The risk of investment in human capital
Raita, S.M.

Citation for published version (APA):
Chapter 6

Education as a Real Option

6.1. Introduction

In this chapter we treat education as a sequential choice that is made under uncertainty. Our model acknowledges the option of leaving now rather than waiting and keeping open the possibility of not leaving should the market wage fall. Uncertainty and irreversibility imply an option value of waiting and therefore greater hesitancy in individuals' going out on the labor market. We also investigate which impact the probability of major shocks has on the decision to further invest in schooling: the individual may discover a sudden drop in the benefits from attending school.

It has been long recognized that individual demand for education should be analyzed in an investment framework. However, most of the enormous empirical literature on human capital and earnings that has been developed from Mincer (1974), Becker (1964, 1993) and other human capital pioneers abstracts from uncertainty about whether a program will be completed and on the role played by risk in education decisions.

While the literature on risk is not abundant some authors have analyzed the topic. Williams (1979) uses a model in which the production of human capital, its depreciation and future outcomes are all stochastic. Higher risk (larger variance in the production of human capital from given inputs) reduces the investment in schooling. Groot and Oosterbeek (1992) define an increase in risk as a mean preserving increase in the spread in the distribution of returns, for any arbitrary distribution. They find that greater dispersion in earnings by schooling level reduces the returns to schooling.
Most investment decisions have three features with different weights: they are partially or completely irreversible, they have uncertain outcomes and their future rewards depend on the timing of the investment. Using the theory of real options, Dixit and Pindyck (1994) focus on the interaction of these three characteristics, in order to determine the optimal decisions of the investors. The option approach is relatively new and it is in essence the view that an individual, while in school, has the option of leaving it whenever he prefers and earn a wage that is related to the time spent in school. The decision to leave the school is irreversible. We base our assumption on the fact that it is uncommon for individuals to go back to school after they have decided to end it. This approach recognizes the option value of waiting for better (but never complete) information.

In Hogan and Walker (2001), a stochastic dynamic programming model is constructed, where being in school has utility value and the shadow wage follows a Brownian motion. Once the student leaves school, this shadow wage becomes the fixed wage for the entire working life. Increasing risk in the post-school wage implies an increase in the upside risk, the probability to obtain a high-wage, while the increase in down-side risk remains ineffective, because at low wage students stay in school anyway. As a result, individuals react by staying in school longer as the risk increases. As individuals can permanently monitor the wage they may get when quitting school, one may be tempted to see this wage as the given reward for the stochastic production of skills while in school. If one then also assumes that less able individuals face greater risk in their production of skills, the model implies that less able individuals will stay in school longer, which is certainly at variance with reality. We will start out from the same model, but adjust for its drawbacks. We will find that increasing risk may not only increase but also decrease expected schooling length.

The chapter is organized as follows: section 6.2. discusses the theoretical model. Section 6.3 solves the model. Section 6.4. extends the model. Section 6.5. summarizes and draws the conclusions.

6.2. Model

We develop a particular model, based on Hogan and Walker (2001), but replace the assumption of fixed post-school wage by a stochastic wage. In our model we assume that when the income reaches a certain level, the threshold wage, the option to leave school is exercised and the individual leaves school and earns an income that evolves, from that initial
level, in a manner dictated by the assumed underlying income process. We assume the return to schooling as being drawn from a normal distribution. Hence, the return from $s$ years of schooling, $r_s$, will be distributed as a normal random variable with mean $\alpha$ and standard deviation $\sigma$ when one additional year of education ($s = 1$) is taken. In other words for small additional intervals, $ds$, $r_s = \left(\frac{dY}{Y}\right)$ will be normally distributed $\mathcal{N}(\alpha ds, \sigma^2 ds)$. Thus, we assume that $Y$ follows a geometric Brownian motion¹, that is if the income at time $t$ is $Y_t$, then the variation in $Y_t$ over a small interval $dt$ is given by:

$$dY_t = \alpha Y_t dt + \sigma Y_t dz$$

(6.1)

We denote $dz$ a Wiener process².

We assume that the decision process is made sequentially: each period the individual has two alternatives: either, he can leave school, lose the value of the option to leave school, enter the labor market and earn a stochastically determined wage that is a function of the accumulated schooling, or he can stay in school for a while, keep the option of leaving school, while receiving some utility from school, $u$, and in the next period he can take again the gamble of exercising the option. If he chooses not to exercise the option, its value evolves according to the underlying income process. Once the option is exercised the individual cannot return to school. The individuals are infinitely lived.

The model of Walker and Hogan has the restriction of ignoring all risk once individuals left school, since they earn a fixed wage for the entire working life. However, we assume that earnings are not fixed but continue to fluctuate with Brownian motion while individual is on the labor market.

As the income flow $Y_t$ follows a Brownian motion in perpetuity, if the initial state is $Y$ and the future earnings are discounted³ at the discount rate $i$, then the expected present value of all future revenues⁴ is given by:

---

¹ See Dixit and Pindyck (1994), page 65.
² A formal description of the Wiener process is presented in the Appendix 6.1.
³ The risk in the wages, $Y$, has zero correlation with the overall labor market risk. Thus the riskless rate is the discount rate for all future revenues, certain or uncertain.
⁴ The proof of this formula is presented in Appendix 6.4.
Chapter 6. Education as a Real Option

\[ E \int_0^\infty Y e^{-\mu t} dt = \frac{Y}{i - \alpha}. \]

Note that the present value is a linear function of the initial value of the wages. We will call it \( f(Y) \). For this to make economic sense we must assume that \( i > \alpha \). If \( i < \alpha \), waiting longer would always be a better policy and an optimum would not exist. We will let \( \delta \) denote the difference between \( i \) and \( \alpha \), \( \delta = i - \alpha \). Thus we are assuming \( \delta > 0 \).

Let \( V(Y) \) be the value of the option to leave school at time \( t \) and start earning income, the current state of the income stream being \( Y_t \). The option value changes over time according to a Bellman equation. The value of the option of leaving school at time \( t \) can be expressed as the sum of the operating profit over the interval \( (t, t+dt) \) and its continuation value beyond \( t+dt \). Since the individual receives net utility \( u \) from participating in education, then:

\[ V(Y) = u dt + E(V(Y + dY) e^{-i dt}). \]

Expanding the right hand side using Ito's Lemma\(^5\), we obtain:

\[ V(Y) = u dt + \left[ \alpha Y V'(Y) + \frac{1}{2} \sigma^2 Y^2 V''(Y) + (1 - i dt) V(Y) + o(dt) \right] \]

where \( o(dt) \) collects terms that go to 0 faster than \( dt \), and \( V' \) and \( V'' \) represent first and second derivative of \( V \) with respect to \( Y \). Simplifying, dividing by \( dt \) and proceeding to the limit as \( dt \rightarrow 0 \) we get the differential equation\(^6\):

\[ \frac{1}{2} \sigma^2 Y^2 V''(Y) + \alpha Y V'(Y) - i V(Y) + u = 0 \]

The general solution for this equation has a homogeneous part: a linear combination of the two power solutions corresponding to the roots \( \beta_1, \beta_2 \) of the quadratic equation:

\(^5\) The option value \( V \), being a constant multiple of \( Y \), also follows a geometric Brownian motion with the same parameters \( \alpha \) and \( \sigma \). A proof can be found in Dixit and Pindyck (1994), page 79. Ito's Lemma is described in the Appendix 6.2.

\(^6\) A generation of this formula is presented in the Appendix 6.5.
Chapter 6. Education as a Real Option

\[ Q_1 : \quad \frac{1}{2} \sigma^2 \beta (\beta - 1) + \alpha \beta - i = 0 \]

and a constant \( \frac{u}{i} \). Therefore we can write \( V(Y) \) as:

\[ V(Y) = B_1 Y^{\beta_1} + B_2 Y^{\beta_2} + \frac{u}{i} \]

where the constants \( B_1, B_2 \) are to be determined.

If the shadow wage tends to 0 the individual will never leave school, so the value of being in school will be the present value of the direct utility of perpetual education:

\[ \lim_{Y \to 0} V(Y) = \frac{u}{i} \]

We focus now on the quadratic equation \( Q_1 \). Since \( i > \alpha \), the larger root \( \beta_1 \) of this equation exceeds 1 and the other root, \( \beta_2 \), is negative. Since \( \beta_2 < 0 \), that power of \( Y \) in \( V(Y) \) goes to \( \infty \) as \( Y \) goes to 0. To prevent the value from diverging, we must set the corresponding coefficient\(^7\), \( B_2 = 0 \). The other term \( B_1 Y^{\beta_1} \) is not easy to get rid of. It represents a component of \( Y \) attributable to speculative bubbles as \( Y \to 0 \).

At a threshold value of the wage, \( Y^0 \), it becomes optimal to exercise the option of leaving school. Once the individual leaves school, the shadow wage becomes his real wage and it continues its Brownian motion (with \( Y^0 \) as initial value) during the entire life. This is where we deviate from the Hogan and Walker model, which assumes a fixed wage after leaving school. In order for the threshold to be chosen optimally, the value of the option at the threshold \( V(Y^0) \), must equal the present value of lifetime income \( f(Y^0) \). This is the value-matching condition:

\[ V(Y^0) = \frac{Y^0}{i - \alpha} . \]

\(^7\) The mechanism is similar to a bubble in the price of a financial asset.
Furthermore, the graphs of the option value and of the present value function should meet tangentially at the optimum. This is the smooth-pasting condition:

\[ V'(Y^0) = \frac{1}{i - \alpha}. \]

Replacing \( V(Y) = B_i Y^R + \frac{u}{i} \) in the matching and smooth pasting conditions yields the value of the threshold wage:

\[ Y^0 = \frac{\beta_1}{\beta_1 - 1} \delta \frac{u}{i} \]  
(6.2)

where

\[ \beta_1 = \frac{\left( \frac{1}{2} \sigma^2 - \alpha \right) + \sqrt{\left( \frac{1}{2} \sigma^2 - \alpha \right)^2 + 2 \sigma^2 i}}{\sigma^2} \]

Sufficient conditions such that \( Y^0 > 0 \) are \( u > 0 \) and \( i > \alpha \). Otherwise \((u < 0 \) and \( i < \alpha)\) school would provide a better return and it would be optimal to stay in school forever.

To see how the decision to stop schooling is influenced by the variance in earnings we differentiate for \( Y^0 \) totally, with respect to \( \sigma^2 \). Hence:

\[
\frac{\partial Y^0}{\partial \sigma^2} = \frac{\partial \beta_1}{\partial \sigma^2} \left( \frac{\beta_1 - 1}{(\beta_1 - 1)^2} \right) - \frac{\partial (\beta_1 - 1)}{\partial \sigma^2} \frac{\beta_1}{i} \delta \frac{u}{i} \]

\[
\frac{\partial Q_1}{\partial \sigma^2} = \frac{1}{2} \beta_1 (\beta_1 - 1) > 0 \quad \text{Implicit Function Theorem:} \quad \frac{\partial Q_1}{\partial \sigma^2} = -\frac{\partial Q_1}{\partial \beta_1} < 0 \]

Therefore, as \( \sigma^2 \) increases, \( \beta_1 \) decreases and therefore the ratio \( \frac{\beta_1}{\beta_1 - 1} \) increases. The greater the level of uncertainty over future values of the earnings, the larger is the excess income the individual will demand before he is willing to make the irreversible choice of leaving school. Hence, allowing for stochastic income after school, rather than assuming
fixed income, is immaterial for the key conclusion obtained by Hogan and Walker that increasing risk extends schooling length.

Note that the basic model of this chapter can be extended and generalized in several ways. One is to allow a more general wage process, for example, a mean-reverting process, or an even more general Itô process. The only difference this makes is that in the differential equation for the option value and in that of the option, the coefficients become more complicated functions of $Y$. In almost all such cases we must rely on numerical solutions. Since there are no new general economic insights to be had from such calculations we will not develop this line of models here. A different direction of extension is worth some attention.

6.2. Allowing for Risk Aversion

In the model described above, individuals are earnings maximizers and hence, risk neutral. We relax the assumption and specify a model with risk aversion where individuals evaluate future income by adding the variance of earnings to their expected income. The variance is weighted by the individuals’ risk attitude parameter, i.e. the certainty equivalent based on the individual’s utility function. The solution to the dynamic programme depends crucially on how we specify the budget constraint and in this section we generalize it by assuming that while in school the individual gets the same utility $\nu$, but after school the utility is provided by labor income via $U$, the instantaneous utility function. The utility function $U$ is increasing and concave. The form of the utility function only affects utility after leaving school as we have precluded the possibility that an agent may borrow against future income in order to subsidize consumption before leaving school. As before, the wage stream $Y_t$ follows a geometric Brownian motion with parameters $\alpha$ and $\sigma$, and we assume a constant income elasticity of the utility function, specifically with Constant Relative Risk Aversion (CRRA):

$$U(Y) = \frac{1}{1 - \rho} Y^{1 - \rho}$$

where $\rho = \frac{U'(Y)}{U(Y)} Y > 0$ is the coefficient of risk aversion.$^8$

---

$^8$ This concept was introduced for the first time by Arrow (1965).
Chapter 6. Education as a Real Option

Following the same steps as before, let $Y'$ denote the threshold wage and $V(Y)$ be the value of the option to leave school. While unexercised, $Y$ in the range $(0, Y')$, the option has merely an expected gain: $udt + E[dV(Y)]$. Setting this equal to the normal return from keeping the option alive for $dt$: $iV(Y)dt$ we get the familiar solution:

$$V(Y) = A_Y Y^\beta + \frac{\mu}{i}$$

where $A_Y$ is a constant to be determined, and $\beta_Y$ is the positive root of the quadratic:

$$Q_2: \quad \frac{1}{2} \sigma^2 \beta (\beta - 1) + \alpha \beta - i = 0$$

Taking the limit of $Y$ to 0 we rule out the term with the negative root. The value of the option to leave school is worth, just at the time of leaving the school, the present value of the discounted sum of utilities from income distributions after graduation. We can calculate this by direct integration assuming as before that income follows the Brownian motion of equation (6.1.). Hence the value matching condition should satisfy:

$$V(Y^*) = E\left[ \int_0^\infty U(Y_t)e^{-\mu t} dt \right]$$

where the initial level of the real wages, $Y_i$, is $Y^*$.

From the smooth pasting condition, $Y^*$ is at the point of tangency of the two curves:

$$\left. \frac{\partial V(Y)}{\partial Y} \right|_{Y = Y^*} = \left. \frac{\partial E\left[ \int_0^\infty U(Y_t)e^{-\mu t} dt \right]}{\partial Y} \right|_{Y = Y^*}$$

To solve for $E\left[ \int_0^\infty U(Y_t)e^{-\mu t} dt \right]$ we retain the first terms in Taylor series expansion of the utility function $U$ around the expected income at time $t$. We have that:
Chapter 6. Education as a Real Option

\[ U(Y_i) = U(Y_i) + (Y_i - Y_i)U'(Y_i) + \frac{1}{2}(Y_i - Y_i)^2 U''(Y_i) + o(dt) \]

where \( U' \) and \( U'' \) are the first and the second derivatives \( U \) with respect to \( Y \).

Hence, multiplying by \( e^{it} \), integrating and taking expectations on both sides:

\[
E \int_0^\infty U(Y_i)e^{-it}dt = \int_0^\infty U(Y_i)e^{-it}dt + \frac{1}{2} \int_0^\infty Var(Y_i)U'(Y_i)e^{-it}dt
\]

(6.3)

where \( V(Y_i) \) is the variance of \( Y_i \).

If the starting value of \( Y_i \) is \( Y \), then the expected value and the variance of \( Y_i \) are respectively:

\[ \bar{Y}_i = EY_i = Ye^{\alpha} \]
\[ Var(Y_i) = Y^2 e^{2\alpha} (e^{\sigma^2} - 1) \]

The proofs of the formulas can be found in the Appendix 6.3.

Working out equation (6.3.), under the assumption that the starting value of \( Y_i \) is \( Y' \), we obtain:

\[
E \int_0^\infty U(Y_i)e^{-it}dt = \int_0^\infty (Y^* e^{\alpha})^{i-\rho} e^{-it}dt + \frac{1}{2} \int_0^\infty (Y^*)^2 e^{2\alpha} (e^{\sigma^2} - 1)(-\rho)(Y^* e^{\alpha})^{-\rho-1} e^{-it}dt
\]

which can be further rewritten as:

\[
E \int_0^\infty U(Y_i)e^{-it}dt = \frac{(Y^*)^{i-\rho}}{1-\rho} \int_0^\infty e^{i\sigma(1-i\rho)} e^{-it}\frac{1}{2}\left(Y^*\right)^{i-\rho-1}(-\rho) \int_0^\infty e^{-\left(i\sigma^2\alpha(\rho+1)-2\alpha\rho\right)\rho} - e^{-\left(i\sigma(\rho+1)-\alpha\rho\right)\rho} dt
\]

Let us denote \( E \int_0^\infty U(Y_i)e^{-it}dt \) with \( g(Y^*) \). Thus, if \( i > \max\left\{\alpha(1-\rho), \sigma^2 - \alpha\rho + \alpha, -\alpha\rho + \alpha\right\} \) then:

\[
g(Y^*) = \frac{(Y^*)^{i-\rho}}{1-\rho} \frac{1}{i-\alpha(1-\rho)} - \frac{1}{2} (Y^*)^{i-\rho} \rho \left[ \frac{1}{i+\alpha(\rho+1)-\sigma^2} - \frac{1}{i+\alpha(\rho+1)} \right] = (Y^*)^{i-\rho} A\]
where $A$ is a constant given by:

$$A = \frac{1}{1 - \rho} \left[ -\frac{1}{i - \alpha(1 - \rho)} \right] - \frac{1}{2} \rho \left[ -\frac{1}{i + \alpha(\rho - 1) - \sigma^2} \right]$$

The threshold $Y^*$ is jointly determined by solving the system:

$$\begin{cases}
A_i(Y^*)^\beta_i + \frac{\mu}{i} = (Y^*)^{1 - \beta_i} A_i \quad \text{"the value-matching condition"} \\
A_i\beta_i(Y^*)^{\beta_i - 1} = (1 - \rho)(Y^*)^{-\rho} A_i \quad \text{"the smooth-pasting condition"}
\end{cases}$$

Solving the system yields the solution:

$$Y^* = \left( \frac{\beta_i}{A[\beta_i - (1 - \rho)]} \right)^{\frac{1}{1 - \rho}} \left( \frac{u}{i} \right)^{\frac{1}{1 - \rho}} \quad (6.4)$$

For the threshold to be positive, $Y^* > 0$, we need $A[\beta_i - (1 - \rho)] > 0$. Because $\beta_i > 1$ and $\rho > 0$, we have that $[\beta_i - (1 - \rho)] > 0$. Therefore a sufficient condition for a positive threshold is that the constant $A > 0$.

The effect of the variance on the threshold level is obtained from totally differentiating $Y^*$ with respect to $\sigma^2$:

$$\frac{\partial Y^*}{\partial \sigma^2} = \frac{1}{1 - \rho} \left[ \frac{\beta_i}{A[\beta_i - (1 - \rho)]} \right]^{\frac{1}{1 - \rho}} \left( \frac{u}{i} \right)^{\frac{1}{1 - \rho}} \frac{\partial A[\beta_i - (1 - \rho)]}{\partial \sigma^2} \frac{\partial \beta_i}{\partial \sigma^2} \frac{\beta_i}{A^2[\beta_i - (1 - \rho)]^2}$$

Rearranging, the denominator of the last ratio in the right hand side will be:

$$A(\rho - 1) \frac{\partial \beta_i}{\partial \sigma^2} \beta_i (\beta_i - 1 + \rho) \frac{\partial A}{\partial \sigma^2}$$

Applying the Implicit Function Theorem to the quadratic $Q_2$, as before, we get that $\frac{\partial \beta_i}{\partial \sigma^2} < 0$.

Differentiating $A$ with respect to $\sigma^2$ we obtain that:
\[
\frac{\delta A}{\delta \sigma^2} = -\frac{1}{2} \rho \frac{1}{(\beta_1 \sigma^2 - \alpha + \alpha \rho)^2} < 0
\]

Since \( \beta_1 > 1 \) and \( 0 < \rho < 1 \), we have that \( A(\rho - 1) \frac{\partial \beta_1}{\partial \sigma^2} - \beta_1 (\beta_1 - 1 + \rho) \frac{\partial A}{\partial \sigma^2} > 0 \). Therefore \( \frac{\partial Y^*}{\partial \sigma^2} > 0 \).

Thus, if the discount rate satisfies \( i > \max\left[\alpha(1-\rho), \sigma^2 - \alpha \rho + \alpha, -\alpha \rho + \alpha\right] \), under risk aversion, the threshold wage rises with earnings variance. Increasing uncertainty over future income induces individuals to stay in school longer, even if they are risk averse. The same is true for risk neutral individuals \( (\rho = 0) \). The conclusion is less clear if the individuals are risk lovers \( (\rho < 0) \), their decision to continue or to leave school will depend on other parameters of the model as well.

### 6.4. An Extension with Jumps in the Income Process

Let us now return to our basic model in which \( Y \) follows a geometric Brownian motion and extend it in a different way. This time we allow for the possibility that at some random point in time \( Y \) will take a Poisson jump\(^9\), which will be described in some detail in Appendix 6.6. This version of the model will describe innate risk of the school career. An education is a sequence of many steps, each with their own feedback. The students may have to take compulsory fields, for which he discovers a lack of ability. There will be many tests and exams, which the student may fail, and realize that his labor market prospects have seriously deteriorated. He may even be banned from school, by not passing minimum requirements. We model this inherent uncertainty of the schooling process as a random event with a major reduction in income. With this in mind, we modify our basic model, assuming that \( Y_t \) follows the mixed Brownian/jump process:

\[
dY_t = aY_t dt + \sigma Y_t dz - Y_t dq
\]

where \( dq \) is the increment of a Poisson process with the mean arrival rate \( \lambda \), and \( dz \) and \( dq \) are independent. We will assume that if an "event" occurs, \( Y \) falls by some fixed\(^{10}\) percentage

\(^9\) Demers (1991) is a more recent contribution to the study of investment when information arrives over time.

\(^{10}\) Notably, an event might cause \( Y \) to fall by some random, rather than fixed amount.
Chapter 6. Education as a Real Option

\[ \theta, \text{ with } 0 \leq \theta \leq 1 \text{ with probability } 1. \] Thus equation (6.5) says that \( Y \) will fluctuate like a geometric Brownian motion but over each time interval \( dt \) there is a small probability \( \lambda dt \) that it will drop to \((1-\theta)\) times its original value, and it will then continue to fluctuate until another event occurs. Note that the expected percentage rate of change in \( Y \) is not \( \alpha \), but instead \( \alpha - \lambda \theta \), because over each interval of time \( dt \) there is a probability \( \lambda dt \) that \( Y \) will fall by 1000 percent. Thus increases in \( \lambda \) reduce the expected rate of capital gain on \( Y \) by increasing the chance of a sudden drop in \( Y \). Second, because a Poisson event occurs only infrequently, most of the time the variance of \( \frac{dY}{Y} \) over a short interval of time \( dt \) is just that of the Brownian motion part \( \sigma^2 dt \). However, if the event occurs, a very large deviation happens, so its contribution to the variance calculated given the information at \( t \) cannot be neglected.

Note that if for simplicity we set \( \alpha=0 \) and write

\[
\begin{align*}
    dY &= \sigma \sqrt{dt} \text{ with probability } \frac{1}{2}(1-\lambda dt), \\
    dY &= -\sigma \sqrt{dt} \text{ with probability } \frac{1}{2}(1-\lambda dt), \\
    dY &= -\theta Y \text{ with probability } \lambda dt.
\end{align*}
\]

then \( EdY= -\lambda dt \theta Y \) and \( \text{Var}(dY) = \sigma^2 Y^2 dt + \lambda \theta Y^2 dt \). This variance has two components. The first component comes from the Brownian motion part of the process, and is conditional on no jump occurring. The second component accounts for the possibility of a jump.

To solve for the optimal stopping rule, a Bellman equation for \( V(Y) \) is used:

\[
V(Y) = u dt + E(V(Y + dY)e^{-\gamma}) \tag{6.6}
\]

We now expand \( E(V(Y + dY)) \) using the version of Ito's Lemma for combined Brownian and Poisson processes:

---

\(^{11}\) When applying Ito's Lemma to a combined Brownian motion/jump process, it is only the first component of this variance that contributes to the new term involving second order derivatives. The jump part contributes a different term involving a difference in values at discretely different points (Dixit and Pyndick (1994), Chapter 3).

\(^{12}\) In the Appendix 6 we present Ito's Lemma for a combination of an Ito process and a jump process.
Chapter 6. Education as a Real Option

\[ iV(Y)dt = udt + \alpha YV'(Y)dt + \frac{1}{2}\sigma^2 Y^2 V(Y)dt - \lambda [V'(Y) - V((l - \vartheta)Y)]dt \]

The solution to equation (5.6.) is again of the form \( V(Y) = CY + \frac{\mu}{i} \), but now \( \beta_1 \) is the positive solution of a slightly more complicated non-linear function:

\[ Q_3: \quad \frac{1}{2} \sigma^2 \beta^2 + \beta(\alpha - \frac{1}{2} \sigma^2) - (i + \lambda) + \lambda(l - \vartheta)^2 = 0 \]

The value of \( \beta_1 \) that satisfies \( Q_3(\beta) = 0 \) can be found numerically. If we evaluate the quadratic \( Q_3 \) at 1 and \( \infty \) we have \( Q_3(1) < 0 \) and \( Q_3(\infty) > 0 \). Therefore \( \beta_1 \) must be larger than 1.

Then given \( \beta_1 \), the values for \( \dot{Y} \) and \( C \) can be found from the boundary conditions\(^\text{13}\). The same boundary conditions apply as before at say \( \dot{Y} \) the threshold. At the optimal exercise point, we have the value matching and smooth matching conditions linking the present value of the expected earnings stream with the appropriate \( V(Y) \).

First we want we express the present value of the expected earnings stream:

\[ P(Y) = E \int_0^\infty Y e^{-it} dt \]

If the wage earned on the labor market follows a Brownian motion all the time and drops by a constant amount \( \theta Y \) at random points in time, where \( \lambda \) is the mean arrival rate of the drops in income. We can treat \( P \) as an asset and equate the normal return on it at rate \( i \) to the sum of dividend (current wage) and the expected capital gain:

\[ iPdt = Ydt + E(dP) \quad (6.7) \]

Since \( Y \) follows a combination of Ito and Jump processes, applying Ito's Lemma to function \( P \), we obtain:

\(^{13}\) If \( \theta = 1 \) (so that the income drops to 0, and remains there forever), \( Q(\beta) \) simplifies to a quadratic equation, which is just like our earlier equation except that the Poisson parameter gets added to the rate of return in the constant term. The positive solution is: \( \beta_1 = \frac{1}{2} \frac{(i - \delta)}{\sigma^2} + \sqrt{\left[ \frac{(i - \delta)}{\sigma^2} \right]^2 + 2 \frac{(i + \lambda)}{\sigma^2}} \).
Chapter 6. Education as a Real Option

\[ E(dP) = \left( \frac{\alpha}{i - \alpha} \right) Ydt + \left( \frac{\lambda \theta}{i - \alpha} \right) \int dt \]

Plugging back the expression of \( E(dP) \) into equation (6.7) we get:

\[ P(Y) = \frac{1}{i - \alpha} Y + \frac{\lambda \theta}{i(i - \alpha)} Y \]

Therefore, \( P \) is equivalent to a perpetuity that pays out forever the current wage plus the capitalized value of the average drop per unit time. The value matching and smooth pasting conditions are:

\[
\begin{cases}
CY^H + \frac{u}{i} = KY \\
C\beta_i Y^{H+1} = K
\end{cases}
\]

where \( K \) is the positive constant, \( K = \frac{i + \lambda \theta}{i - \alpha} > \alpha \).

The system in the unknowns \( C \) and \( Y \) readily yield the solution \( \hat{Y} \) for optimal threshold:

\[ \hat{Y} = \frac{\beta_i}{\beta_i - 1} \frac{i - \alpha}{u + \lambda \theta} \quad (6.8) \]

Thus \( \frac{\partial \hat{Y}}{\partial \lambda} < 0 \) and \( \frac{\partial \hat{Y}}{\partial \theta} < 0 \), therefore the higher the probability that a fall in income will occur (or the larger the drop in income if an event occurs) the less willing will be the individual to stay in school. Thus, we find that in increase in the inherent risk of the schooling process will reduce length of schooling.

Differentiating \( \hat{Y} \) and using the implicit function theorem we get:
Chapter 6. Education as a Real Option

\[ \frac{\partial \hat{Y}}{\partial \sigma^2} = \frac{\partial \beta_1}{\partial \sigma^2} (\beta_1 - 1) - \frac{\partial (\beta_1 - 1)}{\partial \sigma^2} \beta_1 \right) \left( r - \alpha \right) \frac{r + \lambda \theta}{\beta_1} \]

\[ \frac{\partial Q_1}{\partial \sigma^2} = \frac{1}{2} \beta_1 (\beta_1 - 1) > 0 \]

\[ \frac{\partial Q_1}{\partial \beta_1} = \alpha + \sigma^2 (\beta_1 - \frac{1}{2}) + \lambda \beta_1 (1 - \theta)^{n+1} > 0 \]

\[ \text{Implicit Function Theorem} \]

\[ \frac{\partial \hat{Y}}{\partial \sigma^2} < 0 \]

\[ \delta_{U,R,S} \rightarrow \frac{\partial \hat{Y}}{\partial \sigma^2} > 0 \]

Hence, higher the risk, higher the threshold at which the individual is willing to leave school.

Table 6.1. DEPENDENCE OF \( \hat{Y} \) ON \( \theta, \lambda \) AND \( \sigma \) (i=0.08, \( \alpha=0.07 \) and \( \mu=5000 \))

<table>
<thead>
<tr>
<th>( \theta )=0, any ( \lambda )</th>
<th>( \sigma )=0.00</th>
<th>( \sigma )=0.01</th>
<th>( \sigma )=0.03</th>
<th>( \sigma )=0.05</th>
<th>( \sigma )=0.07</th>
<th>( \sigma )=0.09</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{y} )</td>
<td>4999.9</td>
<td>5002.8</td>
<td>5032.0</td>
<td>5088.6</td>
<td>5174.4</td>
<td>5287.0</td>
</tr>
<tr>
<td>( \theta )=0.15, ( \lambda )=0.1</td>
<td>( \hat{y} )</td>
<td>1719.4</td>
<td>1720.9</td>
<td>1732.8</td>
<td>1756.6</td>
<td>1791.8</td>
</tr>
<tr>
<td>( \theta )=0.15, ( \lambda )=0.3</td>
<td>( \hat{y} )</td>
<td>644.4</td>
<td>645.2</td>
<td>651.9</td>
<td>664.6</td>
<td>682.8</td>
</tr>
<tr>
<td>( \theta )=0.15, ( \lambda )=0.5</td>
<td>( \hat{y} )</td>
<td>398.7</td>
<td>399.3</td>
<td>403.6</td>
<td>411.8</td>
<td>423.1</td>
</tr>
<tr>
<td>( \theta )=0.15, ( \lambda )=0.7</td>
<td>( \hat{y} )</td>
<td>305.2</td>
<td>305.6</td>
<td>308.4</td>
<td>313.4</td>
<td>320.3</td>
</tr>
<tr>
<td>( \theta )=0.15, ( \lambda )=0.9</td>
<td>( \hat{y} )</td>
<td>253.2</td>
<td>253.4</td>
<td>255.4</td>
<td>258.8</td>
<td>263.4</td>
</tr>
<tr>
<td>( \theta )=1, ( \lambda )=0.1</td>
<td>( \hat{y} )</td>
<td>454.5</td>
<td>454.8</td>
<td>457.4</td>
<td>462.4</td>
<td>469.6</td>
</tr>
<tr>
<td>( \theta )=1, ( \lambda )=0.3</td>
<td>( \hat{y} )</td>
<td>161.2</td>
<td>161.4</td>
<td>162.2</td>
<td>163.9</td>
<td>166.2</td>
</tr>
<tr>
<td>( \theta )=1, ( \lambda )=0.5</td>
<td>( \hat{y} )</td>
<td>98.0</td>
<td>98.1</td>
<td>98.6</td>
<td>99.6</td>
<td>100.8</td>
</tr>
<tr>
<td>( \theta )=1, ( \lambda )=0.7</td>
<td>( \hat{y} )</td>
<td>70.4</td>
<td>70.4</td>
<td>70.8</td>
<td>71.5</td>
<td>72.3</td>
</tr>
<tr>
<td>( \theta )=1, ( \lambda )=0.9</td>
<td>( \hat{y} )</td>
<td>54.9</td>
<td>54.9</td>
<td>55.2</td>
<td>55.7</td>
<td>56.3</td>
</tr>
</tbody>
</table>
In Table 6.1, we compute the values of the threshold $\hat{Y}$ for various levels of $\lambda$, $\theta$ and $\alpha$. In this table we assume $i=0.08$, $\alpha=0.07$ and $u=5000$. A positive value of $\lambda$ affects the value of stopping school in two ways. First, it reduces the expected rate of capital gain on $Y$ (from $\alpha$ to $\alpha-\lambda$) which reduces $V(Y)$. Second, it increases the variance of percentage changes in $Y$ over finite intervals of time, and this tends to increase $V(Y)$. As Table 6.1. shows this net effect is quite strong; small increases in $\lambda$ lead to substantial drop in $\hat{Y}$. By contrast, the effect of $\sigma$ is quite small. The effects of $\lambda$ and $\sigma$ are virtually independent (i.e. within rows, the ratio of the values of $\hat{Y}$ are barely different between columns). We have increased $\lambda$ while holding $\alpha$ fixed. One could argue that the rate of return to education, $i=\alpha-\lambda \theta$, should remain constant, so that an increase in $\lambda$ should be accompanied by a commensurate increase in $\alpha$ (otherwise nobody would remain in school anymore). Suppose $\theta=1$, the income falls to 0 when an event occurs. Then if $\alpha$ increases exactly as much as $\lambda$, so that $\alpha-\lambda$ remains constant, we would have to replace the terms $(i-\delta)$ in equation $Q(\beta)$ with $(i+\lambda-\delta)$. In this case an increase in $\lambda$ would be equivalent to an increase in the rate of return $i$, and would lead to an increase in $V(Y)$ and $\hat{Y}$. Thus, an increase in risk as an increase in the probability of the major event while holding the expected rate of return constant, will increase schooling length. In that sense, the reduction in schooling length with an increase in the probability $\lambda$ or the size of the shock $\theta$ is apparently due to the decrease in the rate of return. This is a consequence of our return to earnings maximization, or risk neutrality. Under risk aversion, there will also be a compensating increase in the rate of return $\alpha$, but it will now also have to compensate for the utility loss from the uncertainty per se.

The decision to continue studying is fraught with uncertainty concerning acceptance into the next stage and its successful completion. If $\lambda$ represents the probability of failing a class or dropping out from school, the model can be used for determination of appropriate policies for altering flows in the system. Such policies may include scholarships and alterations in the dropout rates.

### 6.5. Conclusions

In this chapter we developed a structural model to reflect sequential choices of education in a world with uncertain prospects. Our starting point is Harmon and Walker's (2001) Real Option Theory application to the problem of education choice where returns to education are
uncertain. Each period the individual may choose to stay in school or to exercise the option of stopping the education and go on the labor market. Harmon and Walker (2001) assume static expectations or myopia, that is when the individual chooses the option to leave school, he receives a wage determined by the level of education via a Brownian motion, but the threshold income that triggers the exercising of the option is earned forever afterwards.

We cast doubt on the realism of this claim, since in most real situations the individual must take into account that the future is and always will be uncertain. Hence a more natural theoretical approach is to assume that the individual has rational expectations about the probabilistic law of motion for its uncertain environment. Our model does just that, the individual’s decisions are optimal given the stochastic process underlying the income. The Brownian motion that governed the motion while in school continues after leaving the school. Hence in our model the wages earned while on the labor market are not fixed but evolve stochastically, with the threshold wage as the initial level.

Even with this modification in the model we find that uncertainty delays individuals from leaving school. This unusual result stems from the irreversibility of the option to leave school. High volatility delays the option of leaving school, to see which way the uncertainty goes. If we adapt our basic model and allow for risk aversion rather than risk neutrality, we still find that increasing risk in the wages will increase the length of schooling. However, if we allow for inherent risk of the schooling process, by modeling the probability of a shock that reduces the wage (like failing tests or exams) we find that increasing risk reduces the length of schooling. Thus, we have shown that risk has two faces, which we might call the labor market risk (wage risk) and the school risk (shocks in educational performance), which will have opposing effects. This result is more in tone with intuitive anticipation of the effect of risk.
Appendices to Chapter 6

Appendix 6.1. Wiener Processes

A Wiener process—also called a Brownian motion—is a continuous time stochastic process with three properties: first it is a Markov process, second it has independent increments and third changes in the process over any finite interval of time are normally distributed with a variance that increases linearly with the time interval.

Formally, if \( z(t) \) is a Wiener process, then any change in \( z \), \( \Delta z \), corresponding to a time interval \( \Delta t \), satisfies the following conditions:

1. \( \Delta z = \varepsilon \sqrt{\Delta t} \), where \( \varepsilon \) is a normally distributed random variable with a mean of zero and a standard deviation of 1.
2. The random variable \( \varepsilon \) is serially uncorrelated (\( \text{E}(\varepsilon_t, \varepsilon_s) = 0 \) for \( t \neq s \)).

By letting \( \Delta t \) become infinitesimally small, the increment of a Wiener process is represented in continuous time as \( \, dz = \varepsilon \sqrt{dt} \). The simplest generalization of this formula is the Brownian motion with drift:

\[
dY = a dt + \varepsilon dz
\]

where \( dz \) is the increment of a Wiener process as defined above.

Appendix 6.2. Ito’s Lemma

If we take the generalization of the simple Brownian motion with drift:

\[
dY = a(Y, t) dt + b(Y, t) dz
\]

where \( dz \) is the increment of a Wiener process and \( a(Y, t) \) and \( b(Y, t) \) are known (nonrandom) functions. In addition we consider a function \( F(Y, t) \) that is at least twice differentiable in \( Y \) and once in \( t \) then the total differential of \( F \), denoted \( dF \) is given as:

\[
dF = \left[ \frac{\partial F}{\partial t} + a(Y, t) \frac{\partial F}{\partial Y} + \frac{1}{2} b^2(Y, t) \frac{\partial^2 F}{\partial Y^2} \right] dt + b(Y, t) \frac{\partial F}{\partial Y} dz
\]
Chapter 6. Education as a Real Option

Appendix 6.3. The Expected Value of \(Y_t\) and The Variance of \(Y_t\)

If \(Y_t\) follows a geometric Brownian motion with drift:
\[
dY_t = \alpha Y_t dt + \sigma Y_t dz
\]
where \(dz\) is the increment of a Wiener process, then \(F(Y_t) = \log(Y_t)\) follows the geometric Brownian motion with drift:
\[
dF(Y_t) = (\alpha - \frac{1}{2} \sigma^2) dt + \sigma dz
\]
Hence over a finite time interval \((0, t)\) the change in the logarithm of \(Y_t\) is normally distributed with mean \((\alpha - \frac{1}{2} \sigma^2)t\) and variance \(\sigma^2 t\).

If we denote \(Y_t\) at \(t=0\) with \(Y_0\) then

\[
e^{\log(Y_t) - \log(Y_0)} \sim \mathcal{N}\left(\alpha - \frac{1}{2} \sigma^2)t, \sigma^2 t\right)
\]

\[
E\left(\frac{Y_t}{Y_0}\right) = e^{(\alpha - \frac{1}{2} \sigma^2)t + \sigma^2 t^2/2} \Rightarrow E(Y_t) = Y_0 e^{\alpha t}
\]

\[
\text{Var}\left(\frac{Y_t}{Y_0}\right) = e^{2(\alpha - \frac{1}{2} \sigma^2)t + \sigma^2 t^2} (e^{\sigma^2 t} - 1) \Rightarrow \text{Var}(Y_t) = Y_0 e^{2\alpha t} (e^{\sigma^2 t} - 1)
\]

Appendix 6.4. Present value of the lifetime revenue

If \(Y_t\) follows a geometric Brownian motion with drift:
\[
dY_t = \alpha Y_t dt + \sigma Y_t dz
\]
where \(dz\) is the increment of a Wiener process.
If we denote \(Y_t\) at \(t=0\) with \(Y_0\) then using the results in the Appendix 6.3, we obtain that:
\[
E \int_0^\infty Y_t e^{-\alpha t} dt = \int_0^\infty Y_0 e^{\alpha t} e^{-\alpha t} dt = \frac{Y_0}{i - \alpha}
\]
Appendix 6.5. Proof of equation: \( \frac{1}{2} \sigma^2 Y^2 \nu''(Y) + \alpha Y \nu'(Y) - i \nu(Y) + u = 0 \)

We start with the arbitrage equation:

\[
V(Y) = u dt + E(V(Y + dY)e^{-idt})
\]

and we obtain the differential equation:

\[
\frac{1}{2} \sigma^2 Y^2 \nu''(Y) + \alpha Y \nu'(Y) - i \nu(Y) + u = 0
\]

\[
V(Y) = u dt + E(V(Y + dY)e^{-idt}) \Leftrightarrow V(Y) = u dt + (1-idt)E(V(Y + dY))
\]

using a Taylor expansion of \( V(Y + dY) \) around \( Y \) we get that

\[
(1-idt)E(V(Y + dY)) = (1-idt)E[V(Y) + V'(Y)dY + \frac{1}{2}V''(Y)(dY)^2 + o(dt)]
\]

therefore

\[
V(Y) = u dt + E[V(Y) + V'(Y)dY + \frac{1}{2}V''(Y)(dY)^2 + o(dt)] - idtE[(V'(Y) + V''(Y)(dY)^2 + o(dt)]
\]

Moreover the last term of the right hand side can be written as:

\[
\text{idtE}[(V(Y) + V'(Y)dY + \frac{1}{2}V''(Y)(dY)^2 + o(dt)] = \text{idtE}[(V(Y)] + \text{idtE}[V'(Y)dY + \frac{1}{2}V''(Y)(E(dY)^2 + o(dt)]
\]

and the last term: \( \text{idtE}[V'(Y)dY + \frac{1}{2}V''(Y)E(dY)^2 + o(dt)] \) is equal to 0.

That is because

\[
\text{idtE}[V'(Y)dY + \frac{1}{2}V''(Y)(E(dY)^2 + o(dt)] = \text{idtE}[V'(Y)\alpha Y dt + \frac{1}{2}V''(Y)\sigma^2 Y^2 dt + o(dt)]
\]

and dt raised at powers higher or equal to 2 disappears as dt goes to 0.
Thus:

\[ V(Y) = u dt + V(Y) + E[V'(Y) dY] + E\left[ \frac{1}{2} V''(Y) (dY)^2 \right] - i dt E[V(Y)] \]

\( V(Y) \) cancels out and we obtain that:

\[ 0 = u dt + V'(Y) E[dY] + \frac{1}{2} V''(Y) E[(dY)^2] - i dt E[V(Y)] \]

But

\[ dY = \alpha Y dt + \sigma Y dz \Rightarrow (dY)^2 = (\alpha Y^2 (dt))^2 + 2(\alpha Y)(\sigma Y)(dt)^2 + (\sigma Y)^2 dt \]

with the same argument that \( dt \) in powers higher or equal to 2 go to zero faster than \( dt \) as \( dt \) goes to 0, the only term that remains in the expression above is \( (\sigma Y)^2 dt \).

Finally,

\[ \frac{1}{2} \sigma^2 Y^2 V''(Y) + \alpha Y V(Y) - iV(Y) + u = 0. \]

**Appendix 6.6. Jump Processes**

A Poisson process is a process subject to jumps of fixed or random size, for which the arrival times follow a Poisson distribution. These jumps are called events. Letting \( \lambda \) denote the mean arrival rate of an event, during a time interval of infinitesimal length \( dt \), the probability that an event will occur is given by \( \lambda dt \), and the probability that an event will not occur is given by \( 1-\lambda dt \). The event is a jump of size \( u \), which can itself be a random variable.

The Poisson process is denoted \( q \), by analogy with the Wiener process. Hence:

\[ dq = \begin{cases} 0 & \text{with probability } 1 - \lambda dt \\ \nu & \text{with probability } \lambda dt \end{cases} \]
Chapter 6. Education as a Real Option

If \( Y_t \) follows a geometric Brownian motion with drift:

\[
dY_t = \alpha Y_t dt + \sigma Y_t dz
\]

where \( dz \) is the increment of a Wiener process

then the stochastic process for the variable \( Y_t \) is written as a Poisson differential equation, as follows:

\[
dY = a(Y,t)dt + b(Y,t)dq
\]

where \( a(Y,t) \) and \( b(Y,t) \) are known (nonrandom) functions.

Appendix 6.7. Ito’s Lemma for combination of an Ito Process and a Jump Process

Sometimes we meet a combination of an Ito process and a jump process. The former goes on all the time, the latter occurs infrequently. The appropriate Ito’s Lemma combines the two effects. Thus if:

\[
dY = a(Y,t)dt + b(Y,t)dz + g(Y,t)dq
\]

and \( G(Y,t) \) is some differentiable function of \( Y \) and \( t \), then the expected change in \( H \) is given by:

\[
dH = \left[ \frac{\partial H}{\partial t} + a(Y,t) \frac{\partial H}{\partial Y} + \frac{1}{2} b^2(Y,t) \frac{\partial^2 H}{\partial Y^2} \right] dt + E \left[ \lambda \left( H(Y + g(Y,t)v,t) - H(Y,t) \right) \right] dq
\]

Appendix 6.8. Density of Optimal schooling Length

Because the evolution of the income is stochastic, we can only simulate the optimal choice of schooling. If the individual starts with an income \( Y_{\text{init}} \), how long will it take for income to reach the threshold value?
If $Y$ evolves according to a Brownian motion law then the expected percentage rate of change in $Y$ is:

$$\frac{dY}{Y} = \alpha ds + \sigma dz$$

The probability that an individual will be in school at time $t$ (so that $S^*$, the optimal schooling choice, is greater than $t$) is equal to the probability that the income process will not have reached the trigger level, $Y^0$, at time $t$.

Since

$$\ln Y_t - \ln Y_{\text{initial}} \sim N\left(\left(\alpha - \frac{\sigma^2}{2}\right)t, \sigma^2 t\right)$$

we obtain that:

$$P(S^* \leq t) = 1 - P(S^* > t) = 1 - P(\ln Y_t < \ln Y^0) = 1 - \Phi \left(\frac{\ln Y^0 - \ln Y_{\text{initial}} - \left(\alpha - \frac{\sigma^2}{2}\right)t}{\sigma \sqrt{t}}\right)$$

(6.9.)

where $\Phi$ is the cumulative distribution function of a standard normal random variable.

If the income evolves as a Brownian motion with Poisson jump then:

$$\frac{dY}{Y} = (\alpha - \lambda \theta) ds + \sigma dz$$

In this case,

$$\ln Y_t - \ln Y_{\text{initial}} \sim N\left(\left(\alpha - \lambda \theta - \frac{\sigma^2}{2}\right)t, \sigma^2 t\right)$$

which implies:
Chapter 6. Education as a Real Option

\[ P(S^* \leq t) = 1 - P(S^* > t) = 1 - P(\ln Y \leq \ln \hat{Y}) = 1 - \Phi \left( \frac{\ln \hat{Y} - \ln Y_{\text{initial}} - \left(\frac{\alpha - 2 \theta - \sigma^2}{2}\right) t}{\sigma \sqrt{t}} \right) \]  

(6.10.)

where \( \hat{Y} \) is the threshold income.

Two numerical simulations of the equations (6.9.) and (6.10.) are shown in the figures 6.1 and 6.2. We see that the density of schooling times is approximately normal with mean around 6. The density tends to flatten as \( \sigma \) or \( \lambda \) decrease, but remains centered around 6. This is also quite clear from Table 6.1. that shows thresholds values for different levels of risk. Figure 6.1. plots a surface where each cross section represents the density function of schooling time \( S^* \) (from equation 6.9.) and Figure 6.2. plots a surface where each cross section represents the density function of schooling time \( S^* \) (from equation 6.10.).

Figure 6.1. Density of \( S^* \) for Different Levels of Risks \( \sigma \) \( (Y_{\text{initial}}=10000, Y^0=15000, \alpha=0.08) \)
Figure 6.2. Density of $S^*$ for Different Levels of Risks $\lambda$. ($Y_{\text{initial}} = 10\,000$, $\hat{Y} = 15\,000$, $\alpha = 0.08$, $\theta = 0.15$, $\sigma = 0.055$)