September 2013

Valuation of the Peterborough Prison Social Impact Bond

Majid Hasan
The University of Western Ontario

Supervisor
Mark Reesor
The University of Western Ontario

Graduate Program in Applied Mathematics

A thesis submitted in partial fulfillment of the requirements for the degree in Master of Science

© Majid Hasan 2013

Follow this and additional works at: http://ir.lib.uwo.ca/etd

Part of the Finance Commons, Numerical Analysis and Computation Commons, Other Applied Mathematics Commons, and the Social Welfare Commons

Recommended Citation

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact tadam@uwo.ca.
VALUATION OF THE PETERBOROUGH PRISON SOCIAL IMPACT BOND

(Thesis format: Integrated Article)

by

Majid Hasan

Graduate Program in Applied Mathematics

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Majid Hasan 2014
Abstract

The Peterborough Prison Bond is a social impact bond (SIB) that was issued by the UK government to reduce recidivism rate in the Peterborough prison. Most of the literature on the SIB so far has been focused on the opportunities, challenges, and the related policy issues (see (Fox), (Strickland), and (Disley)), and little effort has been made to provide a mathematical framework to determine a fair price for such instruments. Here, we aim to provide a pricing framework for the bond. We price the bond both from the issuer's and the buyer's perspective, by adjusting for the bond's risk, ambiguity, and social impact. Results suggest that the issuer could maximize its financial profit by targeting investors with a lower aversion to charity, and it could maximize the bond’s social impact by targeting the investors with lower aversion to risk.

Keywords

Acknowledgments

This thesis would not have been possible without the help of several individuals who contributed their valuable time and assistance in the completion of this work.

I would like to thank my supervisor Prof. Mark Reesor for his support, guidance, and ideas in completing this research. I would also like to thank my thesis examiners for their valuable feedback on this thesis, and the members of financial mathematics group for providing an intellectually rich, friendly and supportive work environment.

In the end, I would like to thank my family for their constant support and encouragement in the pursuit of my goals.
# Table of Contents

Abstract ............................................................................................................................... ii   
Acknowledgments .............................................................................................................. iii  
Table of Contents ............................................................................................................... iv  
List of Tables .................................................................................................................... vii  
List of Figures .................................................................................................................. viii  
List of Appendices ............................................................................................................ xv  

Chapter 1 ............................................................................................................................. 1  
1. INTRODUCTION ......................................................................................................... 1  
   1.1. Peterborough Prison Social Impact Bond (Disley) ................................................. 3  
   1.2. Reference Parameters .......................................................................................... 7  
   1.3. Yield to Maturity ................................................................................................. 8  

Chapter 2 ........................................................................................................................... 10  
2. Modeling the Underlying Reconviction Rate .............................................................. 10  
   2.1. Summary Statistics ............................................................................................. 12  
   2.2. Fitting Probability Distribution to Reconviction Data .......................................... 12  
   2.2.1. Method of Moments ...................................................................................... 13  
   2.2.2. Maximum Likelihood Parameter Estimation ............................................... 14  
   2.2.3. Comparison of Most Likely and Method of Moments Estimates ................. 15  
   2.3. Sensitivity Analysis for Distribution Parameters .............................................. 16  
   2.4. Predictive Data Set ............................................................................................. 19  

Chapter 3 ........................................................................................................................... 24  
3. Pricing Peterborough Prison SIB ................................................................................. 24  
   3.1. Pricing the SIB using Wang Transform (Wang) ................................................. 24  
   3.2. Indifference Pricing ............................................................................................. 32
List of Tables

Table 1: Summary statistics for the raw data sets ................................................................. 12

Table 2: Comparison of summary statistics. 95% confidence interval for the mean is [6.27%, 13.12%] .................................................................................................................................. 23

Table 3: Sensitivity of the skew normal and the beta distribution parameters on the Wang risk aversion parameter ................................................................. 28

Table 4: Slope and convexity of price curves for 1% change in the values of interest rate and risk aversion parameter for $\lambda = 0.1, r = 3\%$, $\Delta \lambda = 0.01$ and $\Delta r = 0.01$. $P +$ refers to the price at lower $\lambda/r$, and $P -$ refers to the price at higher $\lambda/r$. Hence, the slope is defined to be the negative of the actual slope. .................................................................................. 30

Table 5: Bins and weighted frequency ................................................................................. 101

Table 6: Weights assigned to rehab services offered to Peterborough Prisoners ............ 102

Table 7: Historical data for the rehabilitation programs from around the world .......... 106

Table 8: Description of all the variables used and associated symbols is presented here. ... 111

Table 9: All the functions and associated notation used is shown here ......................... 114
List of Figures

Figure 1: The left panel shows the yield to maturity for different notional amounts, and the right panel shows the yield to maturity for different prices. Both plots assume that the reconviction rate drop by same percentage among all cohorts. ................................................ 8

Figure 2: Histogram of Reconviction Rate data. Negative (positive) values of the reconviction rate correspond to an increase (decrease) in the reconviction rate. ................. 11

Figure 3: Method of Moment and the Maximum Likelihood fits to the raw data for the skew normal (left panel) and the beta distribution (right panel). ..................................................... 15

Figure 4: A scatter plot of maximum likelihood parameters for the beta (left panel) and the skew normal (right panel) distributions. Color represents the value of log-likelihood function for the given value of parameters, and it increases from blue to red ( ). Parameters for beta distribution are much less scattered than the skew normal distribution. 17

Figure 5: Histograms for skewness (left panel), position (middle panel), and scale (right panel) of skew normal distribution. ........................................................................................ 17

Figure 6: Histograms for $\rho$ (left panel) and $\delta$ (right panel) parameters of the beta distribution. ......................................................................................................................... 18

Figure 7: Price histograms for the skew-normal (left panel) and the beta (right panel) distributions. Price is calculated using indifference pricing mechanism described later. ...... 19

Figure 8: Predictive probability density functions for the beta and the skew normal distribution. ................................................................................................................................. 20

Figure 9: Maximum likelihood distributions for predictive data set for the skew normal (left panel) and the beta (right panel) distributions. ......................................................... 20

Figure 10: Comparison of the most likely distributions for raw and predicted data sets for the skew normal (left panel) and the beta (right panel) distributions. ............................... 22
Figure 11: Risk adjusted skew normal (left) and beta (right) density functions for different values of risk aversion parameter. .......................................................................................................................... 26

Figure 12: Effect of the Wang risk aversion parameter on distribution statistics. Changes in the distribution mean, standard deviation, skewness and kurtosis are shown in top left, top right, bottom left and bottom right panels, respectively, for both the skew normal and the beta distributions.................................................................................................................................................. 27

Figure 13: Plot in the left panel shows the bond price with risk aversion parameter (\( \lambda \)), and plot in the right panel shows bond price with interest rate (\( r \)). ................................................................. 30

Figure 14: The effect of risk adjustment on the bond price. \( \lambda = 0 \) corresponds to no risk adjustment, and the price there is just the expected present value of all the payouts. ............... 32

Figure 15: Dotted lines indicate the expected present value of the bond, and the solid lines show the risk adjusted price of the bond. Both curves have been drawn for the reference parameters defined in equations (2.11) and (2.12). ................................................................. 34

Figure 16: Utility risk aversion parameter plotted with the Wang risk aversion parameter for both the skew normal and the beta distributions using the reference parameters................. 36

Figure 17: Variation in the fair price with changes in interest rate (left panel) and risk aversion (right panel) for the skew normal distribution...................................................................................... 37

Figure 18: Variation in the fair price with changes in interest rate (risk aversion) and risk aversion (right panel) for the beta distribution.............................................................. 37

Figure 19: The fair price of the bond for different values of skewness and position parameters is shown. Price increases with increasing \( \theta \), all else held fixed. Curves in the left panel are plotted by holding the position parameter (\( \mu \)) fixed at \( \mu = \mu_{\text{ref}} \) and the curves in the right panel are plotted by holding the skewness parameter (\( \alpha \)) fixed at \( \alpha = \alpha_{\text{ref}} \). For the right panel, \( \eta = \mu \) and \( \eta_{\text{ref}} = \mu_{\text{ref}} \) .................................................................................................................................................. 43

Figure 20: Variation of the fair price plot with the skewness parameter (left panel) and the position parameter (right panel) for \( y = 0.2 \). Dotted lines in both panels have been moved up, by an amount equal to the reference price, to be visible on the plot................................. 44
Figure 21: Surface (left panel) and contour (right panel) plots with the skew normal distribution parameters are shown. For both panels, $\gamma = 1$ and $\theta = 5$. ........................................ 45

Figure 22: The left panel shows that with increasing faith in the reference model, candidate model approaches the reference model and two prices get closer. The right panel shows the ambiguity and risk adjusted prices with the faith in reference model. For both panels, $\gamma =1.46$

Figure 23: The left panel shows the ambiguity adjusted price (blue curve) and the risk adjusted price evaluated at the optimal parameters (green curve) as a function of $\theta$. All curves in the left panel are plotted using the same level of risk aversion. The right panel shows the difference between ambiguity adjusted price and the risk adjusted price evaluated at the optimal parameters. This difference can be seen as the effect of ambiguity aversion that cannot be incorporated with the risk aversion only. ............................................................... 47

Figure 24: Difference between the ambiguity adjusted price and the reference risk price is decomposed into the contribution from a change in model parameters (red curve) and a contribution from decrease in risk aversion level (blue curve) is shown in the left panel. The right panel shows the ratio of the red curve to the blue curve. ............................................... 48

Figure 25: Variation of the price function with the faith in reference model for the beta distribution parameters $\rho$ (left panel) and $\delta$ (right panel) is shown. ................................................. 50

Figure 26: Plot of the price function for beta distribution parameters $\rho$ (left panel) and $\delta$ (right panel) for $\gamma = 0.2$ is shown. Point where all curves intersect is the reference value of the parameter................................................................................................................................. 51

Figure 27: Surface (left panel) and contour (right panel) plots of the price function with the beta distribution parameters are shown................................................................................... 52

Figure 28: The left panel shows the variation of the optimal parameters and the right panel shows the variation of the ambiguity adjusted price of the bond with the faith in reference model, for $\gamma = 4.34$................................................................. 52
Figure 29: The left panel compares the ambiguity adjusted price (blue curve) with the risk adjusted price evaluated at the optimal parameters (green curve). The right panel shows the difference between the ambiguity adjusted price and the risk adjusted price. ........................ 53

Figure 30: Difference between the ambiguity adjusted price and the reference risk price is decomposed into the contribution from a change in model parameters (red curve) and a contribution from decrease in risk aversion level (blue curve) is shown in the left panel. Right panel shows the ratio of the red curve to the blue curve. ........................................................ 54

Figure 31: Effect of ambiguity adjustment on the skew normal and the beta distributions. The left panel shows the contribution of the penalty term for both distributions, and right panel shows the ratio of the contribution of the change in parameters to the contribution of the penalty term. ........................................................................................................................... 55

Figure 32: Variation of the fair price function with $\theta$ for different values of skewness (left panel) and position (right panel) parameters is shown. ................................................................. 58

Figure 33: Variation of the fair price function with the skewness and the position parameters is shown in left and right panels respectively. ................................................................. 58

Figure 34: Surface (left panel), contour (right panel) plots of the price function with the skew normal distribution parameters are shown............................................................... 59

Figure 35: The optimal parameters (left panel) and the fair price calculated at the optimal parameters (right panel) are shown.......................................................... 60

Figure 36: Comparison of the ambiguity adjusted price with the risk adjusted price evaluated at optimal parameters is shown in the left panel. The right panel shows the difference between the two prices as a percentage of the risk adjusted price (blue curve in the left panel)................................................................................................................................... 61

Figure 37: Comparison of the two contribution to the ambiguity adjusted price. The left panel shows the individual contributions as a percentage of difference between the ambiguity adjusted price and the reference risk adjusted price, and the right panel shows the ratio of the two contributions. ........................................................................................................... 61
Figure 38: Variation of the price function with $\theta$ for different $\rho$ (left panel) and $\delta$ (right panel) is shown. ................................................................. 63

Figure 39: Variation of the price function with the beta distribution parameters $\rho$ (left panel) and $\delta$ (right panel) is shown. ................................................................. 64

Figure 40: Surface (left panel) and contour (right panel) plots of the price function with $\rho$ and $\delta$ are shown................................................................. 65

Figure 41: The optimal parameters (left panel) and the ambiguity adjusted price (right panel) with the faith in reference model are shown........................................ 66

Figure 42: Comparison of ambiguity adjusted price with the risk only price evaluated at optimal parameters is shown in the left panel. The right panel shows the difference as a percentage of the risk adjusted price......................................................... 66

Figure 43: Comparison of the two contribution to the ambiguity adjusted price. The left panel shows the individual contributions as a percentage of difference between ambiguity adjusted price and reference risk price, and the right panel shows the ratio of the two contributions. 67

Figure 44: The left panel compares the effect of penalty term for the skew normal and the beta distributions, and the right panel compares the ratio of the two contributions for the skew normal and the beta distributions................................................................. 68

Figure 45: Comparison of the Wang and the Utility formulations for the beta distribution. The left panel shows the ambiguity and risk adjusted prices for both pricing mechanisms, and the right panel shows the ratio of the parameters’ contribution to price to the penalty function’s contribution to price................................................................. 69

Figure 46: Comparison of the Wang and the Utility formulations for the beta distribution. The left panel shows the penalty function’s contribution to price, and the right panel shows the ratio of the parameters’ contribution to price to the penalty function’s contribution to price................................................................. 70

Figure 47: The left panel shows that the fair price increases with increasing $A$ (relative strength of social utility), and the right panel shows that the fair price decreases with
increasing $\gamma_2$. For small $A$ and large $\gamma_2$ price approaches the price of the bond in the absence of social utility................................................................. 74

Figure 48: Variation of the fair price with aversion to charity (left panel) and risk aversion parameter (right panel) is shown. For the left panel $\gamma_1 = 1$, and for the right panel $\gamma_2 = 1$.
................................................................................................................................. 75

Figure 49: Variation of the optimal risk aversion parameter with the charity aversion parameter for different values of $A$ is shown in the left panel. The right panel shows the derivative of the fair price with the risk aversion parameter. ............................................. 76

Figure 50: The left panel shows the constant price curves for different prices and same $A$, and the right panel shows the constant price curves for different $A$ and the same price....... 77

Figure 51: The left panel shows the risk adjusted price for $zk = 13\%$, and the right panel shows the price for $k = 0.01\%$. Both panels have $c = 0.01\%$, $\gamma I = 0.1$, and $\gamma B = 1$...... 81

Figure 52: The left panel shows the issuer’s price with the notional ($k$) for $zk = 13\%$. The right panel shows the issuer’s price with the cap rate ($zk$) for $k = 0.008$. For both panels, $c = 0.01\%, pf = 1\%, \gamma I = 0.1$, and $\gamma B = 1$ ................................................................. 83

Figure 53: Plots in the upper panel show the surface plots of the buyer and issuer price, while surface plot in the lower left panel shows issuer’s profit and the contour plot in the lower right panel shows the region where profit exceeds zero. All plots have been obtained for $c = 0.01\%, pf = 1\%$, $\gamma I = 0.1$, and $\gamma B = 1$. All plots have the same x,y, and z axes...... 84

Figure 54: The left panel shows the issuer’s price with the notional amount, and the right panel shows issuer’s price with the cap rate. Total price is represented by the yellow curve. For both panels: $\gamma I = 0.1$, $\gamma S = 0.1$, $\gamma B = 1$, $S = 0.1$, and $B = 1$................................. 86

Figure 55: The upper left and right panels show surface plots of the buyer’s price and the issuer’s price with the bond’s notional and cap rate, respectively. The lower left and right panels show surface and contour plots of the issuer’s profit, respectively. For all panels: $\gamma I = 0.1$, $\gamma S = 0.1$, $\gamma B = 1$, $S = 0.1$, and $B = 1$. All plots have the same x,y, and z axes.
..................................................................................................................................... 88
Figure 56: The left panel shows the buyer’s price with the issuer’s welfare spending, and the right panel shows the buyer’s price less the issuer’s subsidy. ......................................................... 89

Figure 57: The issuer’s profit (left panel) and the issuer’s profit per unit welfare spending (right panel) is shown.......................................................... 90

Figure 58: Relationship between the investor’s risk and charity aversion parameter (left panel) and the optimal welfare spending for a given choice of choice of risk preference (right panel) is shown. .................................................. 91

Figure 59: Variation of the optimal parameters (left panel) and the issuer’s optimal profit (right panel) is shown with the welfare spending.................................................. 92

Figure 60: Net price paid by the buyer (left panel) and the decomposition of the issuer’s profit (right panel) are shown. .......................................................... 93

Figure 61: Contributions of the issuer’s savings and buyer’s price to the issuer’s profit (left panel) and the issuer’s profit margin (right panel) is shown. .............................. 95

Figure 62: Social Impact of the bond (left panel) and the total impact of the bond (right panel) are shown. .......................................................... 95
List of Appendices

Appendix 1: Rehabilitation Program Data..............................Error! Bookmark not defined.101
Appendix 2: Method of Moments........................................Error! Bookmark not defined.107
Appendix 3: Maximum Likely Parameter Estimation.................Error! Bookmark not defined.108
Appendix 4: Notation..........................................................Error! Bookmark not defined.110
Chapter 1

1. INTRODUCTION

Social Impact Bonds are multi-stakeholder partnerships between private, philanthropic, non-profit, and public sectors. Private and philanthropic investors pay for preventive interventions that are carried out by non-profit service providers. If certain social outcomes are achieved, government saves money on its future expenditures, and a part of these savings is paid to the SIB investors. As a result, investors earn a blended social and financial return that is uncorrelated with the financial market; government reduces its risk of spending tax-payers money on unsuccessful projects; service providers get the needed cash upfront; and the society at large benefits as overall quality of life improves.

The first SIB was issued in the UK to reduce reconviction rate in the Peterborough Prison. Since then, SIBs have gained popularity and currently work is underway on several bonds in the UK\(^1\), Canada\(^2\), Australia\(^3\), and the US\(^4\). This rapid increase in the number of SIBs is a clear indication of their appeal to all the stakeholders. In detailed interviews with 22 individuals\(^5\) from organizations involved in the development and implementation of the Peterborough Prison SIB, almost all stakeholders expressed positive sentiments about the project.

\(^1\) http://www.socialfinance.org.uk/work/sibs
\(^3\) http://www.afr.com/p/national/westpac_cba_embrace_social_bonds_myji01zs3Mzr2EVCbfTJrkM
\(^5\) http://www.socialfinance.org.uk/sites/default/files/social-impact-bond-hmp-peterborough%5B1%5D.pdf
Despite their growing popularity, most of the literature on SIBs has been focused on the identification of suitable intervention areas (Fox), development guidelines for the SIB (Finance, Social), and potential benefits to public sectors, but no effort has been made to provide a mathematical framework to value such derivatives. However, in order for SIBs to become publicly traded securities, a rigorous pricing framework that allows investors to gauge the risk and return characteristics of the investment, is necessary. Here we make an effort to provide such a mechanism.

We first obtain a probability distribution for the outcome of Peterborough prison rehab program by analyzing the performance of past rehab programs from around the world. A weighting mechanism is used to take into account the differences between the rehab programs. This distribution then lays the foundation for a discounted cash flow valuation of the bond. Two valuation mechanisms are presented here. One mechanism uses the Wang Transform, and the other uses the Max-Min Expected utility theory. Both mechanisms adjust for both the risk associated with the bond’s cash flows and the ambiguity associated with the model for the underlying reconviction rate. The Peterborough SIB, in addition to a financial return, also provides a social return by offering investors an opportunity to invest in a socially beneficial project. This social return is what enticed many charitable institutions, and socially responsible investors to invest in the bond, and it is important include its effect in determining the fair price of the bond. Therefore, the effect of the social impact of the bond on the utility of a socially responsible investor – one who would be willing to trade off financial return for the social impact – is also considered, and its effect on fair price of the bond is analyzed.

To provide a complete picture, we also price the bond from the issuer’s perspective taking into account the issuer’s risk adjusted price for the net bond payout (bond payout to buyers less the issuer’s savings from the drop in recidivism rate), and the issuer’s willingness for welfare spending. We compare the buyer’s price with the issuer’s price to obtain a set of values for the bond’s notional and cap rate that can lead to voluntary transaction between the buyer and the seller. We also find the optimal values for the bond’s notional and cap rate that maximize the issuer’s profit for a given set of risk and
welfare preferences. In the end, we show how the issuer could benefit from the price information obtained from the first bond to optimize its profit for such bonds in future.

The main conclusions of this work are

- The lack of quantitative data available for the prison rehabilitation programs makes the ambiguity in the model a major source of uncertainty.
- The skew normal distribution provides a better fit for the reconviction rate data as compared to the beta distribution.
- For a given level of risk aversion and pricing mechanism, the beta distribution yields lower price than the price under the skew normal distribution.
- The fair price calculated using the Wang transform is more sensitive to the level of risk aversion and interest rates, as compared to the price calculated using utility mechanism.
- The effect of the ambiguity adjustment can be seen as a trade-off between discount for the risk of the bond and the discount for the ambiguity in the model used for the recidivism rate.
- For a given risk adjusted price and a given level of faith in the reference model used for the recidivism rate, the Wang transform yields a lower ambiguity adjusted price than the price under utility mechanism.
- A socially responsible investor can find a unique value for the risk aversion parameter that minimizes the investor’s price, for a given aversion to charity.
- In the absence of any budget constraints, the issuer maximizes its expected profit by funding the rehabilitation program with its money, without issuing the bond.
- In the presence of budgetary constraints, the issuer can identify a region for the bond’s notional and cap rate where the issuer can earn a profit by issuing the bond.
- The issuer should choose small values for the bond’s notional and cap rate to target investors with lower aversion to charity, and the issuer should choose higher notional and smaller cap rate to target investors with relatively low financial aversion to risk.
- Issuer can increase the combined financial and social impact of the bond, by increasing either the social impact or the financial impact, but not both.
- The issuer should spend very little on welfare to maximize the social impact of the bond, and it should increase its welfare spending to maximize its financial profit.

1.1. Peterborough Prison Social Impact Bond (Disley)

Although the valuation framework can be easily generalized to some of the other SIBs being implemented around the world, we focus on the Peterborough Prison Bond. This is the world’s first SIB issued by the UK government in 2010 to reduce recidivism rate of
the young male prisoners between the ages of 18 to 30, serving a short term sentence of 12 month or less in the Peterborough Prison. According to the estimates of the government of the UK, 60,000 adults per year receive sentences of 12 months or less. This makes up 10% of adult prison population. Short term prisoners make up 65% of all admissions and releases, and they have on average 16 previous convictions. According to the Social Exclusion Unit every re-offender costs the state a minimum of £143,000, and the usual reconviction rate for such offenders is 75% (Strickland).

In September 2010, the UK Ministry of Justice entered into a SIB mechanism for funding interventions for offenders at Peterborough Prison. Social Finance is acting as a financial intermediary, and the rehabilitation services will be provided by St. Giles Trust, a non-profit organization. Social Finance raised £5 million of investment funding from private individuals and charities, and the outcome payments will be made by the Ministry of Justice and the Big Lottery Fund if the reconviction rate drops by at least 10% for each cohort, or 7.5% overall. It is important to note that the Peterborough Prison Bond is not a fixed income instrument, unlike what the name may suggest. It pays absolutely nothing if the reconviction rate doesn't decrease by at least 10% in any one of the cohorts or by 7.5% across all cohorts. Hence, in terms of its payout it is more like call option.

Independent assessors from QinetiQ and the University of Leicester will determine the reconviction rate for prisoners who received interventions under the SIB. The offenders in each cohort will be compared to matched control groups. Each control group will be drawn from all prisoners released from sentences of less than 12 months, within the same time period from other prisons nationally. One-to-many propensity score-matching will be used to select the control group. This means that each cohort prisoner will be matched to up to 10 control group prisoners. According to an interviewee from the Ministry of Justice, the strength of using propensity score-matching (PSM) as an approach to building the control group is that it allows the assessor to control for different characteristics of the offenders.

The Peterborough SIB is a seven year project, and three cohorts of 1000 prisoners each will receive the rehabilitation services under this program. According to the Peterborough
prison statistics, it will take about two years to recruit a cohort of 1000 offenders. Each cohort will last until 1000 program participants have been released or two years have completed, whichever is earlier. Reconviction rate for each cohort is defined as the percentage of prisoners that are convicted for a crime in the first year after their release, and we will be using the terms reconviction rate and recidivism rate interchangeably. Measurement of the reconviction rate will include all prisoners released from the Peterborough Prison during this period, and will not be restricted to the prisoners that actually received the rehabilitation services in the SIB. This procedure is chosen to reduce service provider's incentive to cherry-pick those prisoners that are less likely to get reconvicted. Reconvictions will be measured over a period of one year. Hence, one year will be required to determine the reconviction rate after 1000 prisoners are released from the prison. So the first payment will be made about three years after the recruitment of first cohort i.e. in fall of 2013. While second and third payments, linked to the reconviction rate of the second and third cohorts, respectively, will be made approximately five and seven years after the launch of the project i.e. in 2015 and 2017 respectively.

For the ease of discussion, we will break down the payments into three smaller coupon payments made during 3rd, 5th, and 7th year of the program and a larger fourth coupon payment that would be made in the 7th year if none of the first three coupon payments are made. Each one of the first three coupons make a payment if the reconviction rate of the SIB cohort decreases by at least 10%, which is the minimum value for the drop in reconviction rate to be considered statistically significant for a cohort of thousand prisoners, compared to a similar matched group of prisoners. However, if none of the cohorts show a drop of at least 10%, all of the cohorts (3000 offenders) will be examined together at the end of the entire rehabilitation program, and if the average reconviction rate drops by at least 7.5%, a payout will be made. In order to cap the total cost to the government, payouts are capped from above, and investors’ payout do not increase after recidivism rate exceeds a certain limit.

Mathematically, these coupon payments can be expressed as
\[ C_1(z_1, k, z_k) \equiv k N z_1 1000 I(10\% \leq z_1 \leq z_k) + k N z_k 1000 I(z_1 > z_k) + 0 I(z_1 < 10\%), \]  
\( (1.1) \)

\[ C_2(z_2, k, z_k) \equiv k N z_2 1000 I(10\% \leq z_2 \leq z_k) + k N z_k 1000 I(z_2 > z_k) + 0 I(z_2 < 10\%), \]  
\( (1.2) \)

\[ C_3(z_3, k, z_k) \equiv k N z_3 1000 I(10\% \leq z_3 \leq z_k) + k N z_k 1000 I(z_3 > z_k) + 0 I(z_3 < 10\%), \]  
\( (1.3) \)

\[ C_4(z_1, z_2, z_3, k, z_k) \equiv k N \bar{z} 3000 I(z_1 < 10\%, \quad z_2 < 10\%, \quad z_3 < 10\%, \quad 7.5\% \leq \bar{z} \leq z_k) \]
\[ + k N z_k 3000 I(z_1 < 10\%, \quad z_2 < 10\%, \quad z_3 < 10\%, \quad \bar{z} > z_k) + 0 I(\bar{z} < 7.5\%). \]  
\( (1.4) \)

Here, \( k \) is the notional amount paid per unit decrease in reconviction rate, \( N \) is the number of bonds held by the investor, \( z_1, z_2, z_3 \) are the differences in reconviction rate between control group and the first, second and third cohorts respectively, \( z_k \) is the drop in reconviction rate after which payout remains fixed, \( \bar{z} \) is the average drop in reconviction rate in all three cohorts, and \( I(\cdot) \) is the indicator function.

The present value of the bond’s total payout can be written as

\[ B(z_1, z_2, z_3, k, z_k) = e^{-rT_1} C_1(z_1, k, z_k) + e^{-rT_2} C_2(z_1, k, z_k) + e^{-rT_3} C_3(z_3, k, z_k) + e^{-rT_3} C_4(z_1, z_2, z_3, k, z_k), \]  
\( (1.5) \)

where \( T_1, T_2, T_3 \) are the times for 1\textsuperscript{st}, 2\textsuperscript{nd}, and 3\textsuperscript{rd} payouts, and \( r \) is the risk free rate.
At this point, it is important to note that the Peterborough Bond is not free from shortcoming. The reconviction rate as defined for the Peterborough Bond creates some misplaced incentives. As described earlier, only the reconvictions that take place in the first year are included in the reconviction rate. This may incentivize the service provider to focus on prisoners' short term behavior and crimes that can get reconvictions from the courts quickly, and ignore the prisoners' long term behavior and crimes that may take longer to get reconvictions. Hence, the program may succeed in reducing the reconviction rate in the short term but may not change the prisoners' long term behavior. In addition, the Peterborough project may face internal frictions from different organs of the government. A privately funded prison rehab program that reduces the reconviction rate significantly may reduce for the regular prison staff, and hence may encounter friction. In addition, if investors are allowed to short the bond, then they would have an incentive to work against a drop in the reconviction rate, and may actually try to influence the outcome of the program.

1.2. Reference Parameters

Here, we present the default values for the parameters present in the bond payout function. For simplicity, we assume that the discount rate $r$ is the same for all payouts. Default values are given below:

\[
\begin{align*}
    k &= 0.01£ \\
    r &= 3\% \\
    \zeta_k &= 13\% \\
    T_1 &= 3\ Year \\
    T_2 &= 5\ Year \\
    T_3 &= 7\ Year \\
    N &= 1.
\end{align*}
\]  

(1.6)

The value of notional per bond per prisoner ($k$) is chosen such that each bondholder receives a coupon payment of $1£$ if the reconviction rate drops by 10%. The actual notional amount paid by the issuer of the bond can be different from it. However, any such difference will not affect our analyses because the number of bonds held by an investor can be adjusted to accommodate the difference. For example, If the actual value of $k$ is $1£$, then in our framework the investor will hold 100 bonds.
1.3. Yield to Maturity

The yield to maturity or the internal rate of return for this bond is the discount rate that makes the present value of all the future cash payments equal to the price paid for the bond. This can be easily computed, assuming the price paid for one bond is \( p \), as the \( y \) that satisfies (1.7),

\[
e^{-yT_1}C_1(z_1, k, z_k) + e^{-yT_2}C_2(z_1, k, z_k) + e^{-yT_3}C_3(z_3, k, z_k) + e^{-yT_4}C_4(z_1, z_2, z_3, k, z_k) - p = 0. \tag{1.7}
\]

Assuming recidivism among all three cohorts is decreased by the same amount so that \( z_1, z_2, z_3 \) are the same, yield to maturity for different notional amounts and prices paid is shown below, for \( z_k = 13\% \).

![Graphs showing yield to maturity](image)

**Figure 1:** The left panel shows the yield to maturity for different notional amounts, and the right panel shows the yield to maturity for different prices. Both plots assume that the reconviction rate drop by same percentage among all cohorts.

If change in recidivism \( z \) is smaller than 7.5\%, yield to maturity is undefined as investors don’t get any payout. For 7.5\% \( \leq z < 10\% \), yield to maturity increases linearly as the payout for the 4\textsuperscript{th} coupon increases linearly with \( z \). At \( z = 10\% \), payout for the 4\textsuperscript{th} coupon becomes zero, but the payout for the first three coupons become non-zero. Since the payments for the first three coupons has the same notional amount as the
4th coupon but happens sooner, so the bond’s yield takes a sudden jump, and then increases linearly for $10\% \leq z < 13\%$, as the payout for the three coupons increases linearly. After $z$ exceeds $z_k = 13\%$, yield stays constant as the payout doesn’t increase with increasing $z$. Moreover, yield to maturity curve is shifted upward by increasing $k$ as it increases the notional amount of the bond, which in turn increases every payout by an equal amount. Similarly, increasing $p$ shifts the curves downward.
Chapter 2

2. Modeling the Underlying Reconviction Rate

As discussed earlier, the bond’s payout is linked to the outcome of the rehab program, so we need a probability distribution for the reconviction rate of Peterborough Prison cohorts relative to the control group. This distribution is obtained using the results of past rehab programs from around the world. However, different rehab programs can be of very different natures, so we need a mechanism to adjust for these differences. Notably, the Peterborough prison rehab program is an integrated program that provides rehab services according to the needs of prisoners, while most of the past programs have been focused on one specific need. To adjust for the differences in programs, we calculate an overlap between the historical programs and the program implemented at Peterborough prison.

Since the Peterborough Prison rehab program provides services according to the needs of the prisoners, we cannot exactly determine its nature. However, the service provider for the Peterborough rehab program, St. Giles Trust, carried out a similar Through the Gates (TtG) program in 2009. The Peterborough program is likely to include similar rehab services that were provided in the TtG program. So we use the TtG program as a proxy for the Peterborough rehab program.

The overlap between the past programs and the TtG rehab program is then determined as the percentage of prisoners that received the same services in TtG program that were provided by the compared rehab program (see Appendix 1 for data). Hence, a weight is assigned to each program as determined by its overlap with TtG program. For example, approximately 12% offenders in TtG program received education and employment related services, so the rehab programs that focused exclusively on providing education and employment related services get a weight of 12%. For integrated programs that provide more than one service, a weight is calculated for each service area and then total weight is calculated by adding weights for all services provided. Weight for each bin is
calculated as the average weight for all programs that lie in that bin. These weights are then used to plot a weighted histogram.

However, the weights calculated here are only for illustration purposes, and a more detailed analysis of offenders’ demographics, past convictions, and severity of past crimes should be considered to determine more accurate weights for the rehab programs. Nevertheless, the method here provides enough flexibility to take into account differences between different rehab programs.

Weighted histogram for the choice of bins and weights shown in Appendix 1 is given below.

![Histogram of Reconviction Rate data. Negative (positive) values of the reconviction rate correspond to an increase (decrease) in the reconviction rate.](image-url)
2.1. Summary Statistics

Summary statistics for the outcomes of past rehab programs are shown in the table below.

| Moments | Central | | | | Non-central | | | |
|---------|--------| | | | Raw | Weighted | | | | Raw | Weighted |
| First   | 0.0969 | 0.0849 | | | 0.0969 | 0.0849 | | |
| Second  | 0.01028 | 0.01159 | | | 0.01967 | 0.01880 | | |
| Third   | 1.98 | 2.045 | | | 0.00596 | 0.00612 | | |
| Fourth  | 8.55 | 7.844 | | | 0.00237 | 0.00248 | | |

Table 1: Summary statistics for the raw data sets.

Raw and weighted moments differ by the choice of weights used in calculating moments. Raw moments assume that all rehab programs get identical weights, while the weighted moments use the weights specified in Appendix-1. It is clear from the table that the choice of weights makes a negligible difference on the values of moments.

2.2. Fitting Probability Distribution to Reconviction Data

After obtaining the histogram, we proceed to fit a probability distribution to this histogram. Since the difference between the reconviction rates of the Peterborough prison cohort and the control group has to lie between -100% and 100%, the beta distribution on the interval -1 to +1 is a natural choice. For the purposes of comparison, we also use the skew-normal distribution. Compared to the beta distribution, skew-normal distribution has one additional parameter, and hence provides more flexibility in fitting distribution to the raw data (as will be described in later sections).

The probability density function for the skew normal distribution is
where $\alpha$ is the skewness, $\mu$ is the mean, $\sigma$ is the scale parameter of skew-normal distribution, and $\phi(.)$ and $\Phi(.)$ are the standard normal probability and cumulative distribution functions, respectively. The density function for the beta distribution is

$$f_{\text{beta}}(x, \rho, \delta) = \frac{(1 + x)^{\rho-1}(1 - x)^{\delta-1}}{2^{\rho+\delta-1} \beta(\rho, \delta)},$$  

(2.2)

where $x \in [-1, 1]$, and $\beta(\rho, \delta)$ is the Beta function, defined as

$$\beta(\rho, \delta) = \int_0^1 (x)^{\rho-1}(1 - x)^{\delta-1} dx.$$  

(2.3)

After selecting the probability distributions, we move on to finding the right distribution parameters that fit distribution to the raw data. For the purposes of comparison, we choose two estimation methods: i) methods of moments, and ii) maximum likelihood.

### 2.2.1. Method of Moments

Appropriate skew-normal and beta distribution parameters can be determined using the method of moments. In the method of moments, we choose distribution parameters such that the moments of the fitted distribution match the sample moments (shown in Table 1).

For the skew-normal distribution, mean, variance, and skewness can be written in terms of the distribution parameters, and these equations can then be inverted to estimate the distribution parameters from the sample means (see Appendix 2). However, these inverted equations can only be used if the sample skewness is less than 1. However, for our data, sample skewness is 1.09. Therefore, parameter estimates are calculated numerically by equating the first three sample moments with the distribution moments, without inverting those equations. This leads to
for the skew normal distribution. For the beta distribution, we write the first two moments in terms of the distribution parameters, and then numerically solve for the parameters that make the distribution moments equal to the sample moments. This leads to

\[
\begin{cases}
\hat{\alpha} = 73.37 \\
\hat{\mu} = -0.0028 \\
\hat{\sigma} = 0.083,
\end{cases}
\]

(2.4)

2.2.2. Maximum Likelihood Parameter Estimation

Now we use the maximum likelihood principle to find the most likely parameters for skew normal and beta distributions. Most likely parameters are the ones that maximize the likelihood function for the observed dataset.

The likelihood functions for the skew normal distribution is

\[
L_{skew}(\alpha, \mu, \sigma) \equiv \prod_{i=1}^{n} f_{skew}(x_i)^{f(i)},
\]

(2.6)

where \( f_{skew} \) is the probability density function of the skew-normal distribution and \( f(i) \) is the weighted frequency for the \( i^{th} \) bin (given in Table 1 of Appendix 1). Similarly the likelihood function for the beta distribution is

\[
L_{beta}(\rho, \delta) \equiv \prod_{i=1}^{n} f_{beta}(x_i)^{f(i)},
\]

(2.7)

where \( f_{beta} \) is the probability density function of the beta distribution, and \( f(i) \) is the weighted frequency for the \( i^{th} \) bin.

For convenience, we work with the log likelihood function, and calculate parameters that maximize the log likelihood function for the skew normal and the beta distributions (calculations are shown in the Appendix 3). For the skew normal distribution, this yields

\[
\begin{cases}
\hat{\rho} = 45.92 \\
\hat{\delta} = 38.73.
\end{cases}
\]

(2.5)
Following the same procedure for beta distribution, we get

\[
\begin{align*}
\hat{\alpha} &= 3.69 \\
\hat{\mu} &= -0.002. \\
\hat{\sigma} &= 0.13.
\end{align*}
\] (2.8)

Comparing the most likely estimates with the method of moment estimates, parameter estimates do exhibit some dependency on the choice of method. However, as we will see later, the fair price of the bond is not very sensitive to the variation of this magnitude in parameter values. Plots for the skew-normal and the beta distribution are shown below.

**2.2.3. Comparison of Most Likely and Method of Moments Estimates**

The most likely estimate appears to be a better fit for both the skew-normal and the beta distribution. So, from here onward, we will stick to the maximum likely parameter estimates only.
2.3. Sensitivity Analysis for Distribution Parameters

The parameters estimated so far have been obtained by fitting the distributions to the weighted histogram of the raw data. However, as described earlier, the weights used in the weighted histogram are mainly for illustration purposes, and may or may not be a good representation of the actual overlap between the rehab programs. Therefore, we first analyze the sensitivity of the parameter estimates to the weights used for the rehab programs, and then produce a dataset by averaging over a range of values for the weight of each rehab program. In addition to the weights, we also treat the bin widths as a variable, and instead of sticking to one choice of bin widths, we average over a range choices for bins. In fact, the parameter estimates are more sensitive to the choice of bins than to the weights due to the fewer number of rehab programs. Including results of more rehab programs could reduce the sensitivity to both the bin width and the program weight.

To estimate the sensitivity of most likely parameters on the choice of bin widths and weights associated with the rehab programs, we let the weight of each type of rehab program vary from 5% to 95% in discrete steps of 30%, and choose four different sets of bins. Hence, we obtain 256 sets of weights, and 4 sets of bins. Then we calculate the maximum likelihood parameters for each one of the 256 sets of weights for four different choices of bins. This gives us a total of 1024 different sets of likelihood parameters. A more sophisticated Markov Chain Monte Carlo analysis that randomly selects weights and bin widths from a given set is also carried out, and yields results that are practically indistinguishable from the simpler analysis done by using fixed set of weights and bins. So we only present the results of the simplified analysis here.
Figure 4: A scatter plot of maximum likelihood parameters for the beta (left panel) and the skew normal (right panel) distributions. Color represents the value of log-likelihood function for the given value of parameters, and it increases from blue to red ( ). Parameters for beta distribution are much less scattered than the skew normal distribution.

Figure 5: Histograms for skewness (left panel), position (middle panel), and scale (right panel) of skew normal distribution.
Figure 6: Histograms for $\rho$ (left panel) and $\delta$ (right panel) parameters of the beta distribution.

Figures 4 shows the maximum likelihood parameters for the skew normal (right panel) and beta (left panel) distributions the histograms for these parameters are shown in Figures 5 (skew normal) and 6 (beta). The skew-normal distribution parameters exhibit a much larger variation in the parameter estimates than the beta distribution parameters. This is mainly because the skew-normal distribution has three parameters, while the beta distribution has only two. An additional parameter in the skew-normal distribution provides more flexibility for the fitting the distribution, and hence leads to larger variability in parameter estimates. Therefore, the fair price calculated using the skew normal distribution shows much larger variation as compared to the fair price calculated using the beta distribution.
Figure 7: Price histograms for the skew-normal (left panel) and the beta (right panel) distributions. Price is calculated using indifference pricing mechanism described later.

2.4. Predictive Data Set

Now we are in a position to generate a dataset, which is representative of a range of weights for each program and a range of choices for bins, by using the above calculated most likely parameters, and predicting a new data set using the Bayes principle,

\[
f(D^*|D) \equiv \int_{\text{All } \omega} f(D^*|\omega)f(\omega|D)d\omega, \tag{2.10}
\]

where \(D^* = \{x_i|x_i \in [-1,1]\}\) represents the predictive data set, \(D\) is the observed data set, and \(\omega\) is the vector of distribution parameters, \(f(\omega|D)\) is the joint probability density of \(\omega\) given dataset \(D\), and \(f(D^*|\omega)\) is the probability density of observing dataset \(D^*\) given the distribution parameters vector \(\omega\).

In short, we average over the bin widths and the program weights, and produce one representative dataset. The resulting predictive data sets for the beta and the skew normal distributions are shown in Figure 8.
Figure 8: Predictive probability density functions for the beta and the skew normal distribution.

After getting the predictive data set, we find the most likely parameters for this data set. Resulting most likely distribution are plotted in Figure 9.

Figure 9: Maximum likelihood distributions for predictive data set for the skew normal (left panel) and the beta (right panel) distributions.

The maximum likely parameters for the skew normal distribution turn out to be

\[
\begin{align*}
\alpha^* &= 6.56 \\
\mu^* &= -0.01 \\
\sigma^* &= 0.16
\end{align*}
\]  

(2.11)
and the most likely parameters for the beta distribution are

\[
\begin{align*}
\rho^* &= 66.00 \\
\delta^* &= 55.34
\end{align*}
\]  \hspace{1cm} (2.12)

From now on we will use these parameters as the reference parameters, and all probabilities will be calculated using these parameters, unless otherwise specified.

We now use the asymptotic normal approximation to get an estimate for the confidence intervals of the most likely parameters. In the asymptotic normal approximation, confidence interval for a parameter $\tau$ is

\[
\tau^* \pm \frac{1.96}{\sqrt{-l''(\tau^*, x)}}
\]  \hspace{1cm} (2.13)

where $l''(\tau^*, x)$ is the second derivative of the log likelihood function with respect to $\tau$ evaluated at $\tau^*$. Hence, the 95% confidence intervals for the skew normal distribution are

\[
\begin{align*}
\alpha: [1.356, 6.024] \\
\mu: [-0.036, 0.032] \\
\sigma: [0.117, 0.143]
\end{align*}
\]  \hspace{1cm} (2.14)

Similarly, 95% confidence intervals for the beta distribution are

\[
\begin{align*}
\rho: [56.734, 63.767] \\
\delta: [47.104, 52.936]
\end{align*}
\]  \hspace{1cm} (2.15)
Figure 10: Comparison of the most likely distributions for raw and predicted data sets for the skew normal (left panel) and the beta (right panel) distributions.

Figure 10 compares the maximum likelihood skew-normal and beta distributions for the predicted dataset with the maximum likely distribution for the raw dataset. It is clear from the figure that the skew-normal distribution exhibits a larger change than the beta distribution. This is because the skew-normal distribution has one additional parameter, and is more flexible than the beta distribution. However, the change is not very significant even for the skew normal distribution. This indicates that as long as the skew normal and the beta distributions are good fits for the rehab program data, our choice of reference bin widths and the program weights are a good representation of the raw data.

Comparison of the summary statistics for the original data, and the most likely distributions is shown in table 2. It is clear from this comparison that the fitted distributions have much smaller kurtosis than the original data, and hence underestimate the probability of extreme outcomes.
<table>
<thead>
<tr>
<th>Statistic</th>
<th>Original Data</th>
<th>Skew Normal</th>
<th>Beta</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>9.69%</td>
<td>11.69%</td>
<td>8.78%</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>10.14%</td>
<td>10.02%</td>
<td>9.01%</td>
</tr>
<tr>
<td>Skewness</td>
<td>1.98</td>
<td>0.9071</td>
<td>-0.0316</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>8.55</td>
<td>3.767990</td>
<td>2.953235</td>
</tr>
</tbody>
</table>

Table 2: Comparison of summary statistics. 95% confidence interval for the mean is [6.27%, 13.12%].
Chapter 3

3. Pricing Peterborough Prison SIB

As described in section 1.1, the payout of the Peterborough Prison SIB is linked to the reconviction rate, which cannot be traded. Moreover, the bond is not publicly traded, so it is not possible to short the bond. Hence, the cash flows of the bond cannot be replicated. It is, therefore, an example of an incomplete market derivative that cannot be hedged, and Arbitrage Pricing theory cannot be used to determine a unique price for the bond.

Therefore, we have to use incomplete market pricing mechanisms to adjust for the uncertainties associated with the bond’s cash flows. Here, we present two such mechanisms: the Wang transform, and indifference pricing.

3.1. Pricing the SIB using Wang Transform (Wang)

The Wang transform is an example of a distortion method that transforms one probability measure to another and was proposed by Wang as a unified pricing mechanism for both insurance and financial risks. It can also be seen as a minimum relative entropy distribution that satisfies a certain moment constraint (Reesor). The Wang transform reproduces the Black-Scholes option pricing formula, and the Capital Asset Pricing Model for determining rate of return for a risky asset, if the asset returns are normally distributed.

The transformed probability distribution under the Wang transform is given by

\[ F_{\lambda}(x, \omega, \lambda) \equiv \Phi[\Phi^{-1}(F(x, \omega)) + \lambda], \]

and the corresponding transformed density is

\[ f_{\lambda}(x, \omega, \lambda) = \frac{\phi[\Phi^{-1}(F(x, \omega)) + \lambda]}{\phi[\Phi^{-1}(F(x, \omega))]} f(x, \omega), \]
where $\lambda$ is the risk aversion parameter, $\omega$ is the vector of distribution parameters, and $F_\lambda$ and $f_\lambda$ are risk adjusted cumulative distribution and density functions respectively.

To implement the Wang transform, we first compute the CDF’s of the skew normal, and the beta distributions. For the skew normal,

$$F_{skew}(x, \alpha, \mu, \sigma) \equiv \Phi \left( \frac{x - \mu}{\sigma} \right) - 2 T \left( \frac{x - \mu}{\sigma}, \alpha \right),$$

(3.3)

where $T(\cdot, \cdot)$ is the Owen’s T function defined as

$$T(x, \alpha) \equiv \frac{1}{2\pi} \int_0^\alpha \frac{e^{-x^2(1+t^2)/2}}{1 + t^2} \, dt,$$

(3.4)

and $\Phi(\cdot)$ is the cumulative density function of the standard normal distribution.

The cumulative distribution function for the beta distribution is

$$F_{beta}(x, \rho, \delta) \equiv \frac{\beta_1 \left( \frac{1 + x}{2}, \rho, \delta \right)}{\beta(\rho, \delta)},$$

(3.5)

where $\beta(\rho, \delta)$ is the Beta function defined in equation (2.3), and $\beta_1(\cdot, \cdot, \cdot)$ is the incomplete Beta function defined by

$$\beta_1 \left( \frac{1 + x}{2}, \rho, \delta \right) = \int_0^{\frac{1 + x}{2}} (z)^{\rho - 1} (1 - z)^{\delta - 1} \, dz.$$

(3.6)

Figure 11 shows how the risk adjustment affects the skew normal and the beta probability densities.
To better understand the effect of the Wang transform on the probability distribution, we plot the changes in mean, standard deviation, skewness and kurtosis with the corresponding changes in the Wang risk aversion parameter.

Figure 12 indicates that a positive change in the Wang risk aversion parameter shifts the skew normal distribution backward (lower mean), reduces standard deviation, but increases its skewness and kurtosis. So the overall effect is that the investor with higher risk aversion parameter decreases the probability of smaller and more probable gains, while increases the probability of small and more probable losses, and also increases the probability of higher and less probable gains. For the beta distribution, mean and skewness decrease, while standard deviation and kurtosis remain almost unchanged. The overall effect on the investor’s risk preferences is similar. Therefore, increasing risk aversion parameter in the beta distribution makes investor more risk averse than the skew normal distribution, as there is no increase in skewness or kurtosis for the beta distribution.

Figure 11: Risk adjusted skew normal (left) and beta (right) density functions for different values of risk aversion parameter.
Figure 12: Effect of the Wang risk aversion parameter on distribution statistics. Changes in the distribution mean, standard deviation, skewness and kurtosis are shown in top left, top right, bottom left and bottom right panels, respectively, for both the skew normal and the beta distributions.

Since all the distribution parameters change almost linearly (can be seen in Figure 12) with the Wang risk aversion parameter for small changes in the risk aversion parameter around its reference value, we sum up the effect by calculating the slope of these lines. Slopes are shown in table 3.
<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Skewness</th>
<th>Kurtosis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Skew Normal</td>
<td>-0.097</td>
<td>-0.071</td>
<td>0.055</td>
<td>0.040</td>
</tr>
<tr>
<td>Beta</td>
<td>-0.127</td>
<td>0.001</td>
<td>-0.146</td>
<td>-0.0001</td>
</tr>
</tbody>
</table>

**Table 3: Sensitivity of the skew normal and the beta distribution parameters on the Wang risk aversion parameter.**

This table shows that the beta distribution shows higher negative sensitivity to the change in the Wang risk aversion parameter. Therefore, we expect the beta distribution price to decrease more sharply with increase in the Wang risk aversion parameter.

The Wang transform can be extended to the multivariate distributions, to yield (Kijima)

\[
F_{\lambda}(\mathbf{x}, \theta) \equiv \Phi_{n} \left[ \phi^{-1}(F_{1}[x_{1}, \theta]) + \sum_{j=1}^{n} \lambda_{j}\rho_{1j}, ..., \phi^{-1}(F_{n}[x_{n}, \theta]) + \sum_{j=1}^{n} \lambda_{j}\rho_{nj} \right], \tag{3.7}
\]

where \( \Phi_{n} \) is the multivariate normal distribution, \( F_{1}, ..., F_{n} \) are the marginal CDFs, and \( \lambda_{1}, ..., \lambda_{n} \) are the risk aversion parameters, and \( \rho_{ij} \) is the correlation between \( x_{i} \) and \( x_{j} \).

To simplify calculations, we assume that the reconviction rates for all three cohorts of the Peterborough Prison are uncorrelated with each other. This implies that \( \rho_{ij} \) is equal to zero for all \( i \neq j \), and \( \rho_{ij} \) is equal to 1 for \( i = j \). Hence, the equation (3.7) simplifies to:

\[
F_{\lambda}(\mathbf{x}, \theta) \equiv \Phi_{1}[\phi^{-1}(F_{1}[x_{1}, \theta]) + \lambda_{1}] ... \Phi_{n}[\phi^{-1}(F_{n}[x_{n}, \theta]) + \lambda_{n}]. \tag{3.8}
\]

This shows that the risk adjusted joint distribution can be written as a product of the risk adjusted marginal distributions. Under this assumption of independence, the first three payouts can be priced using the marginal distributions for the reconviction rates of the first three cohorts. This leads to
where \( T_1, T_2, T_3 \) are the times when payments for first, second and third cohorts are made.

Assuming that \( z_1, z_2, z_3 \) are independent and identically distributed random variables, we can then write equation 3.12 as

\[
P_4(k, z_k, \lambda) = e^{-rT_3} \int_{-1}^{1} C_4(z_1, z_2, z_3, z, k) f_{z_1}(z_1, \alpha, \mu, \sigma) f_{z_2}(z_2, \alpha, \mu, \sigma) f_{z_3}(z_3, \alpha, \mu, \sigma) dz_1 dz_2 dz_3, \tag{3.13}
\]

where \( \mathbf{z} \) is the vector of reconviction rates \( z_1, z_2, z_3 \), and the bond price then is

\[
P(k, z_k, \lambda) = P_1(k, z_k, \lambda) + P_2(k, z_k, \lambda) + P_3(k, z_k, \lambda) + P_4(k, z_k, \lambda). \tag{3.14}
\]
Figure 13: Plot in the left panel shows the bond price with risk aversion parameter ($\lambda$), and plot in the right panel shows bond price with interest rate ($r$).

Variation of price with the risk aversion parameter ($\lambda$), and interest rate ($r$) is shown in the Figure 13 above. Both plots in Figure 13 exhibit mild convexity. We compute the first and second derivative of the bond price with respect to both $\lambda$ and $r$ to get a quantitative estimate of the bond’s sensitivity to these parameters.

<table>
<thead>
<tr>
<th>Distribution</th>
<th>$Slope_r$</th>
<th>$Convexity_r$</th>
<th>$Slope_\lambda$</th>
<th>$Convexity_\lambda$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta Distribution</td>
<td>4.95</td>
<td>13.57</td>
<td>9.70</td>
<td>12.85</td>
</tr>
<tr>
<td>Skew Normal Distribution</td>
<td>4.95</td>
<td>13.57</td>
<td>8.92</td>
<td>6.62</td>
</tr>
</tbody>
</table>

Table 4: Slope and convexity of price curves for 1% change in the values of interest rate and risk aversion parameter for $\lambda = 0.1$, $r = 3\%$, $\Delta \lambda = 0.01$ and $\Delta r = 0.01$. $P_+$ refers to the price at lower $\lambda/r$, and $P_-$ refers to the price at higher $\lambda/r$. Hence, the slope is defined to be the negative of the actual slope.
Table 4 shows that, for a given level of interest rate and risk aversion, the price computed with the beta distribution is more sensitive to changes in risk aversion parameter than the skew normal distribution, while both the skew normal and the beta distribution prices have the same sensitivity with interest rate. This higher sensitivity of the beta distribution price on risk aversion parameter ($\lambda$) is consistent with the observations of Figure 12. For the skew normal distribution, increasing $\lambda$ increases skewness and kurtosis, and offsets some of the price reduction coming from decreasing mean. However, no such increase in skewness and kurtosis is observed for the beta distribution, and hence the price decline is steeper.

Next, we look at the sensitivity of the bond price to the distribution parameters. The price of the bond with the skew normal distribution parameters is shown in the left panel of Figure 14. The price decreases linearly with increasing the value of position parameter ($\mu$), as increasing $\mu$ shifts the distribution forward and increases the probability of reduction in reconviction rate. While price appears almost insensitive to the skewness parameter ($\alpha$), though it actually increases with increasing $\alpha$ as more skewed distribution increases the probability of events in the right tail. However, skewness causes much smaller change in price than the position parameter, and makes the change invisible on the graph. Therefore, a small decrease in the position parameter can offset a large change in the skewness parameter. An example of this result has been observed in Section 2.3 during the calculation of sensitivity of distribution parameters to the rehab program weights and bins. While changing the choice of weights and bins affect skewness parameter significantly, these changes are mostly offset by much smaller changes in the position parameter, and the price does not change significantly.
Figure 14: The effect of risk adjustment on the bond price. $\lambda = 0$ corresponds to no risk adjustment, and the price there is just the expected present value of all the payouts.

For the beta distribution, the price’s dependence on each parameter is almost a mirror image of its dependence on the other parameter. This is expected because each parameter’s role in the probability distribution is a mirror image of the other distribution. If we change $z$ to $-z$ in the probability density function of the beta distribution, and switch $\rho$ with $\delta$, the probability density function remains unchanged.

### 3.2. Indifference Pricing

We now use the indifference pricing principle to calculate the fair bond price. We assume that investor’s utility is exponential. This assumption allows us to write down an explicit function for the bond price. Specifically,

$$u(w, \gamma): = \frac{-1}{\gamma} (e^{-\gamma w} - 1),$$

(3.15)

where $w$ is the wealth and $\gamma$ is the risk aversion parameter for the investor.
Now the utility of having the initial wealth $w$ is

$$u_w(w, \gamma) = u(w, \gamma), \quad (3.16)$$

while the utility of investing in the risky bond (assuming $z_1, z_2, z_3$ are i.i.d) is:

$$u_{\text{bond}}(w, p, \gamma) = E\left[u(w - p + B(z_1, z_2, z_3, k, z_k))\right]$$

$$= \int_{-1}^{1} u(w - p) + B(z_1, z_2, z_3, k, z_k) f(z_1) f(z_2) f(z_3) dz_1 dz_2 dz_3, \quad (3.17)$$

where $p$ is the price paid for the bond, $B$ is the payout function of the bond, and $f(\cdot)$ is the probability density function for the drop in reconviction rate.

The fair price of the bond is then calculated as the price that makes the utility of holding an initial wealth $w$, given by $u_w$, equal to the expected utility of investing in the risky asset, given by $u_{\text{bond}}$. This leads to:

$$p(k, z_k, \gamma) = -\frac{1}{\gamma} \ln \left( \int_{-1}^{1} e^{-\gamma B(z_1, z_2, z_3, k, z_k)} f(z_1) f(z_2) f(z_3) dz_1 dz_2 dz_3 \right). \quad (3.18)$$

Variation of the fair price with the risk aversion parameter is shown in Figure 15 below.

Having obtained the fair price function, we first see which value of the risk aversion parameter corresponds to the risk neutral price of the bond. Since a risk neutral investor has a linear utility function, its risk aversion parameter would be zero, as it would make the utility function equal to the investor's wealth $w$. This can be easily seen by expanding the investor's utility function given in equation (3.15). That is, we start by expanding the exponential function

$$u(w) = -\frac{1}{\gamma} (e^{-\gamma w} - 1) = -\frac{1}{\gamma} (1 - \gamma w + (\gamma w)^2 + \cdots - 1) \quad (3.19)$$
\[ u(w) = w - \gamma w^2 + \cdots, \]  
(3.20)

and for \( \gamma = 0 \), we get

\[ u(w) = w, \]  
(3.21)

Hence the risk-neutral value of the risk aversion parameter is \( \gamma^* = 0 \). Using equation (3.18), we see that the right hand side of equation (3.21) is simply the risk adjusted price, and hence the risk neutral price is simply the present value of the future payout

\[ p(k, z_k, \gamma^*) = E[B(z_1, z_2, z_3, k, z_k)]. \]  
(3.22)

Figure 15: Dotted lines indicate the expected present value of the bond, and the solid lines show the risk adjusted price of the bond. Both curves have been drawn for the reference parameters defined in equations (2.11) and (2.12).
Chapter 4

4. Connection between the Wang and Utility Formulations

Having calculated the risk adjusted price in two different ways, we compare the key results of the two methods. First, we obtain a relationship between the Wang risk aversion parameter and the utility risk aversion parameter, such that we get the same price in both formulations. We start with a value of the Wang risk aversion parameter, calculate the corresponding risk adjusted price, and then find the utility risk aversion parameter that leads to the same price. The curve thus obtained is plotted in the Figure 16.

Now we compare the bond’s sensitivity to the risk aversion parameter and the interest rate under the two pricing mechanisms for both the skew normal and the beta distributions. For that purpose, we choose $\lambda = 1$. This corresponds to $\gamma_{beta} = 4.34$, and $\gamma_{skew} = 4.10$. Then we plot the price calculated using the Wang transform and the utility formulation for small changes in interest rate and risk aversion levels around their reference values. Results are shown in Figure 17 and 18.
Figure 16: Utility risk aversion parameter plotted with the Wang risk aversion parameter for both the skew normal and the beta distributions using the reference parameters.
Figure 17: Variation in the fair price with changes in interest rate (left panel) and risk aversion (right panel) for the skew normal distribution.

Figure 18: Variation in the fair price with changes in interest rate (risk aversion) and risk aversion (right panel) for the beta distribution.

These figures show that the price calculated using the Wang distortion measure is more sensitive to changes in both interest and risk aversion, for both the skew normal and the
beta distributions. Moreover, the percentage change in price with either interest rate or the risk aversion parameter is higher in the case of the skew normal distribution than in the case of the beta distribution. This also confirms the result obtained in Table 3, which showed that the beta distribution is more sensitive to the changes in the Wang risk aversion parameters, as compared to the skew normal distribution.
Chapter 5

5. Pricing under Ambiguity Aversion

As described in Section 2.1, the reference model for the probability of a reduction in reconviction rate is based on a relatively small dataset, and may not be a very accurate representation of the outcome of the Peterborough Prison rehab program. Moreover, this lack of quantitative data would be an essential feature of any such SIB program, because the outcomes of prison rehabilitation programs have historically not been measured very rigorously. Therefore, any pricing mechanism for such bonds has to take into account this deficiency in order to determine the fair price.

Here, we incorporate this lack of data as ambiguity in the reference model, and treat it as a source of uncertainty, separate from the risk of the bond’s payout for a given model. To adjust for this ambiguity, we allow investors to consider probability distributions that are different from the reference distribution at the cost of paying a penalty. In other words, the investor is allowed to consider models other than the reference model at the cost of increasing its utility for a given model. We choose the relative entropy between the relative model and the candidate model as the penalty term (Jaimungal)

\[
h(Q'|Q) = \theta E_Q \left[ \frac{dQ'}{dQ} \ln \left( \frac{dQ'}{dQ} \right) \right],
\]

(5.1)

where \( \frac{dQ'}{dQ} \) is the Radon-Nykodym derivative of \( Q' \) with respect to \( Q \), \( \theta \) is a measure of faith in the reference model \( (Q) \) and \( E_Q \) represents expected value under model \( Q \). Radon-Nykodym derivative \( \frac{dQ'}{dQ} \) is

\[
\frac{dQ'}{dQ} = \frac{f_{Q'}(X)}{f_{Q}(X)}
\]

(5.2)

where \( f_{Q'}(X) \) is the probability density function of the candidate measure evaluated at \( X \) and \( f_{Q}(X) \) is the probability density function of the reference model evaluated at \( X \). Hence, the penalty function for the skew normal and the beta distributions become:
\[ h_{\text{skew}}(\alpha, \mu, \sigma) \equiv \int_{-\infty}^{\infty} f_{\text{skew}}(x, \alpha, \mu, \sigma) \times \ln \left( \frac{f_{\text{skew}}(x, \alpha, \mu, \sigma)}{f_{*}\text{skew}(x)} \right) dx, \quad (5.3) \]

\[ h_{\beta\text{eta}}(\rho, \delta) \equiv \int_{-1}^{1} f_{\beta\text{eta}}(x, \rho, \delta) \times \ln \left( \frac{f_{\beta\text{eta}}(x, \rho, \delta)}{f_{*}\beta\text{eta}(x)} \right) dx \quad (5.4) \]

where \( f \) and \( f^* \) are the PDFs of candidate and reference models respectively.

In the coming sections, we see how the ambiguity adjustment affects the bond price, and compare these effects for every choice of probability distribution and the pricing mechanism. Each subsection shows the calculation of the ambiguity adjusted price for a given choice of pricing mechanism and probability distribution, and concludes with a comparison of prices obtained under different methodologies.

### 5.1. Robust Indifference Pricing

To include the ambiguity aversion in the indifference pricing mechanism, we add the penalty term defined in equation (5.3) to the expected utility given in equation (3.17). This leads to

\[
 u_{\text{bond}}(w, p, \gamma, \theta) \\
 \equiv \theta \ h(\omega) \\
 + \left( \prod_{-1}^{1} u[w - p + B(z, k, w, z_k), \gamma] \ f(z_1)f(z_2)f(z_3)dz \right) \quad (5.5)
\]

where \( \omega \) is the set of distribution parameters, \( B(., \ldots, .) \) is the payout function for the bond, and \( f(.) \) is the probability density function for the drop in reconviction rate.

The fair price is then calculated as the minimum price that equates the penalized expected utility to the utility of investing in the risk free asset. With the addition of the penalty term, price function becomes
Comparing equation (5.6) with (3.18), we see that the effect of the penalty term is to add a $\theta$ dependent term in the price function. This term, however also depends on the investor’s initial wealth ($w$). While, in reality, an investor’s ambiguity aversion may depend on the initial wealth of the investor, we would like to remove this dependence to be able to study the effect of ambiguity in the reference model alone. Therefore we replace $\theta$ by $\theta e^{-yw}$ in equation (5.6). The modified price function then becomes

$$p(k, z_k, \gamma) = -\frac{1}{\gamma} \ln \left( \prod_{-1}^{1} e^{-\gamma B(z_1, z_2, x_3, k, x_k)} f(z_1) f(z_2) f(z_3) dz_1 dz_2 dz_3 \right)$$

$$+ \frac{1}{\gamma} \ln \left( 1 + \theta e^{yw} h(\omega) \right).$$

(5.7)

Since the penalty term is always positive by definition, the effect of the penalty term in the price function is to increase the price for a given level of risk aversion and a given set of distribution parameters. At this point, it suggests that increasing the ambiguity aversion increases the price. But it is not the case, as the actual price under ambiguity aversion is calculated for that model that minimizes the price function in equation (5.6). Hence, the ambiguity adjusted price cannot exceed the risk adjusted price. Complete formula for the price function can be written as

$$p(k, z_k, \gamma, \theta) = \min_{a,\mu} \left[ -\frac{1}{\gamma} \ln \left( \prod_{-1}^{1} e^{-\gamma B(z_1, z_2, x_3, k, x_k)} f(z_1) f(z_2) f(z_3) dz_1 dz_2 dz_3 \right) \right.$$

$$\left. + \frac{1}{\gamma} \ln \left( 1 + \theta h(\omega) \right) \right],$$

(5.8)
while the corresponding risk adjusted price is given in equation (3.18).

Therefore, while (5.7) has an additional positive term compared to the risk adjusted price given in equation (3.18), the freedom to choose $\alpha, \mu$ different from the reference parameters $(\alpha^*, \mu^*)$ can make the ambiguity adjusted price smaller than the risk adjusted price.

5.1.1. Robust Indifference Pricing for the Skew Normal Distribution

Using equation (5.8), we can write down the ambiguity adjusted price for the skew normal distribution simply by replacing $h(\omega)$ with $h_{\text{skew}}(\alpha, \mu, \sigma)$. Hence the ambiguity adjusted skew normal price is

$$p(k, z, \gamma, \theta) = \min_{\alpha, \mu} \left[ -\frac{1}{\gamma} \ln \left( \int_{-1}^{1} e^{-\gamma \beta(z_1, z_2, z_3, k, x_k)} f(z_1) f(z_2) f(z_3) dz_1 dz_2 dz_3 \right) 
+ \frac{1}{\gamma} \ln (1 + \theta h_{\text{skew}}(\alpha, \mu, \sigma)) \right].$$

(5.9)

Before calculating the actual ambiguity adjusted price (given in equation (5.9)), we first plot the price function given in equation (5.7) to develop some intuition about the penalty function. Plots of this price function with respect to the model parameters are shown in Figures 19 and 20. Figure 19 shows the variation of the ambiguity adjusted price versus $\theta$ for different values of skewness and position parameters of the candidate model. The figures show that the slope of the fair price is almost independent of $\theta$, but increases as the candidate measure deviates away from the reference measure in either direction. This is an expected result because any deviation from the reference model increases the value of the penalty function $h_{\text{skew}}(\alpha, \mu, \sigma)$, which is the coefficient of $\theta$ in equation (5.7), and hence increases the derivative of price with respect to theta.

It is also clear from the Figure 19 that the price curves in the left panel exhibit a higher a slope than the curves in the right panel. Since the only difference between the curves in
the left and right panels is that the curves in the left panels are evaluated at $\mu^*$, while the curves in the right panel are evaluated at $\alpha^*$. Therefore, we can conclude that the penalty function is more sensitive to changes in the skewness parameter than changes in the position parameter.

**Figure 19:** The fair price of the bond for different values of skewness and position parameters is shown. Price increases with increasing $\theta$, all else held fixed. Curves in the left panel are plotted by holding the position parameter ($\mu$) fixed at $\mu = \mu_{\text{ref}}$ and the curves in the right panel are plotted by holding the skewness parameter ($\alpha$) fixed at $\alpha = \alpha_{\text{ref}}$. For the right panel, $\eta = \mu$ and $\eta_{\text{ref}} = \mu_{\text{ref}}$.

In order to further confirm this intuition, we plot the price with the skewness and the position parameter in Figure 20. Here, the solid curves represent the total price, and the dotted curves represent the ambiguity component (total price less the risk only price) of the total price. Dotted price curve in the left panel of Figure 20 (almost identical to the solid curve) shows a much larger variation in price than the curve in the right panel of Figure 20, confirming that the penalty function is more sensitive to the variation in the skewness parameter. Moreover, these plots also show that increasing $\theta$ shifts the price curves upward, reducing the region where price lies below the reference price. Hence, for very large values of $\theta$ the price cannot be minimized below the reference price, and
ambiguity adjusted price would equal the risk adjusted price. Therefore, we interpret $\theta$ as the faith in reference model. The larger the $\theta$, more confident the investor is about the reference model, and closer the ambiguity adjusted price is to the risk adjusted price.

Figure 20: Variation of the fair price plot with the skewness parameter (left panel) and the position parameter (right panel) for $\gamma = 0.2$. Dotted lines in both panels have been moved up, by an amount equal to the reference price, to be visible on the plot.

Now, we move on to the actual ambiguity adjusted price given in equation (5.8). Based on the results obtained so far, we can see that changing the skewness parameter in any direction about its reference level increases the price, and increasing position parameter above its reference level increases the price function. Therefore the minimum price is expected to be achieved by decreasing the position parameter below its reference level, and by keeping the skewness parameter at or above its reference level.

A surface plot of the fair price with skewness and position parameters is shown in Figure 21. The minimum level of this surface plot is the ambiguity adjusted for a given $\theta$ (faith in reference model). A contour plot of the price function is shown in the right panel of the Figure 21. It shows that the price lies in a triangular region in the lower right corner of the plot.
Figure 21: Surface (left panel) and contour (right panel) plots with the skew normal distribution parameters are shown. For both panels, $\gamma = 1$ and $\theta = 5$.

Minimizing the price function with respect to both $\alpha$ and $\mu$ simultaneously for various values of $\theta$ yields a three dimensional plot of optimal values of the skew normal parameters for each value of $\theta$. These parameters are plotted in the left panel of Figure 22, and it can be seen that the parameters approach the reference parameters for higher values of $\theta$ (defined in equation (2.11)). This is because increasing $\theta$ increases the relative effect of the ambiguity term in the price function, and for sufficiently large values of $\theta$, minimizing the price function is equivalent to minimizing the penalty function, which is minimized when candidate model is the same as the reference model. Hence, for sufficiently large $\theta$, the optimal ambiguity adjusted model is simply the reference model. This result is also consistent with our interpretation of $\theta$ as the faith in reference model.
Figure 22: The left panel shows that with increasing faith in the reference model, candidate model approaches the reference model and two prices get closer. The right panel shows the ambiguity and risk adjusted prices with the faith in reference model. For both panels, $\gamma = 1$.

Finally, the ambiguity adjusted price as a function of $\theta$ is shown in the right panel of Figure 22. For small values of $\theta$, the investor has little faith in the reference model, and would pay only a small price due to high ambiguity. As $\theta$ increases, the investor puts more faith in the reference model and the resulting decrease in ambiguity allows her to pay a higher price. For infinitely large values of $\theta$, the investor has full faith in the reference model, and the ambiguity adjusted price approaches the price obtained by reference model.

In the end, we compare the ambiguity adjusted price curve with the risk adjusted price curve evaluated at the optimal parameters calculated by minimizing the penalty term in the ambiguity function (shown in the left panel of Figure 23). The first thing we notice is that the ambiguity adjusted price is not the same as the risk adjusted price evaluated at the optimal parameters. The ambiguity adjusted price (blue curve) is slightly higher than the risk adjusted price (green curve) for small $\theta$, but both price curves (blue and black curves) approach the same asymptotic limit (red curve) for large $\theta$. Therefore, we can
conclude that the effect of ambiguity adjustment cannot be completely modeled by changing the candidate model and making a constant shift in the level of risk aversion in the risk adjusted price function. The difference between the ambiguity adjusted price and the risk only price is shown in the right panel of Figure 23, and it exhibits a sharp decline as $\theta$ increases.

Figure 23: The left panel shows the ambiguity adjusted price (blue curve) and the risk adjusted price evaluated at the optimal parameters (green curve) as a function of $\theta$. All curves in the left panel are plotted using the same level of risk aversion. The right panel shows the difference between ambiguity adjusted price and the risk adjusted price evaluated at the optimal parameters. This difference can be seen as the effect of ambiguity aversion that cannot be incorporated with the risk aversion only.

Figure 24 decomposes the difference between the ambiguity adjusted price and the reference risk adjusted price into the contribution from change in the optimal parameters.
and the contribution from the decrease in risk aversion. The red curve shows the change in price that results from a change in parameters of the model, and the blue curve shows the change in price that comes from decrease in risk aversion level. The difference between the two curves is always equal to hundred percent. The ratio of the price change attributed to the change in parameters to the price change attributed to the change in risk aversion is shown in the panel of the Figure 24. Whether this ratio increases or decreases depends on how quickly the risk adjusted price approaches the ambiguity adjusted price. If the risk adjusted price quickly approaches the ambiguity adjusted price, the denominator shrinks faster than the numerator, and the ratio increases. However, if the risk adjusted price approaches the ambiguity adjusted price very slowly, the denominator stays almost constant and the ratio decreases. From Figure 24, we can conclude that the change in parameters plays a less important role in determining the ambiguity adjusted price for higher values of \( \theta \), and that the risk adjusted price approaches the ambiguity adjusted price slowly.

![Components of Ambiguity Adjusted Skew Normal Price](image1)

![Ratio of the Two Components of Ambiguity Adjustment](image2)

**Figure 24:** Difference between the ambiguity adjusted price and the reference risk price is decomposed into the contribution from a change in model parameters (red curve) and a contribution from decrease in risk aversion level (blue curve) is shown in the left panel. The right panel shows the ratio of the red curve to the blue curve.
5.1.2. Robust Indifference Pricing for Beta Distribution

Now we follow the same procedure to calculate ambiguity adjusted price for the beta distribution. The ambiguity adjusted price for the beta distribution can be obtained by following the same steps as above. First, after the addition of penalty term, the expected utility of investing in the bond is given by

\[ u_{\text{bond}}(w, p, \gamma, \theta) \equiv \theta h_{\text{beta}}(\rho, \delta) + \int_{-1}^{1} u[(w - p) + B(z_1, z_2, z_3, k, z_k), \gamma] f(z_1)f(z_2)f(z_3)dz_1dz_2dz_3, \]

where \( h_{\text{beta}} \) is the penalty term for beta distribution defined in (4.4).

The fair price of the bond is then obtained by equating the utility given in (5.10) with the utility of having initial wealth \( w \). The fair price is shown below

\[ p(k, z_k, \gamma) = \min_{\rho, \delta} [\frac{1}{\gamma} \ln \left( \int_{-1}^{1} e^{-\gamma B(z_1, z_2, z_3, k, z_k)} f(z_1)f(z_2)f(z_3)dz_1dz_2dz_3 \right) + \frac{1}{\gamma} \ln(1 + \theta h_{\text{beta}}(\rho, \delta))]. \]

Now we look at the variation of the price function (price before minimizing over the distribution parameters) with the distribution parameters \( (\rho, \delta) \), and it is shown in Figures 25 and 26. We notice from Figure 25 that as \( \rho \) and \( \delta \) move away from their reference values, the price becomes more sensitive to theta. This is, as described in the previous section, due to the fact that the slope of the price with theta is proportional to the penalty function, and the penalty function increases as parameters move farther away from their reference values.
Figure 25: Variation of the price function with the faith in reference model for the beta distribution parameters $\rho$ (left panel) and $\delta$ (right panel) is shown.

Figure 26 shows the variation of the price function with the parameters of the beta distribution. We notice that for $\theta = 0$, the price curve with $\delta$ is a mirror image of the price curve with rho, an outcome of the symmetry of the underlying beta distribution (with $\rho - \delta$). The ambiguity component appears to be less sensitive to the changes in parameters above their reference values than the changes in parameters below their reference values. This tiny asymmetry in the ambiguity component makes the total price curve plotted against $\rho$ nearly symmetric around its reference value. However, for the price curve with $\delta$, this makes the total price curve more asymmetric around its reference value. Hence, the price can lowered by moving $\delta$ above its reference value, but the effect of $\rho$ is ambiguous.
Figure 26: Plot of the price function for beta distribution parameters $\rho$ (left panel) and $\delta$ (right panel) for $\gamma = 0.2$ is shown. Point where all curves intersect is the reference value of the parameter.

Now, we draw a surface plot of the price function with both parameters, and show it along with the price contours in Figure 27. Both plots of Figure 27 show that the minimum price lies along a thin region around the line $\rho = \delta$. Actual values of the optimal parameters that minimize the price are shown in Figure 28, and the plot confirms our earlier intuition about the behavior of price with $\rho$ and $\delta$ as the optimal $\delta$ is higher than its reference value, while $\rho$ oscillates around its reference value. And the minimum price for a given value of $\theta$ obtained by minimizing the price over distribution parameters ($\rho, \delta$) is shown in right panel of Figure 28. We can see in Figure 28 that optimal parameters, which lead to minimum price for a given value of $\theta$ approach the reference parameters, as was the case for the skew normal distribution. As a result the ambiguity adjusted price (shown in the right panel of Figure 28) also approaches the price under reference model for large values of $\theta$. 
Figure 27: Surface (left panel) and contour (right panel) plots of the price function with the beta distribution parameters are shown.

Figure 28: The left panel shows the variation of the optimal parameters and the right panel shows the variation of the ambiguity adjusted price of the bond with the faith in reference model, for $\gamma = 4.34$. 
Figure 29: The left panel compares the ambiguity adjusted price (blue curve) with the risk adjusted price evaluated at the optimal parameters (green curve). The right panel shows the difference between the ambiguity adjusted price and the risk adjusted price.

Here again, we see that the ambiguity adjusted price approaches the risk adjusted price calculated at reference parameters, but lies always above the risk only price for $\theta < \infty$.

Figure 30 decomposes the difference between the ambiguity adjusted price and the reference risk adjusted price into the contribution from change in optimal parameters and the contribution from the decrease in risk aversion. The red curve shows the change in price that results from a change in parameters of the model, and the blue curve shows the change in price that comes from a decrease in risk aversion level. The difference between the two curves is always equal to one hundred percent. Plot in the right panel of the Figure 30 shows the ratio of the red curve to the blue curve. Similar to the case of the skew normal distribution, we notice that for the beta distribution, the parameters’ contribution to the price increases with increasing $\theta$. 
5.1.3. Comparison of Ambiguity’s Effect on the Skew Normal and the Beta Distribution under Utility Formulation

We are now in a position to compare the effects of the ambiguity adjustment on the bond price under the beta and the skew normal distribution. As described earlier, this effect can be broken down into a change in price that results from the change in model parameters, and a change in price that comes from the change in the level of risk aversion. Since both components of ambiguity adjustment have identical dependence on $\theta$, it is enough to compare any one of the two components. We, therefore, compare only the contribution that comes from the change in the risk aversion level, along with the ratio of the two contributions. Results are shown in Figure 31. Both curves show different behavior. Contribution of the penalty term for the beta distribution decrease, while it increases for the skew normal distribution. The ratio plotted in the right panel shows the opposite behavior, as it has the contribution of the penalty term in the denominator. This implies that the risk adjusted price approaches the ambiguity adjusted price more quickly for the skew normal distribution as compared to the beta distribution.

Figure 30: Difference between the ambiguity adjusted price and the reference risk price is decomposed into the contribution from a change in model parameters (red curve) and a contribution from decrease in risk aversion level (blue curve) is shown in the left panel. Right panel shows the ratio of the red curve to the blue curve.
5.2. Ambiguity Adjustment with the Wang Transform

In this section, we incorporate the effect of ambiguity in the Wang pricing mechanism. We use the same penalty functions that we used in the utility formulation (defined in equations (5.1) through (5.4)). However, for the Wang transform we do not have a utility function to add this penalty term to. We, therefore, make the Wang risk aversion parameter a function of this penalty term. That is, we modify the Wang transform given in equation (3.1) to

$$F_{X}[x, \omega, \lambda, \theta] = \Phi[\Phi^{-1}P[x, \omega] + \lambda(\theta, \omega)]$$

(5.12)

where $\omega$ is the set of distribution parameters, and the $\lambda(\theta, \omega)$ is the Wang risk aversion parameter, which now depends on the ambiguity aversion parameter and the distribution parameters.

We define the dependence of risk aversion on ambiguity aversion as
\[ \lambda(\theta, \omega) = \lambda - \theta \times h(\omega) \]  

(5.13)

where \( h(\omega) \) is the penalty function for distribution parameters \( \omega \) defined in equations (5.1) through (5.4).

Hence, the effect of ambiguity aversion is to allow investors to consider models different from the reference model at the cost of reducing its risk aversion for a given model. This treatment is similar to the one used in the utility formulation, where the presence of ambiguity allows investors to consider other models for the outcome at the cost of increasing its utility for a given model. However, the ambiguity aversion directly reduces the risk aversion parameter in the Wang transform, while in the utility formulation it does not directly affect the risk aversion parameter.

The Wang adjusted probability density function, in the presence of ambiguity term, then becomes

\[
f_\omega(x, \omega, \lambda, \theta) = \frac{\phi \left[ \Phi^{-1} F(x, \omega) \right] + \lambda - \theta \times h(\omega) }{\phi \left[ \Phi^{-1} F(x, \omega) \right]} f(x, \alpha, \mu, \sigma),
\]

(5.14)

where \( f_\omega(x, \omega, \lambda, \theta) \) is the risk and ambiguity adjusted probability distribution function.

The ambiguity adjusted price of the bond under the Wang transform can then be written as

\[
P(k, z_k, \lambda, \theta) = \min_{\omega} \left[ E_\omega \left[ B(z_1, z_2, z_3, k, z_k) \right] \right],
\]

(5.15)

where \( E_\omega \left[ B(z_1, z_2, z_3, k, z_k) \right] \) is the expected value of the Bond’s payout under the risk and ambiguity adjusted probability distribution given in (5.14). That is, we find the ambiguity adjusted price by finding the expected value of the bond’s payout under the probability distribution that minimizes the expected present value of the bond.
5.2.1. Ambiguity Adjustment for the Skew Normal Distribution

Now we apply the mechanism devised above to calculate the ambiguity adjusted price for the skew normal distribution. For the skew normal distribution, the fair price can be written explicitly as

\[
P(k, z_k, \lambda, \theta) = \min_{\alpha, \mu} \left[ \int_{-1}^{1} B(z_1, z_2, z_3, k, z_k) f_{z_1, z_2, z_3, \theta}(z_1, z_2, z_3, \alpha, \mu, \sigma, \lambda, \theta) \, dz_1 \, dz_2 \, dz_3 \right],
\]

(5.16)

where \(f_{z_1, z_2, z_3, \theta}(z_1, z_2, z_3, \alpha, \mu, \lambda, \theta)\) is the risk and ambiguity adjusted joint probability density function for the skew normal distribution.

For simplicity, we assume that the outcomes of three cohorts are uncorrelated. Under this assumption, the joint probability density function can be written as the product of the marginal density functions of the outcomes of the three cohorts,

\[
f_{z_1, z_2, z_3, \theta}(z_1, z_2, z_3, \alpha, \mu, \lambda, \theta) = f_{z_1, \theta}(z_1, \alpha, \mu, \sigma, \lambda, \theta) f_{z_2, \theta}(z_2, \alpha, \mu, \sigma, \lambda, \theta) f_{z_3, \theta}(z_3, \alpha, \mu, \sigma, \lambda, \theta).
\]

(5.17)

Now we study the dependence of the price function (price before minimizing over the distribution parameters) on the distribution parameters. Variation of the price with respect to the skewness and the position parameters is shown in Figures 32 and 33. As observed in the utility formulation, the price function is much more sensitive to the changes in the skewness parameter below its reference value than the changes in the skweness parameter above its reference value or changes in the position parameter. This implies that the price can be minimized by reducing the position parameter below its reference value.

Moreover, for higher \(\theta\), the possible reduction in price that can be achieved by moving away from reference parameters decreases. Hence, we can interpret \(\theta\) as the faith in reference model.
Figure 32: Variation of the fair price function with $\theta$ for different values of skewness (left panel) and position (right panel) parameters is shown.

Figure 33: Variation of the fair price function with the skewness and the position parameters is shown in left and right panels respectively.
Figure 34 shows a surface and contour plot of the fair price with the skew normal parameters. We see an elliptical region in the contour plot (only a part of it is visible) where the lowest of the price occurs. The region is almost identical to the one observed in the utility formulation. Figure 35 shows the optimal parameters and the price evaluated at those parameters. As expected, the optimal parameters as well as the price approach their reference values as the faith in reference model increases. Behavior of the optimal parameters with the faith in reference model is also consistent with the utility formulation.

Figure 34: Surface (left panel), contour (right panel) plots of the price function with the skew normal distribution parameters are shown.
Figure 35: The optimal parameters (left panel) and the fair price calculated at the optimal parameters (right panel) are shown.

In the end, we compare the ambiguity adjusted price with the risk adjusted price only calculated evaluated at the optimal parameters (shown in Figure 36). And we notice that the difference is always bigger than zero, and it decreases as the faith in the reference model increases. This is consistent with the result obtained in the utility formulation, and with our interpretation that the ambiguity aversion decreases the investor’s risk aversion, and the lower the faith in reference model, the greater the reduction in risk aversion compared to its reference level.
Figure 36: Comparison of the ambiguity adjusted price with the risk adjusted price evaluated at optimal parameters is shown in the left panel. The right panel shows the difference between the two prices as a percentage of the risk adjusted price (blue curve in the left panel).

Figure 37: Comparison of the two contribution to the ambiguity adjusted price. The left panel shows the individual contributions as a percentage of difference between the ambiguity adjusted price and the reference risk adjusted price, and the right panel shows the ratio of the two contributions.
Figure 37 compares the contribution to the ambiguity adjustment that come from the change in parameters and a change in risk aversion. For the Wang transform, the percentage contribution of both terms increases with increasing $\theta$, and the relative contribution of the change in parameters decreases with increasing $\theta$. This is in contrast to the utility pricing mechanism, where the change in percentage contributions decrease and the relative contribution of the parameters increase with $\theta$. This implies that for the Wang transform, increasing $\theta$ increases the effects of both the change in parameters and the risk aversion level, while both these effects decrease with increasing $\theta$ in the utility formulation.

5.2.2. Ambiguity Adjustment with Wang Transform for Beta Distribution

Now we apply the same procedure to the beta distribution to incorporate ambiguity aversion in it. The ambiguity adjusted CDF and PDF are

$$F_0[x, \rho, \delta, \lambda, \theta] = \Phi[\Phi^{-1}F[x, \rho, \delta] + \lambda - \theta \times h_{\text{beta}}(\rho, \delta)]$$ and

$$f_0(x, \rho, \delta, \lambda, \theta) = \frac{\Phi[\Phi^{-1}F[x, \rho, \delta] + \lambda - \theta \times h_{\text{beta}}(\rho, \delta)]}{\Phi[\Phi^{-1}F[x, \rho, \delta]]} f(x, \rho, \delta),$$

respectively, where $h_{\text{beta}}(\rho, \delta)$ is the penalty function for the beta distribution, and $\rho, \delta$ are the parameters of the beta distribution.

Using these functions, ambiguity adjusted price under the beta distribution becomes

$$P(k, z_k, \lambda, \theta) = \text{Min}_{\alpha, \mu} \left[ \int_{-1}^1 B(z_1, z_2, z_3, k, z_k) f_{z_1, z_2, z_3, \theta}(z_1, z_2, z_3, \rho, \delta, \lambda, \theta) \, dz_1 \, dz_2 \, dz_3 \right],$$

where $f_{z_1, z_2, z_3, \theta}(z_1, z_2, z_3, \rho, \delta, \lambda, \theta)$ is the joint beta probability density function of the outcome of the reconviction rate for the three cohorts.

Now we look at the effect of the distribution parameters on the price function. Figure 38 shows the variation in the price function with the ambiguity parameter ($\theta$) for different values of $\rho$ and $\delta$. As seen earlier in the case of the skew normal distribution, price
function becomes more sensitive to \( \theta \) as \( \rho \) and \( \delta \) move away from their reference values. This again reflects the fact that penalty term’s contribution to price becomes more significant as the candidate model becomes significantly different from the reference model.

![Figure 38: Variation of the price function with \( \theta \) for different \( \rho \) (left panel) and \( \delta \) (right panel) is shown.](image)

Figure 39 shows the variation of the price function with respect to the distribution parameters \((\rho, \delta)\). Overall behavior of the price function is similar to the behavior observed using the utility formulation. The price function increases (decreases) with an increase (decrease) in \( \rho \), while it decreases (increases) with an increase (decrease) in \( \delta \).

However, there is also an important difference in the behavior of the ambiguity component of the price function compared to its behavior in the utility formulation. In the case of the utility pricing mechanism, the ambiguity component of the price always increases as the distribution parameters move away from their reference values. However, this is not the case here, as the ambiguity component increases for small changes around the reference parameters, but then it starts decreasing, and eventually approaches zero. This appears counter-intuitive as moving the distribution parameters
away from their reference values increases the penalty term, which decreases the investor’s risk aversion, and hence, it should increase the price. But this is not what we observe. And the reason is that for the Wang transform, when the beta distribution parameters become sufficiently different from the reference parameters, the beta distribution becomes more peaked and highly concentrated. This, in turn, decreases the expected payout of the bond, and leads to a smaller price. Therefore, this counter-intuitive result is a limitation of defining the ambiguity adjusted Wang transform in this way. We do not see such result in utility based pricing, because the presence of an additional penalty dependent term in the price function (defined in equation ) keeps the price from decreasing, even if the first term decreases. The price function of the ambiguity adjusted Wang transform, on the other hand, does not have any such term, and therefore price decreases. However, we can avoid this problem numerically by restricting the range of $\rho$ and $\delta$ values such that the ambiguity adjusted price does not fall below the its risk adjusted level. Keeping this limitation in mind, we show a surface and contour plot for the price function for a range for $\rho$ and $\delta$ where the price function does not hit zero. These plots are shown in Figure 40.

Figure 39: Variation of the price function with the beta distribution parameters $\rho$ (left panel) and $\delta$ (right panel) is shown.
Figure 40: Surface (left panel) and contour (right panel) plots of the price function with $\rho$ and $\delta$ are shown.

Comparing Figure 40 with Figure 27, we see that the minimum price region for the Wang transform appears identical to the region obtained using utility formulation. So, despite the above mentioned limitation of the Wang transform, it can still be used to calculate the optimal parameters and the optimal price.

The ambiguity adjusted price for the beta distribution under the Wang transform is shown in Figure 41. For large values of $\theta$, the price approaches the reference price, as expected from earlier discussions. In the end, we compare the ambiguity adjusted price with the risk only price in Figure 42, and we obtain the expected result that the ambiguity price lies above the risk only price for small $\theta$, but approaches the same asymptotic limit for large $\theta$.

Figure 43 compares the contributions of the penalty term and the change in parameters to the ambiguity adjusted price, and the percentage contribution of both components increase with increasing $\theta$. This result is similar to the one obtained for the skew normal distribution under the Wang transform.
Figure 41: The optimal parameters (left panel) and the ambiguity adjusted price (right panel) with the faith in reference model are shown.

Figure 42: Comparison of ambiguity adjusted price with the risk only price evaluated at optimal parameters is shown in the left panel. The right panel shows the difference as a percentage of the risk adjusted price.
Figure 43: Comparison of the two contribution to the ambiguity adjusted price. The left panel shows the individual contributions as a percentage of difference between ambiguity adjusted price and reference risk price, and the right panel shows the ratio of the two contributions.

This implies that an increase in individual contributions due to a change in parameters and the level of risk aversion is a feature of the Wang transform, and not of the choice of distribution. Moreover, the ratio of the two contributions also decreases with $\theta$, which is again consistent with the skew normal result under the Wang transform.

5.2.3. Comparison of the Ambiguity’s Effect on the Skew Normal and the Beta Distribution under the Wang Transform

Here, we compare the effect of incorporating ambiguity adjustment in the Wang transform on the skew normal and the beta distributions. We compare the contribution of the penalty term (change in risk aversion level) and the ratio of the contribution due to change in parameters to the contribution due to change in risk aversion, in Figure 44.
Figure 44: The left panel compares the effect of penalty term for the skew normal and the beta distributions, and the right panel compares the ratio of the two contributions for the skew normal and the beta distributions.

Figure 44 shows that the beta distribution is more sensitive to the level of ambiguity aversion than the skew normal distribution, as the blue is more concave (convex) than the red curve in left (right) panel. This is in contrast to the case of the utility formulation, where ambiguity adjustment makes a similar impact on both the skew normal and the beta distributions.

5.3. Comparison of the Wang and the Utility Formulations

Now we compare the effect of ambiguity under the Wang and utility formulations for both the skew normal and the beta distributions. Figures 45 and 46 compare the ambiguity and risk adjusted prices (left panel), and the ratio of the two contributions to the ambiguity adjusted price (right panel). Both price curves (solid blue and green lines) in the left panel of Figure 45 have the same asymptotic limit. This implies that the ambiguity adjusted price under the Wang transform is more sensitive to the faith in reference model, as this curve has a higher curvature, and is always lower than the price under utility formulation. The ratio of the two contributions to the ambiguity adjusted price shows a similar pattern for both pricing mechanisms. The red curve lies slightly above the blue curve as the difference between the ambiguity and risk adjusted price...
under the utility formulation (difference between solid and dotted blue lines in the left panel) is smaller than the difference between the ambiguity and risk adjusted price under the Wang formulation (difference between solid and dotted green lines in the left panel).

For the beta distribution (shown in Figure 46), we see the same behavior for the ambiguity and risk adjusted prices (shown in the left panel). The price under the Wang transform is more curved than the price under the utility mechanism. The ratio of the two contributions to the ambiguity adjustment (shown in the right panel) in this case shows different behavior under the two pricing mechanisms. Both pricing mechanisms show different behavior. The ratio decreases for the utility formulation for small values of $\theta$, as the difference between the ambiguity and risk adjusted price (difference between solid and dotted blue lines in the left panel) increases sharply for small $\theta$.

Figure 45: Comparison of the Wang and the Utility formulations for the beta distribution. The left panel shows the ambiguity and risk adjusted prices for both pricing mechanisms, and the right panel shows the ratio of the parameters’ contribution to price to the penalty function’s contribution to price.
Figure 46: Comparison of the Wang and the Utility formulations for the beta distribution. The left panel shows the penalty function’s contribution to price, and the right panel shows the ratio of the parameters’ contribution to price to the penalty function’s contribution to price.

This analysis of the decomposition of the ambiguity adjusted price can help investors determine the appropriate choice of pricing mechanism, for a given faith in reference model. This is because the change in price attributed to the change in parameters can be seen as a pure ambiguity effect, and the change in price attributed to the penalty function can be seen as a change due to a decrease in the investor’s risk aversion level. Therefore, the ratio of the change in price attributed to the change in parameters to the change in price attributed to the penalty function, can be interpreted as the discount that investor requires for ambiguity for giving up one unit of discount for risk. Hence, this ratio represents the investor’s ambiguity aversion relative to its risk aversion. We will call it the ambiguity to risk ratio. An investor with higher relative risk aversion (lower relative ambiguity aversion) would require more ambiguity discount to give up one unit of risk discount. An investor with lower relative aversion to ambiguity should choose a pricing mechanism that has the highest ambiguity to risk ratio, for a given level of faith in the reference model. Therefore, for an investor with lower relative ambiguity aversion, the
Wang transform is a better choice for the beta distribution, and the utility pricing mechanism is more suitable for the skew normal distribution.
6. Effect of Social Utility

So far we have assumed that investors’ utility is derived solely from the financial return of the bond. However, a main incentive for investing in these bonds is their social impact, as evident by many charities and foundations investing in the bond. Therefore, we include another utility term to reflect the investor’s utility arising from the social impact of the bond. So the overall utility of investor is combination of two terms: a financial utility that arises from the bond’s financial return and a social utility that arises from the bond’s impact on recidivism, specifically

\[
\begin{align*}
\text{Financial Utility: } u_1(w, \gamma_1) &= \frac{-1}{\gamma_1}(e^{-\gamma_1 w} - 1) \\
\text{Social Utility: } u_2(z_1, z_2, z_3, \gamma_2) &= \frac{-1}{\gamma_2}(e^{-\gamma_2 (z_1 + z_2 + z_3)} - 1),
\end{align*}
\]

(6.1)

where \( w \) is the investor’s wealth, \( z_1, z_2, z_3 \) are the change in reconviction rates for the three cohorts, \( z_1 + z_2 + z_3 \) is the total drop in recidivism rate, and \( \gamma_1 \) measures the aversion to financial risk. \( \gamma_2 \) plays the same role for the social utility that \( \gamma_1 \) does for the financial utility. Therefore, we interpret \( \gamma_2 \) as the investor's aversion to social risk, or uncertainty in the social impact of the program.

The total utility is given by

\[
u(w, \gamma_1, \gamma_2) := u_1(w, \gamma_1) + A(E[B(z_1, z_2, z_3, k, z_k)], Z) u_2(z_1, z_2, z_3, \gamma_2(z_k))
\]

(6.2)

Where \( A(E[B(z_1, z_2, z_3, k, z_k)], Z) \) is the strength of social utility relative to the financial utility, \( E[B(z_1, z_2, z_3, k, z_k)] \) is the expected payout of the bond, and \( Z \) is the current recidivism rate in the country.

We make \( A(E[B(z_1, z_2, z_3, k, z_k)], Z) \) a function of the bond’s expected payout and \( Z \), as we expect investors to be more inclined to do charity if the bond’s expected payout and the recidivism rate is higher. Moreover, we expect the aversion to social risk (\( \gamma_2 \)) to be a
function of \( z_k \). This is because the cap rate is the drop in the reconviction rate above which government does not pay more for further decreases in reconviction rate. Hence, the cap rate \( z_k \) provides a natural measure for the drop in recidivism rate above which the investor’s social utility would stay almost constant with improving outcomes. For now, we will consider the \( A(E[B(z_1, z_2, z_3, k, z_k)], Z) \) to be constant, independent of both \( E[B(z_1, z_2, z_3, k, z_k)] \) and \( Z \).

### 6.1. Pricing Under Social Utility with Constant Relative Strength

We now calculate the indifference pricing in the presence of social utility function. The indifference equation is given by

\[
u(w_0, \gamma_1, \gamma_2) = E[u(w_0 - p + B(z_1, z_2, z_3, k, z_k), \gamma_1, \gamma_2)],
\]

where \( p \) is the price paid for the bond, and \( B(z_1, z_2, z_3, k, z_k) \) is the payout of the bond.

Using the utility function given in (6.2), we get

\[
p(k, z_k, \gamma_1, \gamma_2) = \frac{1}{\gamma_1} \ln \left( \frac{\left(1 + A(\gamma_1 / \gamma_2)e^{\gamma_1 w_0} (1 - E[e^{-\gamma_2(z_1 + z_2 + z_3)}])\right)}{E[e^{-\gamma_1 B(z_1, z_2, z_3, k, z_k)}]} \right).
\]

Comparing (6.4) with (2.2) we see that the addition of social utility inserts an extra factor of \( A(\gamma_1 / \gamma_2)e^{\gamma_1 w_0} (E[e^{-\gamma_2(z_1 + z_2 + z_3)}] - 1) \) in the numerator, which becomes negligible as \( A \) goes to zero.

Equation (6.4) can be written as

\[
p(k, z_k, \gamma_1, \gamma_2) = \frac{-1}{\gamma_1} \ln \left( E[e^{-\gamma_1 B(z_k, z_k)}] \right) + \ln \left( 1 + A e^{\gamma_1 w_0} \frac{\gamma_1}{\gamma_2} (1 - E[e^{-\gamma_2(z_1 + z_2 + z_3)}]) \right).
\]

First term in equation (6.5) can be seen as the price for the financial return of the bond, while the second term can be interpreted as the price for the social return of the bond. For
\( \gamma_1 = \gamma_2 = 1 \), and \( A = 1 \), the price for the social return is about 47\% of the price for the financial return. Hence, for a 100£ paid for the financial return, investor will pay about 47£ as charity.

In order to see the effect of relative strength of social utility and aversion to charity on fair price, the price of bond is plotted against \( A \) and \( \gamma_2 \) in the Figure 47.

**Figure 47:** The left panel shows that the fair price increases with increasing \( A \) (relative strength of social utility), and the right panel shows that the fair price decreases with increasing \( \gamma_2 \). For small \( A \) and large \( \gamma_2 \) price approaches the price of the bond in the absence of social utility.

The price increases logarithmically with \( A \), and decreases with \( \gamma_2 \). Increasing \( A \) means that the investor is willing to spend more money for a given expected drop in the reconviction rate, and hence the fair price increases. Increasing \( \gamma_2 \) makes investor more averse to uncertainties in the social outcome. Hence, increasing \( \gamma_2 \) for given uncertainty in the social outcome of the bond decreases the price for the bond. Therefore, we can interpret \( \gamma_2 \) as investor's aversion to charity as well.
Figure 48 shows the variation of fair price with aversion to charity (left panel), and risk aversion parameter of the investor. As expected, the price decreases monotonically with the charity aversion parameter. The price of the bond as a function of the risk aversion parameter, however, decreases initially but then changes its behavior as $\gamma_1 > \gamma_2$, and increases afterwards. This is because as $\gamma_1$ exceeds $\gamma_2$, investor becomes more averse to financial risk than charity, and pays more attention to the social outcome of the bond than its financial outcome. As a result the price moves towards the level that an investor solely investing for the social impact of the bond would pay. Moreover, this figure suggests that all else being equal, a more risk averse investor would pay higher charity than an investor with lower risk aversion, for certain risk and charity preferences. This is possible because a more risk investor would be willing to pay more for a given change in its utility than an investor with a lower risk aversion, and hence could pay more for charity. It is also clear from Figure 48, that for some of the curves with $A > 0$, the price can be minimized for a certain choice of risk aversion and charity aversion parameters.

We explore this optimal point corresponding to the minimum price further in Figure 49. The right panel of Figure 49 shows the derivative of the price with respect to the level of
risk aversion parameter for different values of $A$. This plot shows that the price is very sensitive to the risk aversion parameter in the region where the price decreases with increasing risk aversion (negative slope). However, the price’s sensitivity to risk aversion level does not always stay negative, and it actually becomes positive once the risk aversion crosses a threshold. Hence, increasing risk aversion only decreases the price to a certain extent. Moreover, we notice that the value of the derivative, for a given risk aversion, decreases with $A$. This is consistent with our intuition that investor pays less attention to its risk aversion as $A$ increases. Plot in the left panel of Figure 49 shows the optimal risk aversion ($\gamma_1$) with the charity aversion parameter ($\gamma_2$) for three different level of $A$. We see that the optimal risk aversion increases with the charity aversion parameter. Hence, an investor with a higher level of charity aversion could decrease its price by further increasing its financial risk aversion. In addition, the curves move downward with increasing $A$. This is because an investor with higher $A$ would be willing to spend more on charity for given risk preferences, and hence would quickly lose its ability to reduce its price by increasing its risk aversion.

![Figure 49: Variation of the optimal risk aversion parameter with the charity aversion parameter for different values of $A$ is shown in the left panel. The right panel shows the derivative of the fair price with the risk aversion parameter.](image)
Figure 50: The left panel shows the constant price curves for different prices and same A, and the right panel shows the constant price curves for different A and the same price.

Figure 50 shows the set of $\gamma_1, \gamma_2$ values that lead to the same price. The negative slope of the curves at all points indicates that an increase in one parameter would require a decrease in the other parameter to keep the price at the same level. In other words, if an investor becomes more risk averse, she would pay the same price for the bond if her aversion to charity decreases. Curves move downward with increasing price as the investor needs to be less charity averse to pay the higher price for the same level of risk aversion. Similarly, increasing $A$ shifts curves upward as an investor with higher $A$ derives higher utility from the social impact of the bond and would be willing to pay a higher price even for relatively higher aversion to charity than the investor with lower $A$. In the left panel of Figure 50, we see an interesting behavior in terms of the convexity of the curves. For small values of the financial risk aversion parameter, the convexity is negative. This implies that an investor with a higher level of risk aversion would be willing to become more charitable to increase her risk aversion by one unit, as compared to an investor with a lower risk aversion. That is, investor values her risk aversion more
as her risk aversion increases. However, this behavior changes for larger values of the risk aversion parameter. As convexity becomes positive, investor values her risk aversion less, and becomes willing to trade off one unit of risk aversion for a lower amount of charity aversion.

At this point, investors with different levels of risk and charity aversion parameters that lie on a constant price curve are indistinguishable from each other. However, these investors would likely respond differently to changes in the notional and the cap rate of the bond. For a fixed expected payout of the bond, an investor with a lower level of aversion to charity is likely to pay a higher price for the bond that has a higher cap rate and lower notional, as compared to a bond that has a higher notional and lower cap rate. This is because, a bond with a higher cap rate would pay more when the reconviction rate drops by a higher amount, and the higher drop in reconviction rate would make a socially responsible investor pay more for the bond. Therefore, the investor’s social utility can be useful in determining the optimal choice for the notional and the cap rate for the bond. We will further explore this point in the next section, when we price the bond from the issuer’s perspective, and aim to find the optimal choice for the bond parameters.
Chapter 7

7. The Fair Price of the SIB From the Issuer’s Perspective

In this section, we calculate the bond price from the issuer’s perspective, and study the effects of the issuer’s budget constraints, risk aversion, and its propensity to do social welfare on the bond price. When pricing from the issuer’s perspective, we consider the bond’s notional \((k)\) and the cap rate \((z_k)\) to be variable, as the values of these parameters are chosen by the issuer. For a given choice of issuer’s utility parameters, and budget constraints, we then obtain a set of \(k\) and \(z_k\) values that make the bond a feasible investment for both the issuer and the investor. This also allows us to identify the optimal values of \(k\) and \(z_k\) such that the issuer’s risk adjusted profit is maximized within its budget constraints.

7.1. Risk Adjusted Price for the Issuer

We start by modeling the issuer’s risk aversion with the exponential utility. Specifically

\[
u_i(w_i, \gamma_i) := \frac{-1}{\gamma_i} \left( e^{-\gamma_i w_i} - 1 \right),
\]

(7.1)

where \(u_i(w_i, \gamma_i)\) is the issuer’s utility, \(\gamma_i\) is the issuer’s risk aversion, and \(w_i\) is the wealth of the issuer.

Now we compare the issuer’s utility of funding the Peterborough Prison rehabilitation program with its utility to fund the cost of keeping prisoners incarcerated without any rehabilitation program. Without the rehab program, the government\(^6\) would have to pay the cost of keeping prisoners incarcerated, while upon funding the rehab program, the government has to pay the bond’s payout, but it saves the amount spent on keeping

---

\(^6\) The bond is issued by the Social Finance, a financial intermediary, and the payout is made by the UK’s Ministry of Justice, and the Big Lottery Fund. But, in this section, we make no distinction between the issuer of the bond and the UK government.
prisoners in prison if the recidivism rate drops. Hence, we calculate the fair price of the
bond as the minimum price that government needs to charge for the bond in order to
equate its expected utility of funding the rehab program to its utility of funding the cost of
prison without the rehab program. This leads to

\[ u_i(w_i - C(Z_0)) = E[u_i(w_i + p - C(Z) - B(z_1, z_2, z_3, k, z_k))], \]  \quad (7.2)\]

where \( C(Z_0) \) is the cost of keeping prisoners in prison per bondholder when the
recidivism rate is \( Z_0 \), and \( C(Z) \) is the cost of keeping in prisoners in prison per
bondholder when the recidivism rate is \( Z \). \( p \) is the minimum price the issuer is willing to
accept for the bond, and \( B(z_1, z_2, z_3, k, z_k) \) is the bond’s payout.

We define the government’s savings from issuing the bond as

\[ S(z_1, z_2, z_3) = C(Z_0) - C(Z), \]  \quad (7.3)\]

where \( z_1, z_2, z_3 \), are the drops in reconviction rate for the 1\textsuperscript{st}, 2\textsuperscript{nd}, and 3\textsuperscript{rd} cohorts, respectively.

This leads to the following form for the price

\[ p_i(k, z_k, y_i) = \frac{1}{y_i} \ln \left( \int_{z_1} \int_{z_2} \int_{z_3} e^{y_i[B(z_1, z_2, z_3, k, z_k) - S(z_1, z_2, z_3)]} f(z_1, z_2, z_3, \omega) \, dz_1 \, dz_2 \, dz_3 \right), \]  \quad (7.4)\]

where \( f(z_1, z_2, z_3, \omega) \) is the joint probability density function for the change in
reconviction rate in the three cohorts of the Peterborough Prison, and \( \omega \) is the set of
distribution parameters.

We now introduce another parameter \( c \), which denotes the cost of keeping one prisoner in
prison for one year divided by the total number of bonds sold. The parameter \( c \) therefore
represents the amount that government saves per bondholder if one prisoner does not go
back to prison. Hence, the government’s savings can be written as

\[ S(z_1, z_2, z_3) = C(Z_0) - C(Z) = c \times (z_1 + z_2 + z_3). \]  \quad (7.5)\]
Depending on the relative values of \( c \) and \( k \), government’s price can be either positive or negative. If \( c > k \), government’s price is likely to be negative as it saves more money per prisoner than it pays to the bondholders. By default, we take \( c = k = 0.01£ \).

Figure 51: The left panel shows the risk adjusted price for \( z_k = 13\% \), and the right panel shows the price for \( k = 0.01£ \). Both panels have \( c = 0.01£, \gamma_I = 0.1, \) and \( \gamma_B = 1 \).

Plots of the issuer’s price versus notional amount (left panel) and cap rate (right panel) are shown in Figure 51. These plots indicate that for \( k = 0£ \) the government saves approximately 2£ per bondholder, if it funds the rehab program, and that the government’s savings exceed its risk adjusted cost for \( k < 0.012£ \), if the cap rate is fixed at 13%. Notice, however, that \( k = 0£ \) corresponds to the scenario where the government funds the rehabilitation program on its own without issuing any bond. Expected profit at \( k = 0£ \) is, therefore, the profit that the government can earn without issuing any bond. Moreover, for \( k = c = 0.01£ \), the issuer’s savings exceed the risk adjusted payout for all \( z_k < 30\% \). Thus if the government chooses to pay a notional of 0.01£ per prisoner, it can increase its cap rate to 30% before it depletes its savings. The difference between the black curve (buyer’s price) and the red curve (issuer’s price) can be considered the issuer’s risk adjusted profit. The region where red curve is below the black curve is the profitable region for the issuer, because the minimum price issuer is willing to receive for
the bond lies below the maximum price the buyer is willing to pay. To maximize its profit, issuer would want to choose \( k \) and \( z_k \) where its profit is the largest.

### 7.2. Risk Adjusted Price Under Budget Constraints

Our utility analysis assumes that government would be willing to issue the social impact bond irrespective of the price buyers pay for it. But that is not the case. As indicated in the introductory section of this report, part of the SIB’s appeal to the government is that it saves the government from paying the money upfront, and the government cannot always fund such rehab programs due to budgetary constraints. Therefore, it would only be willing to issue the bond if it can raise enough money from buyers to cover a significant portion of the upfront cost of the rehab program.

We include this budgetary constraint by introducing a new parameter \( p_F \) in our model, which refers to the minimum amount the government needs to raise from a bondholder in order to fund the rehab program. So if the price that the buyer is willing to pay for the bond is less than the minimum amount needed to be raised to fund the program \( (p_F) \), then the project will not be initiated. Here we also assume that the number of bonds sold is independent of the price charged. This is a reasonable assumption because the price charged never exceeds the maximum price that buyer is willing to pay for the bond, and hence, the buyer should be mostly indifferent between different prices. If the bond price increases that is because the risk adjusted payout of the bond increases, and not because the buyer lowers her risk aversion. In other words, the buyer pays more for a higher expected payout and not for the same payout.

To take into account this budget constraint, we modify the price function to

\[
p_I(k, z_k, y_1) = \begin{cases} 
p_F; & \text{if } p_B < p_F \\
p_I(k, z_k, y_1); & \text{if } p_B > p_F' 
\end{cases}
\]  

where \( p_F \) is the minimum amount per bondholder that the government needs to raise in order to fund the rehabilitation program.
This makes the price function discontinuous, and puts a lower limit on $k$ and $z_k$. If the government selects $k$ and $z_k$ that are too small, buyers will not be willing to pay enough for the bond to cover the minimum funding cost required to fund the rehab program.

Figure 52: The left panel shows the issuer’s price with the notional ($k$) for $z_k = 13\%$. The right panel shows the issuer’s price with the cap rate ($z_k$) for $k = 0.008$. For both panels, $c = 0.01 \£$, $p_F = 1 \£$, $\gamma_I = 0.1$, and $\gamma_B = 1$.

The price function is plotted in Figure 52. In the presence of minimum funding cost ($p_F$), the issuer’s profitability region shrinks. Hence, the higher the minimum amount that the government wants to raise from the investor ($p_F$), the narrower the profitability region.
Figure 53: Plots in the upper panel show the surface plots of the buyer and issuer price, while surface plot in the lower left panel shows issuer’s profit and the contour plot in the lower right panel shows the region where profit exceeds zero. All plots have been obtained for $c = 0.01\text{£}, p_f = 1\text{£}, \gamma_I = 0.1$, and $\gamma_B = 1$. All plots have the same x,y, and z axes.

Figure 53 shows the surface plots of the buyer and the issuer prices with $k$ and $z_k$.

Looking at the contour plot in the lower right panel of Figure 53, we notice that the width of the profitability region shrinks as we increase the notional ($k$). Moreover, the region
also shifts leftward along the cap rate \((z_k)\) axis. This is because increasing the notional amount increases the bond’s payout, which increases both the buyer’s and the issuer’s price for a given \(z_k\). An increase in the buyer’s price shifts the lower limit of the profitability band backward because the minimum funding cost can be raised at lower \(z_k\). Similarly, increase in issuer’s price for a given \(z_k\), shifts the upper end of the profitability band backward as the issuer’s price exceeds the buyer’s price for smaller \(z_k\). Somewhere within this region lies the optimal choice of \(k\) and \(z_k\) that maximizes the issuer’s profit.

### 7.3. Effect of Social Welfare Spending on the Fair Price

Before we try to find the optimal choice of parameters, we take into account the effect of the issuer raising more than the bare minimum cost of the program. We can expect that with the higher upfront funding, the issuer would be able to invest more in the rehabilitation program and achieve better outcomes. Therefore, if the price that the buyer pays for the bond exceeds the minimum amount needed to fund the program, the probability distribution for the outcome of the rehabilitation program will move forward. That is, we expect the position parameter to move forward by

\[
\Delta \mu = S(p_B - p_F),
\]  

(7.7)

where \(S\) is the drop in recidivism rate that can be achieved by an additional unit of initial funding. We will assume that the \(S\) is constant, and by default we will take \(S = 0.1\). That is, we expect one additional unit of upfront funding to reduce recidivism rate by an additional 10%. This is reasonable because we have taken the minimum cost to fund the program to be \(1£\) and it is expected to decrease the recidivism rate by 10% from its normal level.

Moreover, we assume that the government would be willing to spend extra money for social welfare if it can achieve higher reduction in recidivism rate. Therefore, we add another term to the issuer’s utility function, which increases when \(p_B\) exceeds \(p_F\). We model this with an exponential utility function, and we write the change in issuer’s social utility as
\[
\Delta u_S(y_S) = I(p_B - p_F > 0) \frac{B}{y_S} (1 - e^{-y_S (p_B - p_F)}),
\]  

(7.8)

where \(y_S\) is the issuer’s aversion to social welfare, and \(B\) is the amount issuer is willing to spend for social welfare for a given decrease in recidivism rate, \(I(\cdot)\) is the indicator function.

The price function then becomes

\[
p_I(k, z_k, y_I, y_S) = \frac{1}{y_I} \ln \left( E \left[ y_I \left( B(z_1, z_2, z_3, k, z_k) - S(z_1, z_2, z_3) \right) \right] \right) \\
- \frac{1}{y_I} \ln \left[ 1 + I(p_B - p_F > 0) B \frac{y_I}{y_S} (1 - e^{-y_S (p_B - p_F)}) \right].
\]  

(7.9)

Figure 54: The left panel shows the issuer’s price with the notional amount, and the right panel shows issuer’s price with the cap rate. Total price is represented by the yellow curve. For both panels: \(y_I = 0.1, y_S = 0.1, y_B = 1, S = 0.1, \) and \(B = 1.\)

Figure 54 shows the variation of the price with respect to the bond’s notional (left panel) and cap rate (right panel). Here we see that the red curve, which represents the issuer’s price for risk, bends downward as \(p_B\) exceeds \(p_F\). This is a result of making the
probability distribution for the outcome of the rehab program a function of the buyer’s price. As the buyer’s price exceeds the minimum funding cost \((p_F)\), the rehab program can deliver better results, which result in higher savings for the government. The green curve represents the effect of the social welfare spending of the government. As the expected drop in the reconviction rate increases (due to higher upfront money raised), the government’s willingness to subsidize the bond increases, and results in a lower price that the government is willing to accept for the bond. The yellow curve is the total price that the government is willing to receive for the bond. Once \(p_B\) exceeds \(p_F\), the yellow curve is simply the sum of the red curve and the green curve. The region where the yellow curve lies below the black curve is the profitable region for the issuer, and the presence of the green curve increases this region.

Figure 55 shows surface plots of the buyer’s and the issuer’s price (upper left and right panels), and surface and contour plots for the issuer’s profit (lower left and right panels). Comparing the plots of Figure 55 with the plots of Figure 53, we notice little difference. This is because for small values of \(B\) and \(S\), the effect on the prices and the issuer’s profit is relatively small. The most important effect of the non-zero values of \(B\) and \(S\) is on the optimal values for the bond’s notional and cap rate that maximize the issuer’s profit. If \(B\) and \(S\) are both zero, the optimal parameters are always the ones that make the buyer’s price equal to the minimum funding cost required to fund the program. However, non-zero values for \(B\) and \(S\), even if small, can change the optimal parameters.
Figure 55: The upper left and right panels show surface plots of the buyer’s price and the issuer’s price with the bond’s notional and cap rate, respectively. The lower left and right panels show surface and contour plots of the issuer’s profit, respectively. For all panels: $\gamma_I = 0.1$, $\gamma_S = 0.1$, $\gamma_B = 1$, $S = 0.1$, and $B = 1$. All plots have the same $x$, $y$, and $z$ axes.

7.4. The Optimal Bond Parameters

Now we look at how the optimal bond parameters vary with the issuer’s willingness to subsidize a higher reduction in recidivism rate. We first calculate the optimal parameters
for a given value of $B$, and then calculate the buyer’s price at those parameters. The left panel of Figure 56 shows the buyer’s price evaluated at the optimal parameters as a function of issuer’s subsidy, and the right panel shows the net buyer’s price (buyer’s price less issuer’s subsidy) as a function of issuer’s subsidy. We see that the buyer’s price increases with the issuer’s subsidy, however, the net buyer’s price starts decreasing after the issuer’s subsidy exceeds a certain limit. This provides one way for the issuer to choose the optimal subsidy such that it maximizes the net money raised from the buyers.

**Figure 56:** The left panel shows the buyer’s price with the issuer’s welfare spending, and the right panel shows the buyer’s price less the issuer’s subsidy.

Figure 57 shows the effect of government’s welfare spending on the government’s optimal profit. The left panel plots the government’s profit with its welfare spending, and the right panel shows the government’s profit per unit welfare spending. It is clear from the figure that while the government can increase its net profit (profit less welfare spending) by increasing its welfare spending, its profit per unit welfare spending decreases. The issuer profit increases with the welfare spending because the higher welfare spending increases the bond’s notional and/or cap rate, which increases the buyer’s price and the expected savings. Therefore, for small amounts of welfare spending, the government can extract more money from the buyers and the savings than it spends on welfare. Hence, the net profit increases.
Figure 57: The issuer’s profit (left panel) and the issuer’s profit per unit welfare spending (right panel) is shown.

Now, we look at how the optimal parameters and the optimal profit vary with the buyer’s risk and charity aversion parameters. We consider a set of buyers with different risk and charity aversion such that they pay the same price for the bond with $k = 0.01\£$, and $z_k = 13\%$. Next, we calculate the optimal parameters for each one of these buyers, and then calculate the issuer’s profit at these optimal parameters. This profit represents the maximum profit that the issuer can earn from a buyer with the given risk preferences. In order to calculate the optimal net profit, we subtract the subsidy paid by the issuer from the issuer’s total profit.
Figure 58: Relationship between the investor’s risk and charity aversion parameter (left panel) and the optimal welfare spending for a given choice of choice of risk preference (right panel) is shown.

The left panel of Figure 58 shows the set of investors that pay the same price for the bond issued with the reference parameters. Each point on this curve represents an investor with a different choice of risk preferences. We can see from the plot that as the investor becomes more averse to charity, its risk aversion decreases to keep the price constant. The right panel of Figure 58 shows the optimal welfare spending for each investor. We see that the optimal welfare spending decreases as the aversion to charity increases. That is, an investor with higher aversion to charity would optimize the profit when the government spends greater amount on the welfare spending.

7.4.1. Optimal Bond Parameters for Buyers with Different Risk Preferences

Next, we look at the dependence of the optimal profit on the government’s welfare spending for the set of investors shown in Figure 58. Figure 59 shows the optimal parameters (left panel) and the corresponding net profit (right panel) with the government’s welfare spending. We see from Figure 59 that the optimal profit initially decreases with welfare spending but then starts to increase again. The red curve in the right panel of Figure 59 shows the profit that the government earns at the reference
parameters \( (k = 0.01\pounds, \text{ and } z_k = 13\%) \). The optimal parameter touches the reference profit for a subsidy of about \(0.5\pounds\), and is always bigger than the reference profit for other values of welfare spending.

**Figure 59:** Variation of the optimal parameters (left panel) and the issuer’s optimal profit (right panel) is shown with the welfare spending.

The point where the red curve intersects with the green curve corresponds to the investor for whom the optimal parameters are the reference parameters. This is the investor that the issuer should aim to target if it issues the bond with the reference parameter. However, targeting this investor is not advisable because all investors, including the ones that could lead to higher profit for other choices of the bond parameters, would pay the same price at the reference parameters, and the government would be earning suboptimal profit. If the government sets the optimal parameters different from the reference parameters, then the optimal buyer would be the one that pays the highest price, and other buyers would simply not buy the bond. This will allow the issuer to earn the maximum profit. Moreover, this curve implies that the issuer can achieve the same level of profitability with at two different levels of welfare spending. This higher profit earned at different choices of the bond parameters is a consequence of utilizing the information that the buyers pay the same price at the reference parameters. This, however, comes with a
risk. The issuer is taking the risk that the investors would maintain their risk preferences even if the bond’s payout changes, and would be willing to transact, and buy the same number of bonds. The farther away the issuer moves from the reference parameters, the higher these risks, and the higher the profit opportunity. Hence, the additional profit that the government can earn by targeting buyers with optimal parameters different from the reference parameters can be seen as a reward for the risk that the government is taking based on the information of price that the buyers paid at the reference parameters. However, the government can only take advantage of this information in its future projects, as it would not have the price information before the issuance of the first bond. Therefore, the issuer can use this model to learn more and more about the buyers’ preferences from the repeated issuances of the bond and maximize its total profit over the total bond issuances.

![Graphs showing buyer's price and issuer's profit](image)

**Figure 60:** Net price paid by the buyer (left panel) and the decomposition of the issuer’s profit (right panel) are shown.

In order to determine which level of welfare spending (and the corresponding bond parameters) the issuer should choose, we first look at the variation of the buyer’s price (left panel of Figure 60) and the sources of the issuer’s profit (right panel of Figure 60). The red curve in the left panel represents the price that the buyer is willing to pay at the reference parameters ($k = 0.01\text{£}$, and $z_k = 13\%$). Unlike the issuer’s profit, which is
always bigger than reference profit, the buyer’s price is lower than the reference price when the welfare spending is lower than 0.5£, and increases with the welfare spending. Note that the government’s welfare spending has been subtracted from the buyer’s price in Figure 59 to yield the price that buyer is paying after compensating for the government’s subsidy. This plot helps us understand the sources of the issuer’s profit (shown in the right panel). For small subsidies the net profit comes almost exclusively from the issuer’s savings. However, the relative significance of the savings drops as the welfare spending increases, and the significance of the buyer’s price increases.

Figure 61 further clarifies this trend by plotting the contributions of the buyer’s price and the issuer’s savings to the issuer’s net profit. Issuer’s savings remain dominant, however, their relative contribution to the profit decreases. Therefore, if the issuer prefers savings over the upfront cash raised from buyers then it should lower its welfare spending. On the other hand, if the issuer prefers the upfront cash more than the savings, then it should increase its welfare spending. The right panel of Figure 61 shows the issuer’s profit margin (net profit per unit welfare spending), and it decreases exponentially. Therefore, while the issuer can increase its profit by increasing its welfare spending, its profit margin decreases with the welfare spending. This suggests that for a given level of profit, the issuer should choose the lowest welfare spending to maximize its profit margin. However, as shown in the left panel of Figure 61, for lower welfare spending, most of the issuer’s profit comes from the savings, which reduces the government’s ability achieve better social outcomes due to lower cash raised.
Figure 61: Contributions of the issuer’s savings and buyer’s price to the issuer’s profit (left panel) and the issuer’s profit margin (right panel) is shown.

Figure 62: Social Impact of the bond (left panel) and the total impact of the bond (right panel) are shown.

Therefore, in order to get a complete picture, we look at the variation of the social impact (as measured by the expected drop in recidivism rate) of the bond with the issuer’s welfare spending. In order to estimate the bond’s social impact, we use equation (7.7),
where we assume that the average drop in reconviction rate is proportional to the amount raised from the buyers above the minimum funding cost required for the rehab program. Hence, the social impact of the bond increases with the welfare spending, due to the rising buyer’s price. We multiply the issuer’s profit margin with the social impact of the bond in the right panel of Figure 61 to obtain a measure for the total impact of the bond. This plot shows that the same level of total impact can be achieved with two different levels of welfare spending. Lower welfare spending achieves higher total impact by increasing the profit margin, while the higher welfare spending increases the total impact by increasing the social impact of the bond. Moreover, since the different levels of welfare spending optimize profits for different types of investors, the issuer would essentially be choosing one of the two target investors by setting a level of welfare spending, to achieve a given level of total impact. And the higher the impact that the issuer wants to achieve, the higher would be the difference between the nature of investors. Therefore, to maximize the impact of the bond, the issuer would have to target either the investors with a relatively high risk aversion and very small (negative) charity aversion, or the investors with a relatively low risk aversion and a very high charity aversion. The investors in the middle with medium risk aversion and medium charity aversion would lead to the smallest impact.
8. Summary and Future Work

In this thesis, we have analyzed the social impact bond from the perspective of the buyer and the issuer. For the buyer, we have studied the combined effect of the buyer’s aversion to risk and ambiguity, and the buyer’s aversion to risk and charity. We have presented two pricing mechanisms with two different probability distribution functions for the underlying random variable, and examined the effect of these choices on the fair price and the sensitivity of the fair price to key variables. Our analysis has also shown that the interplay of the investor’s risk and charity preferences yields some interesting results. A buyer with a higher risk aversion may be willing to pay a higher price for the bond, and buyers with very different risk and charity preferences may pay the same price for the bond.

For the issuer, we have limited ourselves to the combined effect of the issuer’s aversion to risk and social welfare. However, even this simple analysis helps us gain some useful insights about how the issuer can adjust the bond’s parameters to achieve a required level of impact. We see that the issuer can increase its financial profit by targeting buyers with relatively lower aversion to risk, and it can increase the social impact of the bond by targeting buyers with relatively lower aversion to charity. In addition, we can conclude that the issuer should choose small values for notional and cap rate of the bond to target investors with relatively lower aversion to charity, and the issuer should choose higher notional and lower cap rate to target investors with relatively low aversion to financial risk. Issuer can use this model to learn about buyers' preferences and optimize the bond for specific buyers in each subsequent issuance, and maximize its total profit over all the bond issuances.

We could further extend this work by taking into account all three sources of uncertainty – financial risk, soical risk, and ambiguity – at once, for both the buyer and the issuer, and determine optimal bond parameters. In addition, here we have limited ourselves to
calculating the fair price for the buyer and the expected profit for the issuer. In future, this can be further extended to include value at risk models that allow the buyer and the issuer to do a more sophisticated risk and return analysis and adjust the bond’s notional and cap rate accordingly. The assumption that the bond’s payout structure is fixed can also be relaxed. And the effect of the changes in the bond’s payout structure can be studied from the perspective of both the issuer and the buyer, to find the payout structure that maximizes the bond’s total impact for the society.

Moreover, we have only considered how the social and financial returns of the bond may appeal different buyers with different risk and charity preferences differently, and how that may affect the fair bond price. This can be further extended by looking at the correlations of the social impact bond with different asset classes to incorporate diversification benefits of the social impact bonds. The fair bond price then will not only depend on the social and financial returns of the bond but also on its diversification benefits.
Bibliography


## Appendix 1: Rehabilitation Program Data

<table>
<thead>
<tr>
<th>Bins (i)</th>
<th>Average weight of a program in the bin ((w_i))</th>
<th>Frequency((f_i))</th>
<th>Weighted frequency ((f_i^* = f_i w_i / \sum_{i=1}^{n} w_i))</th>
</tr>
</thead>
<tbody>
<tr>
<td>-10%</td>
<td>0%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-5%</td>
<td>0%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>-2.50%</td>
<td>17%</td>
<td>1</td>
<td>1.258122744</td>
</tr>
<tr>
<td>0%</td>
<td>18%</td>
<td>6</td>
<td>7.844765343</td>
</tr>
<tr>
<td>7.50%</td>
<td>12%</td>
<td>13</td>
<td>11.39711191</td>
</tr>
<tr>
<td>12.50%</td>
<td>13%</td>
<td>11</td>
<td>10.87906137</td>
</tr>
<tr>
<td>17.50%</td>
<td>18%</td>
<td>6</td>
<td>8.140794224</td>
</tr>
<tr>
<td>22.50%</td>
<td>3%</td>
<td>2</td>
<td>0.493381468</td>
</tr>
<tr>
<td>27.50%</td>
<td>0%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>32.50%</td>
<td>5%</td>
<td>1</td>
<td>0.370036101</td>
</tr>
<tr>
<td>50%</td>
<td>0%</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>70%</td>
<td>5%</td>
<td>1</td>
<td>0.370036101</td>
</tr>
<tr>
<td>100%</td>
<td>0%</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Table 5: Bins and weighted frequency.
<table>
<thead>
<tr>
<th>Type</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education and Training</td>
<td>12%</td>
</tr>
<tr>
<td>Housing</td>
<td>40%</td>
</tr>
<tr>
<td>Cognitive</td>
<td>5%</td>
</tr>
<tr>
<td>Addiction</td>
<td>17%</td>
</tr>
</tbody>
</table>

Table 6: Weights assigned to rehab services offered to Peterborough Prisoners.

<table>
<thead>
<tr>
<th>Program</th>
<th>Type of Program</th>
<th>Drop in Reconviction Rate</th>
<th>Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key-Crest Program (Key only) (Gaes, 1999)</td>
<td>Drug-Addiction</td>
<td>-3.0%</td>
<td>17%</td>
</tr>
<tr>
<td>Straight Thinking On Probation (STOP) - 24 month follow up (Raynor, 1996)</td>
<td>Cognitive Skills</td>
<td>-2.0%</td>
<td>5%</td>
</tr>
<tr>
<td>Behavioral Change Program (Part of Final Warning Scheme) (Hine, 2001)</td>
<td>Cognitive Skills</td>
<td>0.0%</td>
<td>5%</td>
</tr>
<tr>
<td>The Intensive after care program (Griffiths, 2007)</td>
<td>Integrated</td>
<td>0.0%</td>
<td>12%</td>
</tr>
<tr>
<td>The Harlem Parole Reentry Court (Griffiths, 2007)</td>
<td>Integrated</td>
<td>0.0%</td>
<td>17%</td>
</tr>
<tr>
<td>Project Greenlight (Griffiths, 2007)</td>
<td>Integrated</td>
<td>0.0%</td>
<td>57%</td>
</tr>
<tr>
<td>Prison Labor and Prison (Gaes, 1999)</td>
<td>Employment and Education</td>
<td>0.0%</td>
<td>12%</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>--------------------------</td>
<td>------</td>
<td>-----</td>
</tr>
<tr>
<td>The Thinking Skills Program (Gaes, 1999)</td>
<td>Cognitive Skills</td>
<td>2.0%</td>
<td>5%</td>
</tr>
<tr>
<td>Summary of Vocational Training (Gaes, 1999)</td>
<td>Employment and Education</td>
<td>3.5%</td>
<td>12%</td>
</tr>
<tr>
<td>Summary of Vocational Programs - less than 100 hours (Gaes, 1999)</td>
<td>Employment and Education</td>
<td>4.1%</td>
<td>12%</td>
</tr>
<tr>
<td>Summary of Vocational Programs - more than 300 hours (Gaes, 1999)</td>
<td>Employment and Education</td>
<td>4.5%</td>
<td>12%</td>
</tr>
<tr>
<td>CDATE - average drop in adult reconviction (Gaes, 1999)</td>
<td>Cognitive Skills</td>
<td>5.0%</td>
<td>5%</td>
</tr>
<tr>
<td>Cognitive Thinking Skills Program(using largest scale of evaluation to date) Canada (Gaes, 1999)</td>
<td>Cognitive Skills</td>
<td>5.1%</td>
<td>0%</td>
</tr>
<tr>
<td>Prison Education and Prison Behavior (Adult Basic Education ABE - dropped out group) (Gaes, 1999)</td>
<td>Employment and Education</td>
<td>5.4%</td>
<td>0%</td>
</tr>
<tr>
<td>Program</td>
<td>Category</td>
<td>Integrated</td>
<td>0%</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>------------</td>
<td>-----</td>
</tr>
<tr>
<td>Maryland Reentry partnership (REP) (House of Commons Home Affairs Committee, 2010)</td>
<td>Integrated</td>
<td>5.6%</td>
<td>0%</td>
</tr>
<tr>
<td>Straight Thinking On Probation (STOP) - 12 month follow up (Gaes, 1999)</td>
<td>Cognitive Skills</td>
<td>7.0%</td>
<td>5%</td>
</tr>
<tr>
<td>Summary of Academic programs - fewer than 100 hours (Gaes, 1999)</td>
<td>Employment and Education</td>
<td>7.0%</td>
<td>12%</td>
</tr>
<tr>
<td>Cognitive Skill Courses (Gaes, 1999)</td>
<td>Cognitive Skills</td>
<td>8.0%</td>
<td>74%</td>
</tr>
<tr>
<td>Summary of Academic programs - more than 300 hours (Gaes, 1999)</td>
<td>Employment and Education</td>
<td>8.4%</td>
<td>5%</td>
</tr>
<tr>
<td>New Vision Program (Gaes, 1999)</td>
<td>Drug-Addiction</td>
<td>9.0%</td>
<td>12%</td>
</tr>
<tr>
<td>Cognitive Thinking Skills Program (Pilot in Correctional Services Canada) (Gaes, 1999)</td>
<td>Cognitive Skills</td>
<td>10.4%</td>
<td>5%</td>
</tr>
<tr>
<td>Key-Crest (Crest Only) (Gaes, 1999)</td>
<td>Cognitive Skills</td>
<td>11.0%</td>
<td>12%</td>
</tr>
<tr>
<td>Rehabilitation for Addicted Prisoners Trust (RAPt) program (House of Commons Home Affairs Committee,</td>
<td>Drug-Addiction</td>
<td>11.0%</td>
<td>17%</td>
</tr>
<tr>
<td>Program</td>
<td>Type</td>
<td>2010 (%)</td>
<td>2011 (%)</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>-----------------------</td>
<td>----------</td>
<td>----------</td>
</tr>
<tr>
<td>Prison Education and Prison Behavior (Adult Basic Education ABE - group who completed) (Gaes, 1999)</td>
<td>Employment and Education</td>
<td>11.5%</td>
<td>5%</td>
</tr>
<tr>
<td>Residential Treatment in Bureau of Prisons (Gaes, 1999)</td>
<td>Drug-Addiction</td>
<td>11.9%</td>
<td>5%</td>
</tr>
<tr>
<td>The Prolific and Other Priority Offender Program PPO (U.K) (Griffiths, 2007)</td>
<td>Integrated</td>
<td>12.0%</td>
<td>17%</td>
</tr>
<tr>
<td>Specialized Drug Offender Program (SDOP non-cognitive group) (Gaes, 1999)</td>
<td>Drug-Addiction</td>
<td>12.3%</td>
<td>0%</td>
</tr>
<tr>
<td>ASSET (Samo, Hearnden, &amp; Hedderman)</td>
<td>Employment and Education</td>
<td>13.0%</td>
<td>12%</td>
</tr>
<tr>
<td>Circles of Support and Accountability (Griffiths, 2007)</td>
<td>Integrated</td>
<td>15.1%</td>
<td>17%</td>
</tr>
<tr>
<td>Amity Right Turn Project (Gaes, 1999)</td>
<td>Integrated</td>
<td>15.8%</td>
<td>40%</td>
</tr>
<tr>
<td>Prison Labor and Prison Behavior (Gaes, 1999)</td>
<td>Employment and Education</td>
<td>16.0%</td>
<td>17%</td>
</tr>
<tr>
<td>Program</td>
<td>Focus Area</td>
<td>Improvement</td>
<td>Success Rate</td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>---------------------</td>
<td>-------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Specialized Drug Offender Program (SDOP) - Cognitive group only (Gaes, 1999)</td>
<td>Drug-Addiction</td>
<td>16.20%</td>
<td>12%</td>
</tr>
<tr>
<td>R &amp; R program (Gaes, 1999)</td>
<td>Cognitive Skills</td>
<td>19.0%</td>
<td>52%</td>
</tr>
<tr>
<td>Life Skills group, Cognitive Thinking Skills Program (Reasoning and Rehabilitation) Canada (Gaes, 1999)</td>
<td>Cognitive Skills</td>
<td>22.0%</td>
<td>17%</td>
</tr>
<tr>
<td>Key-Crest(Key-Crest group) (Gaes, 1999)</td>
<td>Cognitive Skills</td>
<td>31.0%</td>
<td>12%</td>
</tr>
<tr>
<td>Cognitive Skills Group, Cognitive Thinking Skills Program(Reasoning &amp; Rehabilitation) Canada (Gaes, 1999)</td>
<td>Cognitive Skills</td>
<td>51.40%</td>
<td>17%</td>
</tr>
</tbody>
</table>

Table 7: Historical data for the rehabilitation programs from around the world.
Appendix 2: Method of Moments

For the skew-normal distribution, mean, variance, and skewness are

\[
\begin{align*}
\text{mean} &= \mu + \sqrt{\frac{2}{\pi}} \frac{\sigma \alpha}{\sqrt{1 + \alpha^2}} \\
\text{variance} &= \sigma^2 \left(1 - \frac{2}{\pi} \frac{\alpha^2}{1+\alpha^2}\right), \\
\text{skewness} &= \frac{4-\pi}{2} \left(\frac{\sigma}{\sqrt{1 + \alpha^2}}\right)^3 \frac{1}{\left(1 - \frac{2}{\pi} \frac{\alpha^2}{1+\alpha^2}\right)^{3/2}}
\end{align*}
\] (A2.1)

These equations can be inverted to calculate \(\mu, \sigma, \alpha\) for the skew-normal distribution. This leads to

\[
\begin{align*}
|\delta| &= \frac{\pi}{2} \left|\gamma_3\right|^{\frac{2}{\pi}} \left|\gamma_3\right|^{\frac{4-\pi}{2}} \\
\Rightarrow \hat{\alpha} &= \frac{\delta}{\sqrt{1-\delta^2}}, \\
\hat{\delta}^2 &= \frac{Y_2}{1 - \frac{2}{\pi} \frac{\alpha^2}{1+\alpha^2}} \\
\hat{\mu} &= Y_1 - \sqrt{\frac{2}{\pi}} \frac{\alpha \sigma}{\sqrt{1 + \alpha^2}}
\end{align*}
\] (A2.2)

where \(Y_1, Y_2, Y_3\) are the sample mean, variance, and skewness respectively, and \(\delta = \frac{\alpha}{1+\alpha^2}\).

For the beta distribution, first two moments are

\[
\begin{align*}
E[x] &= \frac{\rho - \delta}{2\rho \delta} \left[1 + \rho \delta \Gamma(\rho)H(1,1-\delta,2+\rho,-1) + \delta \rho \Gamma(\delta)H(1,1-\rho,2+\delta,-1) \right] \\
E[x^2] &= \frac{1}{4} \left[\delta + \delta^2 + \rho + \rho^2 - 2\rho \delta \right] \left[\Gamma(\rho)H(1,1-\delta,3+\rho,-1) \right] + \frac{2+\delta+\rho+\rho(1+\rho)\Gamma(2+\delta)H(1,1-\rho,3+\delta,-1)}{\delta(1+\rho)\rho(1+\rho)}
\end{align*}
\] (A2.3)

where \(\Gamma(\cdot)\) is the Gamma function, and \(H(\cdot,\ldots,\cdot)\) is the Hypergeometric function.
Appendix 3: Maximum Likelihood Parameter Estimation

The likelihood functions for the skew normal distribution is

\[ A_{\text{skew}}(\alpha, \mu, \sigma) \equiv \prod_{i=1}^{n} f_{\text{skew}}(x_i)^{f(i)}, \tag{A3.1} \]

where \( f_{\text{skew}} \) is the probability density function of the skew-normal distribution and \( f(i) \) is the weighted frequency for the \( i^{th} \) bin (given in Table 1 of Appendix 1).

Similarly, the maximum likelihood function for the beta distribution is

\[ A_{\beta}(\rho, \delta) \equiv \prod_{i=1}^{n} f_{\beta}(x_i)^{f(i)}, \tag{A3.2} \]

where \( f_{\beta} \) is the probability density function of the beta distribution and \( f(i) \) is the weighted frequency for the \( i^{th} \) bin (given in Table 1 of Appendix 1).

The most likely parameters are then obtained by maximizing the log-likelihood function with respect to the distribution parameters.

For the skew-normal distribution, we get the following system of equations

\[
\begin{align*}
\frac{\partial \ln(A_{\text{skew}}(\alpha, \mu, \sigma))}{\partial \alpha} &= 0 \\
\frac{\partial \ln(A_{\text{skew}}(\alpha, \mu, \sigma))}{\partial \mu} &= 0 \\
\frac{\partial \ln(A_{\text{skew}}(\alpha, \mu, \sigma))}{\partial \sigma} &= 0
\end{align*}
\]

\[ \Rightarrow \left\{ \begin{array}{c}
\sum_{i=1}^{n} \frac{f(i)}{f_{\text{skew}}(x_i)} \frac{\partial f_{\text{skew}}(x_i)}{\partial \alpha} = 0 \\
\sum_{i=1}^{n} \frac{f(i)}{f_{\text{skew}}(x_i)} \frac{\partial f_{\text{skew}}(x_i)}{\partial \mu} = 0 \\
\sum_{i=1}^{n} \frac{f(i)}{f_{\text{skew}}(x_i)} \frac{\partial f_{\text{skew}}(x_i)}{\partial \sigma} = 0
\end{array} \right. \tag{A3.3} \]
Similarly, for the beta distribution the system of equations is

\[
\begin{align*}
\sum_{i=1}^{n} f(i) (x_i - \mu) \frac{\phi(\frac{\alpha x_i - \mu}{\sigma})}{\phi(\frac{\alpha x_i - \mu}{\sigma})} &= 0 \\
\sum_{i=1}^{n} f(i) \left[ \frac{\phi' \left( \frac{\alpha x_i - \mu}{\sigma} \right)}{\phi \left( \frac{\alpha x_i - \mu}{\sigma} \right)} + \frac{\phi(\frac{\alpha x_i - \mu}{\sigma})}{\phi(\frac{\alpha x_i - \mu}{\sigma})} \right] &= 0 \\
\sum_{i=1}^{n} f(i) \left[ \alpha \left( \frac{x_i - \mu}{\sigma} \right) \frac{\phi \left( \frac{\alpha x_i - \mu}{\sigma} \right)}{\phi \left( \frac{\alpha x_i - \mu}{\sigma} \right)} + \frac{x_i - \mu}{\sigma} \right] \frac{\phi' \left( \frac{\alpha x_i - \mu}{\sigma} \right)}{\phi \left( \frac{\alpha x_i - \mu}{\sigma} \right)} + 1 &= 0
\end{align*}
\]

(A3.5)

Similarly, for the beta distribution the system of equations is

\[
\begin{align*}
\frac{\partial \ln(A_{\text{beta}}(\rho, \delta))}{\partial \rho} &= 0 \\
\frac{\partial \ln(A_{\text{beta}}(\rho, \delta))}{\partial \delta} &= 0
\end{align*}
\]

(A3.6)

\[
\begin{align*}
\sum_{i=1}^{n} f(i) \left( \ln(1 + x_i) - \ln(2) - \frac{\partial B(\rho, \delta)/\partial \rho}{B(\rho, \delta)} \right) &= 0 \\
\sum_{i=1}^{n} f(i) \left( \ln(1 - x_i) - \ln(2) - \frac{\partial B(\rho, \delta)/\partial \delta}{B(\rho, \delta)} \right) &= 0
\end{align*}
\]

(A3.7)

Where \(B(\rho, \delta)\) is the Beta function defined in equation (2.3).
## Appendix 4: Notation

<table>
<thead>
<tr>
<th>Description of the Variable</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed, and predictive dataset for the drop in recidivism rate</td>
<td>$D, \bar{D}$</td>
</tr>
<tr>
<td>Drop in the reconviction rate of the 1\textsuperscript{st}, 2\textsuperscript{nd}, and 3\textsuperscript{rd} cohorts respectively</td>
<td>$z_1, z_2, z_3$</td>
</tr>
<tr>
<td>Average drop in the reconviction rate among the three cohorts</td>
<td>$\overline{z}$</td>
</tr>
<tr>
<td>Notional per prisoner for the bond</td>
<td>$k$</td>
</tr>
<tr>
<td>Cap rate for the bond</td>
<td>$z_k$</td>
</tr>
<tr>
<td>Total number of bonds held by an investor</td>
<td>$N$</td>
</tr>
<tr>
<td>The risk free interest rate</td>
<td>$r$</td>
</tr>
<tr>
<td>Time of payment for the 1\textsuperscript{st}, 2\textsuperscript{nd}, and 3\textsuperscript{rd} coupons respectively</td>
<td>$T_1, T_2, T_3$</td>
</tr>
<tr>
<td>The bond’s yield to maturity</td>
<td>$y$</td>
</tr>
<tr>
<td>The fair price of the bond under the Wang and utility pricing mechanism, respectively</td>
<td>$P, p$</td>
</tr>
<tr>
<td>Skewness, position, and scale parameters for the skew normal distribution</td>
<td>$\alpha, \mu, \sigma$</td>
</tr>
<tr>
<td>The beta distribution parameters</td>
<td>$\rho, \delta$</td>
</tr>
<tr>
<td>Method of Moments estimates for the skew normal distribution parameters</td>
<td>$\hat{\alpha}, \hat{\mu}, \hat{\sigma}$</td>
</tr>
<tr>
<td>Method of Moments estimates for the beta distribution parameters</td>
<td>$\hat{\rho}, \hat{\delta}$</td>
</tr>
<tr>
<td>Maximum likelihood estimates for the skew normal distribution parameters</td>
<td>$\tilde{\alpha}, \tilde{\mu}, \tilde{\sigma}$</td>
</tr>
<tr>
<td>Maximum likelihood estimate for the beta distribution parameters</td>
<td>$\tilde{\rho}, \tilde{\delta}$</td>
</tr>
<tr>
<td>Parameter</td>
<td>Symbol</td>
</tr>
<tr>
<td>--------------------------------------------------------------------------</td>
<td>--------</td>
</tr>
<tr>
<td>Reference parameters for the skew normal distribution</td>
<td>$\alpha^<em>, \mu^</em>, \sigma^*$</td>
</tr>
<tr>
<td>Maximum likelihood estimate for the beta distribution parameters</td>
<td>$\rho^<em>, \delta^</em>$</td>
</tr>
<tr>
<td>Set of distribution parameters for a general distribution</td>
<td>$\omega$</td>
</tr>
<tr>
<td>Risk aversion parameter for the Wang transform</td>
<td>$\lambda$</td>
</tr>
<tr>
<td>Risk aversion parameter for the utility formulation</td>
<td>$\gamma$</td>
</tr>
<tr>
<td>Initial wealth of the investor</td>
<td>$w$</td>
</tr>
<tr>
<td>Ambiguity aversion parameter</td>
<td>$\theta$</td>
</tr>
<tr>
<td>Investor’s aversion to charity</td>
<td>$\gamma_2$</td>
</tr>
<tr>
<td>Risk aversion parameter for the issuer</td>
<td>$\gamma_I$</td>
</tr>
<tr>
<td>Aversion to social welfare parameter for the issuer</td>
<td>$\gamma_S$</td>
</tr>
<tr>
<td>Initial wealth of the issuer</td>
<td>$w_I$</td>
</tr>
<tr>
<td>The cost per bondholder of keeping one prisoner in prison for one year</td>
<td>$c$</td>
</tr>
<tr>
<td>Minimum funding required to fund the rehabilitation program</td>
<td>$p_F$</td>
</tr>
<tr>
<td>The fair price of the bond for the issuer</td>
<td>$p_I$</td>
</tr>
<tr>
<td>Current recidivism rate for the Peterborough Prisoners</td>
<td>$Z_0$</td>
</tr>
<tr>
<td>Recidivism rate of the Peterborough Prisoners after the rehab program</td>
<td>$Z$</td>
</tr>
<tr>
<td>Buyer’s price received by the issuer</td>
<td>$p_B$</td>
</tr>
</tbody>
</table>

Table 8: Description of all the variables used and associated symbols is presented here.
<table>
<thead>
<tr>
<th>Description of the Function</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Payout for the (i^{th}) coupon</td>
<td>(C_i(z_i, k, z_k); i = 1, 2, 3)</td>
</tr>
<tr>
<td>Payout for the 4(^{th}) coupon</td>
<td>(C_4(z_1, z_2, z_3, k, z_k))</td>
</tr>
<tr>
<td>Payout for the bond</td>
<td>(B(z_1, z_2, z_3, k, z_k))</td>
</tr>
<tr>
<td>Probability density function for the skew normal</td>
<td>(f_{skew}(z, \alpha, \mu, \sigma))</td>
</tr>
<tr>
<td>Probability density function for the beta distribution</td>
<td>(f_{beta}(z, \rho, \delta))</td>
</tr>
<tr>
<td>Reference probability density function for the skew normal</td>
<td>(f^*_{skew}(z, \alpha, \mu, \sigma))</td>
</tr>
<tr>
<td>Probability density function for the beta distribution</td>
<td>(f^*_{beta}(z, \rho, \delta))</td>
</tr>
<tr>
<td>The Beta function</td>
<td>(\beta(\rho, \delta))</td>
</tr>
<tr>
<td>Maximum likelihood function for the skew normal distribution</td>
<td>(A_{skew}(\alpha, \mu, \sigma))</td>
</tr>
<tr>
<td>Maximum likelihood function for the beta normal distribution</td>
<td>(A_{beta}(\rho, \delta))</td>
</tr>
<tr>
<td>The Wang adjusted cumulative distribution function</td>
<td>(F_\lambda(z, \omega, \lambda))</td>
</tr>
<tr>
<td>The Wang adjusted probability distribution function</td>
<td>(f_\lambda(z, \omega, \lambda))</td>
</tr>
<tr>
<td>Cumulative distribution function for the skew normal distribution</td>
<td>(F_{skew}(z, \alpha, \mu, \sigma))</td>
</tr>
<tr>
<td>Cumulative distribution function for the beta distribution</td>
<td>(F_{beta}(z, \rho, \delta))</td>
</tr>
<tr>
<td>The Incomplete Beta function</td>
<td>(\beta_1(z, \rho, \delta))</td>
</tr>
<tr>
<td>Description</td>
<td>Formula</td>
</tr>
<tr>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------------------------------</td>
</tr>
<tr>
<td>The fair price for the $i$\textsuperscript{th} coupon under the Wang transform</td>
<td>$P_i(z_i, k, z_k); i = 1,2,3$</td>
</tr>
<tr>
<td>The fair price for the 4\textsuperscript{th} coupon under the Wang transform</td>
<td>$P_4(z_1, z_2, z_3, k, z_k)$</td>
</tr>
<tr>
<td>The fair price for the bond under the Wang Transform</td>
<td>$P(z_1, z_2, z_3, k, z_k)$</td>
</tr>
<tr>
<td>Financial utility function for the investor</td>
<td>$u_1(w, y_1)$</td>
</tr>
<tr>
<td>Penalty function for the ambiguity aversion</td>
<td>$h(Q'Q)$</td>
</tr>
<tr>
<td>Ambiguity adjusted CDF and PDF for the skew normal, respectively</td>
<td>$F_\theta(z, \alpha, \mu, \sigma, \lambda), f_\theta(z, \alpha, \mu, \sigma, \lambda)$</td>
</tr>
<tr>
<td>Ambiguity adjusted CDF and PDF for the beta, respectively</td>
<td>$F_\theta(x, \rho, \delta, \lambda, \theta), f_\theta(x, \rho, \delta, \lambda, \theta)$</td>
</tr>
<tr>
<td>Social utility function for the investor</td>
<td>$u_2(z_1, z_2, z_3, y_2)$</td>
</tr>
<tr>
<td>Coefficient of the social utility for the investor</td>
<td>$A(w, Z)$</td>
</tr>
<tr>
<td>Total utility function for the investor</td>
<td>$u(w, y_1, y_2)$</td>
</tr>
<tr>
<td>The risk adjusted fair price under utility formulation</td>
<td>$p(k, z_k, y)$</td>
</tr>
<tr>
<td>The ambiguity adjusted fair price under utility formulation</td>
<td>$p(k, z_k, y, \theta)$</td>
</tr>
<tr>
<td>The fair price for the bond under social utility</td>
<td>$p(k, z_k, y_1, y_2)$</td>
</tr>
<tr>
<td>Issuer’s utility for wealth</td>
<td>$u_1(w, y_1)$</td>
</tr>
<tr>
<td>Issuer’s utility for social welfare</td>
<td>$u_5(w, y_5)$</td>
</tr>
<tr>
<td>Government’s savings per bondholder</td>
<td>$S(z_1, z_2, z_3)$</td>
</tr>
<tr>
<td>The fair price for the issuer</td>
<td>$p_l(k, z_k, y_1)$</td>
</tr>
<tr>
<td>The fair price for the issuer willing to do spend for social welfare</td>
<td>$p_l(k, z_k, y_1, y_5)$</td>
</tr>
</tbody>
</table>
Table 9: All the functions and associated notation used is shown here.
# Curriculum Vitae

<table>
<thead>
<tr>
<th>Name:</th>
<th>Majid Hasan</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Post-secondary Education and Degrees:</strong></td>
<td>University of the Punjab Lahore, Punjab, Pakistan 2004-2008 B.Sc.</td>
</tr>
<tr>
<td><strong>Honours and Awards:</strong></td>
<td>President Talent Farming Scholarship 2004-2007</td>
</tr>
<tr>
<td></td>
<td>B.Sc. Gold Medalist, University of the Punjab 2008</td>
</tr>
<tr>
<td><strong>Related Work Experience:</strong></td>
<td>Teaching Assistant The University of Western Ontario 2011-2013</td>
</tr>
</tbody>
</table>