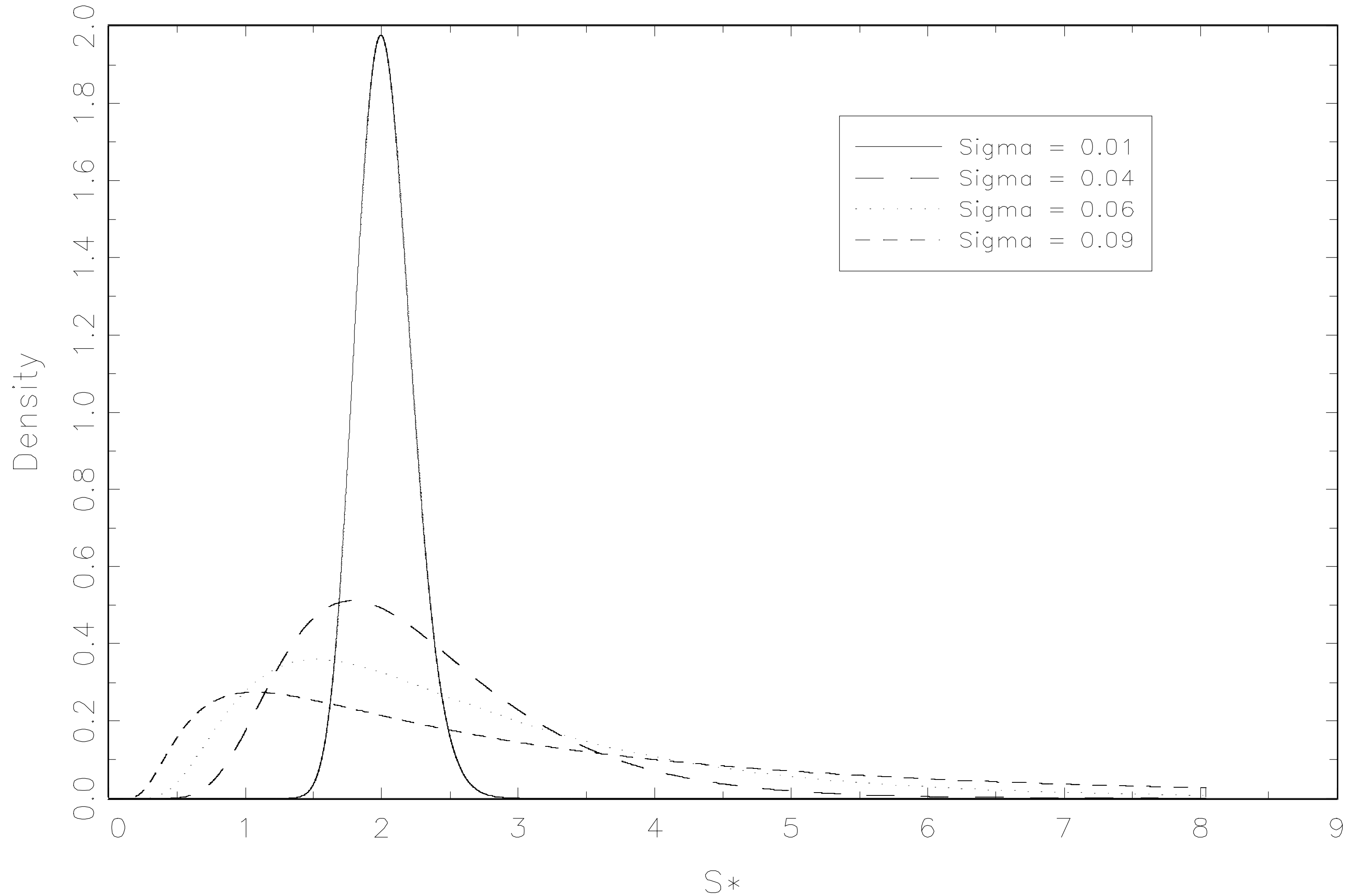

Figure 7: The Effect of rA/F on Y*

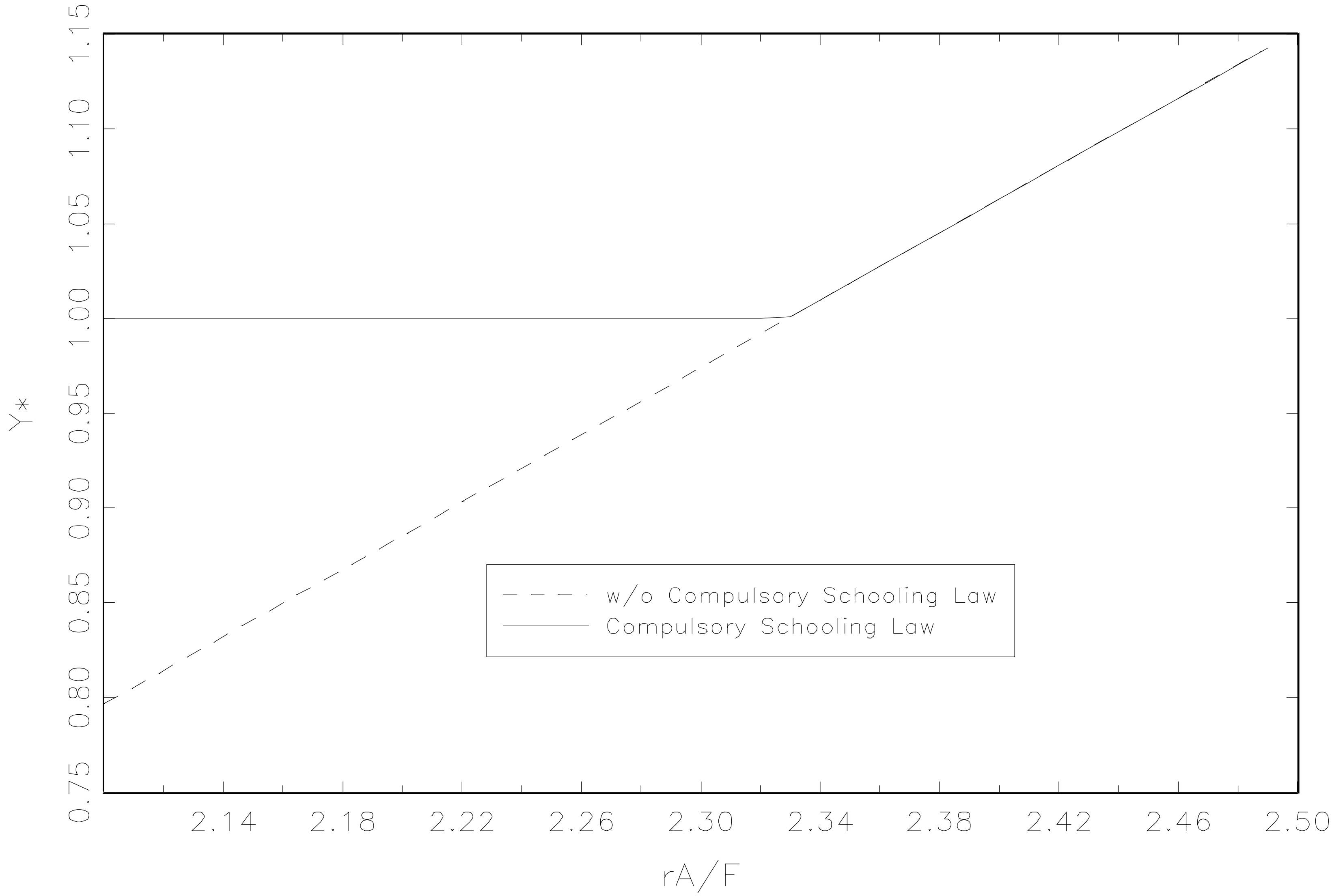

Figure 8: A Proportional Tax

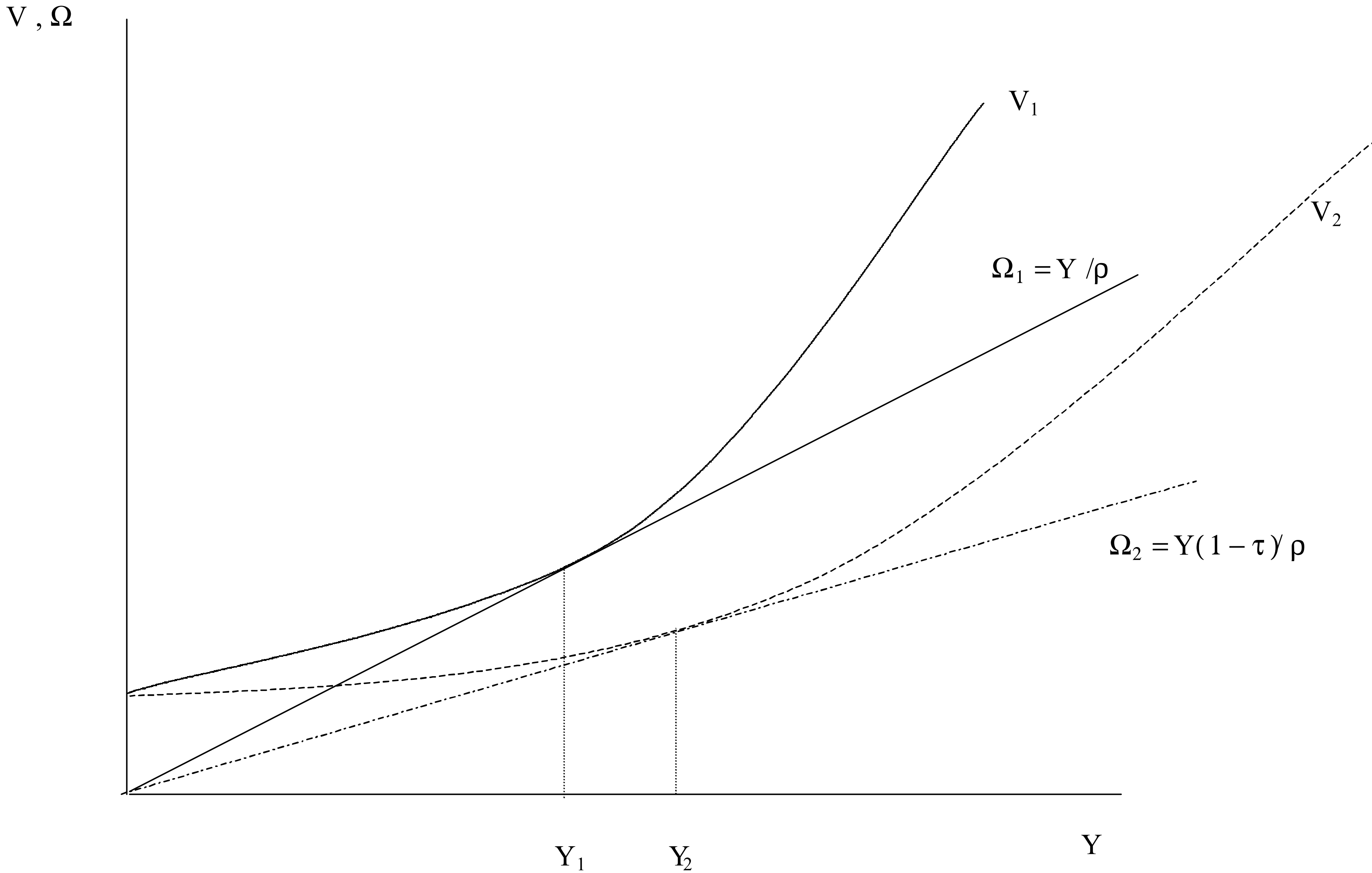

Figure 9: The Effect of Tax Progressivity on Y*

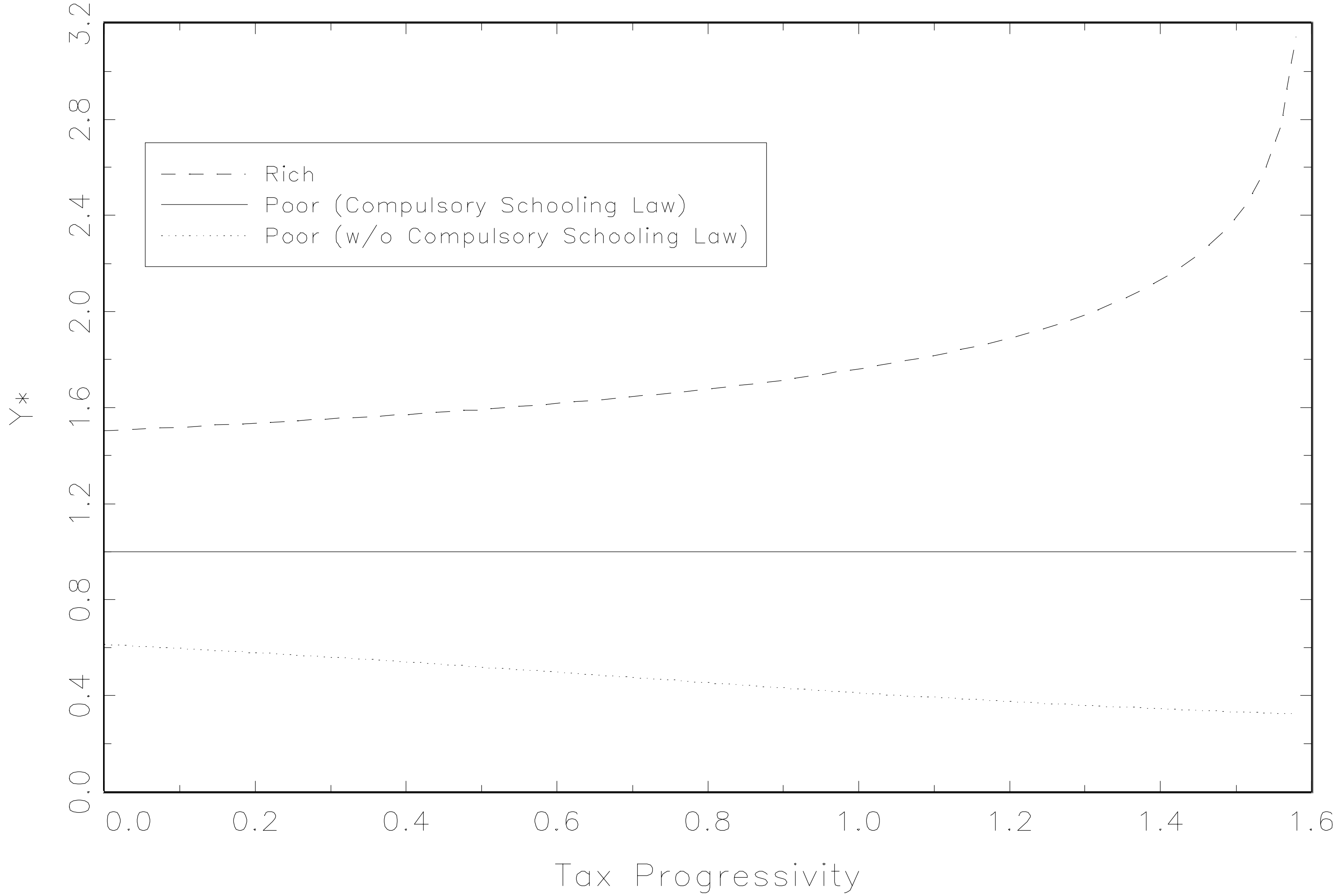

Figure 10: The Effect of a Repayable Subsidy on Y*