A Practical Real Options Approach to Valuing High-Frequency Trading System R&D Projects

Article *in* The Journal of Trading · June 2013

DOI: 10.3905/jot.2013.8.3.040

2 authors:

Andrew Kumiega
Illinois Institute of Technology
25 PUBLICATIONS **49** CITATIONS
SEE PROFILE

Ben Van Vliet
Illinois Institute of Technology
34 PUBLICATIONS **60** CITATIONS
SEE PROFILE

A Practical Real Options Approach to Valuing High Frequency Trading System R&D Projects

Andrew Kumiega, Ph.D.
Illinois Institute of Technology
565 W. Adams, Chicago, IL 60661
p. 312 906-6513 e. akumiega@iit.edu

Ben Van Vliet, Ph.D., *Corresponding*
Stuart School of Business,
Illinois Institute of Technology
565 W. Adams, Chicago, IL 60661
p. 312 906-6513 e. bvanvliet@stuart.iit.edu

February 15, 2013

Abstract

In the age of automation, trading and market making is about estimating the fair price of automated trading system research and development projects. This requires a new methodology to arrive at such a fair price. A real options framework is a natural choice. In this paper we review a methodology for automated trading system R&D as well as a practical real option model for valuing such projects so as to enable rapid strategy cycling.

Key Words

High-frequency trading, R&D project management, Real options

Trading and market making used to be about estimating the fair price of a financial instrument at a future time. The trader bought below the fair price estimate and sold above it, or vice versa. However, trading and market making has evolved into a complex information technology business. Trading is done by computer systems by way of quantified and codified trading strategies.

Trading and market making is now about estimating the fair price of an automated trading system research and development (henceforth R&D) project. This requires a new methodology to determine such a fair price. A real options framework is a natural choice. In the trading industry, options are of course well-known, and traders intuitively understand managing a portfolio of options. But, the trader's new job is not to manage a portfolio of options on stocks (say), but rather to manage a portfolio of real options on trading system R&D projects.

These projects depend upon uncertain costs and uncertain payoffs. They are characterized by large up-front investment in quantitative research and technology, and high probabilities of project failure. The key to automated trading is to rapidly and inexpensively cycling through and evaluate different strategies and technologies, spending additional time and money on only those few strategies and technologies that show promise. The purpose of this paper is to map the various components of such R&D in order to arrive at a practical model that enables valuation of these real options.

The paper proceeds as follows. In Section II we provide background on real options thinking. In Section III we provide an overview of the trading system development methodology (henceforth K|V methodology) put forth in Kumiega and Van Vliet (2008). In Section IV we describe revenue estimation for trading systems. In

Section V we parameterize costs and times for the three R&D stages of trading system development. In Section VI we discuss the real option valuation. Section VII concludes.

II. REAL OPTIONS

A vast literature on real options exists. We admire Datar and Mathews' method (DM) (see Mathews et al. 2007) for its simplicity and transparency. Our model uses DM as a foundation and like DM our model has the feel of NPV analysis. Additionally, we draw upon new product development research (see Cooper 2001) and the ensuing research of Gunther McGrath, et al. (2004), Adner (2007), Hackett and Dilts (2004) and Onno and Pennings (2001), who marry the Stage-Gate®[1] methodology with real options.

In a multi-stage R&D process, gates between stages represent option expirations. Gates predefine incremental releases of capital. An optional release structure limits R&D capital providers' loss potential by tying capital to deliverables, real option valuation techniques, and gate-passage criteria. Essentially, at each successive gate management must make a stronger and stronger commitment to the development project. The option on the ensuing stage gets exercised with some level of probability. Higher probabilities of technological and strategic success suggest a greater probability of continuing R&D in the next stage.

Projects which are found to contain little expected return or that no longer meet management's business objectives can be terminated prior to larger expense of time and money. The greater the uncertainty of payoffs, the more valuable the flexibility afforded by delayed investment processes. Through the R&D stages, uncertainties are resolved and at each gate management revalues the option with new distributions. Such

[1] "Stage-Gate" is a registered trademark of Product Development Institute Inc.

quantification coupled with rigorous a R&D process should ensure that politics do not drive project asset allocation.

The value of the call option on an R&D project at any given gate is the maximum of the present value of the expected future cash flows generated by that project, minus the R&D costs for project completion, or zero:

$$\text{Call Option Value} = Max(\text{Present Value} - \text{R \& D Costs}, 0) \quad (1)$$

For a trading firm to survive, new strategies must continually be researched and developed. As technological speed is a source of competitive advantage, new hardware and software must be also be researched continually. Because the value of a strategy is a function of the duration of expected cash flows and strategies are perishable, the ability to quickly deploy capable strategies on a technological platform means a firm is able to capture more of the trading opportunity.

II. R&D METHODOLOGY

Our K|V methodology progresses in a four-stage waterfall—design-test-implement-manage (DTIM)—as in Royce (1970), focusing on quantitative methods, data, technology, and risk management respectively. At the outset ($t = 0$), project revenues, costs and times are estimated. At each inter-stage gate, a decision is made to go or kill the project. After Stage 3 comes Gate 3, a decision to launch (time ℓ). Given launch, the trading system operates until τ, when the system loses its edge and is shut down. The gates correspond to nodes in Figure 1:

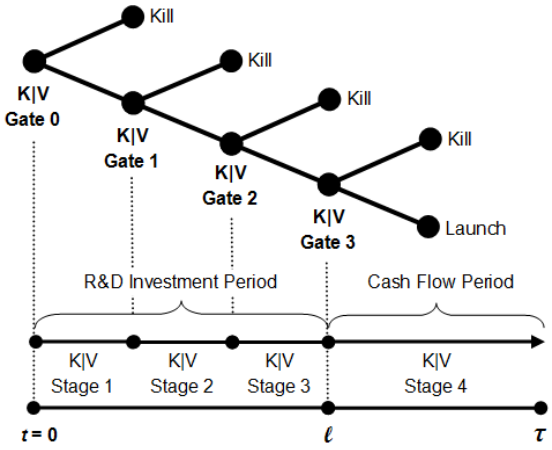


Figure 1: Compound Option Tree and R&D Methodology

Within each stage, four steps spiral (Boehm 1988) to capture learning through iteration. The activities of each time-boxed spiral are organized into a four step plan-benchmark-do-check (PBDC[2]) framework.

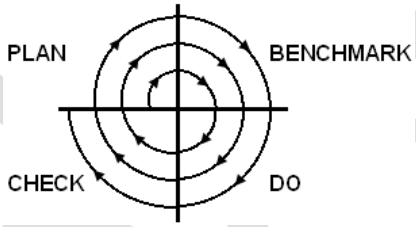


Figure 2. Plan-benchmark-do-check[3]

- **Plan.** What do we need to do?
- **Benchmark.** What's the best way to do it
- **Do.** We do it.
- **Check.** How did we do?

At the completion of each stage is a gate that gives top management the optionality. They can kill the project or continue to the next stage of development. Well-

[2] K|V's PBDC framework in Figure 3 differs from the traditional Six Sigma plan-do-check-act (PDCA) methodology from quality due to the heavy research component in quantitative finance.
[3] To simplify the model, each spiral is condensed to three loops as in Figure 7. In practice they may consist of more or fewer as needed. Each loop consists of one pass over each of the steps in the stage spiral.

organized gate meetings will each have a unique set of metrics and criteria for exercising the option on the next stage and funding further development. Gates minimize the probability of investing more R&D capital on unsuccessful trading system projects.

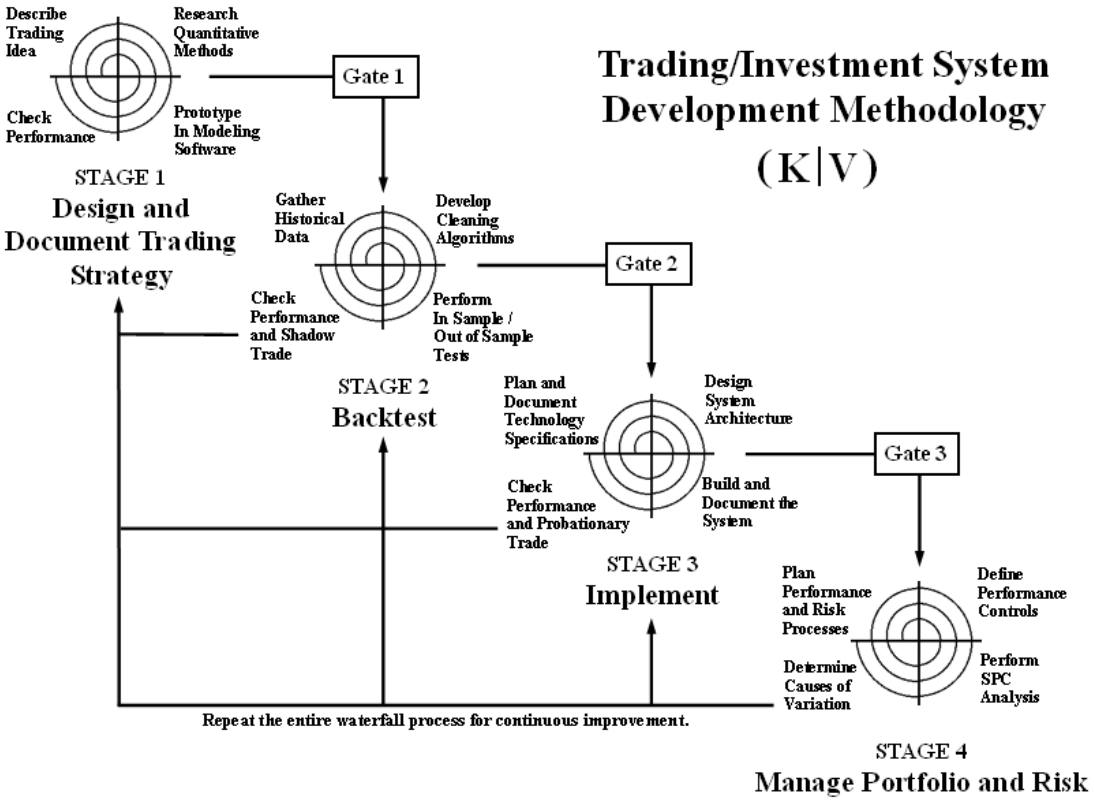


Figure 3. K|V Development Methodology

III. REVENUES

The payoff of a trading system is the stream of daily cash flows generated by the running technological implementation. The future performance of such a system can estimated by way of a stable reference distribution of returns established in a backtest, or in simulated and/or probationary trading. Time τ occurs when the performance of the working trading system violates its reference distribution (see Bilson, et al. 2012). As

with later parameter estimates which use the triangular distribution[4], trading groups can arrive at consensus min-max-most-likely parameter estimates for τ. Estimating τ takes into account that complex systems may have a longer life than simple ones.

The estimated distribution of returns of the trading system may be non-normal (but stable) (see Bilson et al. (2010), Cooper and Van Vliet (2012)), with potentially long tails. Myriad families of distributions able to fit combinations of skewness and kurtosis exist—Johnson, Pearson, beta. For ease of calculation and use in simulation, we will use the four-parameter generalized lambda distribution (henceforth GLD) (see Cooper and Van Vliet (2012) for the use of this distribution to model non-normal automated trading system returns). Given $GLD(\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, the average profit per time period is:

$$\bar{\pi}_d = \lambda_1 + \frac{A}{\lambda_2} \qquad \text{where,} \quad A = \frac{1}{1+\lambda_3} - \frac{1}{1+\lambda_4} \tag{2}$$

(See Appendix 1 for GLD specifications.)

The distribution of revenues is estimated at $t = 0$, and then re-estimated at each gate. These estimates evolve from opinion to empirically validated distributions as R&D progresses.

IV. COSTS

While an automated trading system may demand a relatively small amount of trading capital, it (more than likely) will require a large investment in technological infrastructure—software, servers, routers, and IT personnel. Thus, profitable trading does not guarantee positive cash flow after expenses. To be capable, an automated trading system must at a minimum cover its own costs (Kumiega and Van Vliet 2013).

[4] The triangular distribution provides for intuitive parameter estimation and well as straight-forward computation. The triangular distribution is regularly used in project management since it allows a team to estimate using intuitive logic of min-max-most-likely parameters. See Klastorin (2004) for the use of this distribution in project management.

For all projects, Stage 1 time and costs are incurred. Stage 2 time and costs are only incurred on the chance that Stage 1 proves successful in the form of a well-defined trading strategy. Likewise, Stage 3 time and costs are only incurred if Stage 2 backtesting proves a stable and capable return distribution. And, the cash flows are realized only if Stage 3 technological implementation and probationary trading proves successful.

At the outset, of course, these costs and times are fuzzy. To estimate the times to complete each stage and the associated costs, we use the triangular distribution plus a probability of success. For each stage, traders can arrive at minimum *a*, most likely *m*, and maximum *b* estimates for both the time and cost to complete each stage using (say) a Delphi approach[5]. As R&D progresses, better and better triangular estimates can be made of the future costs and completion times. For K|V Stages 1, 2, and 3, the parameters for costs c_i are:

$$c_i \sim Triangular(a_{c,i}, m_{c,i}, b_{c,i}) \tag{3}$$

So that the expected cost for each stage is $\bar{c}_i$. Figure 4 shows that the expected cost of each stage increases. Prototypes are the least expensive to design, then more to build and implement in production. At each gate, management commits more capital to the project.

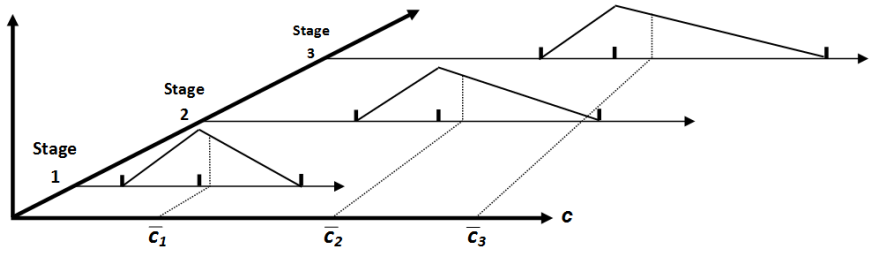


Figure 4. Expected Costs for Each Stage Increase

[5] This is intuitive for traders who often use range-based estimators of risk.

Likewise, the estimated time to completion t_i are of each stage is:

$$t_i \sim Triangular(a_{t,i}, m_{t,i}, b_{t,i}) \tag{4}$$

So, the expected time to completion of each stage is $\bar{t}_i$. Figure 5 shows that the time to complete each stage increases.

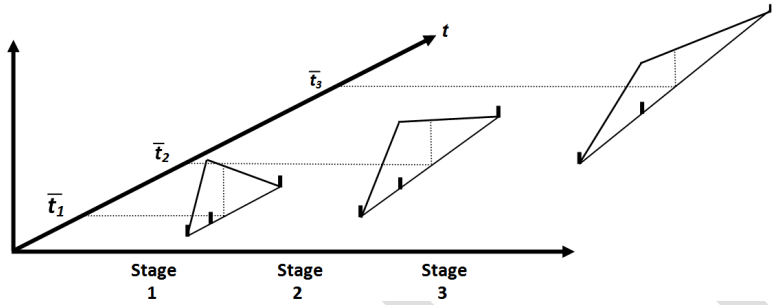


Figure 5. Expected Times to Completion of Each Stage Increase

Also, there is a probability that the time and effort will be wasted on a project that turns out to have no chance of success. Thus, for each stage, we estimate the probability of success p_i. This is the probability the option on the next stage will be bought.

Stage 1: Design and Document Trading/Investment Strategy. Stage 1 is the quantitative research step—constructing testable, "well-defined" strategies for Stage 2. The goal of this stage is to find trading strategies potentially capable of meeting risk-return specifications.

The costs associated with Stage 1 are people, software, office space, and basic data. The present value of the future costs for Stage 1 at $t = 0$ is:

$$PVCosts_1 = \bar{c}_1 \cdot \exp(-r_f \cdot \bar{t}_1) \tag{5}$$

Gate 1. Gate meeting attendees evaluate each project according to rigorous criteria for justifying continuation. The gate meeting is the expiration of the first project

option. Continuation of the project beyond this gate is contingent upon finding the expected net present value to be higher than the termination value.

In order to receive additional funding for continued R&D in Stage 2, management must revalue the next option, which requires re-estimation of revenues and τ, times and costs to completion for the remaining stages, and the probabilities of success. These estimates will be made with more clarity and understanding than those estimates at $t = 0$.

Stage 2: Backtest. Stage 2 is the empirical validation phase—testing strategies defined in Stage 1. A backtest is a simulation of a trading strategy using historical data. The goal of this stage is to refine the reference distribution of expected returns. Costs associated with Stage 2 are people, data, backtesting software, office space, servers. Stage 2 costs are incurred only for those projects that pass Gate 1. That is, the Stage 2 costs are multiplied by the probability that Stage 1 was completed successfully. The present value of the future costs for Stage 2 at $t = 0$ is:

$$PVCosts_2 = \bar{c}_2 \cdot \exp(-r_f \cdot (\bar{t}_1 + \bar{t}_2)) \cdot p_1 \tag{6}$$

Gate 2. As with Gate 1, continuation of the project beyond this gate is contingent upon finding the expected net present value to be higher than the termination value. In order to receive additional funding for continued R&D in Stage 3, management must value the next project option. The parameter estimates for the remaining R&D stage will be made now with greater clarity. This is the key gate because the estimate of profitability distribution goes from consensus opinion to an empirically validated distribution through backtesting.

Stage 3: Implement. Stage 3 is the technology development phase. The goal of this stage is to build a technology infrastructure that correctly implements the trading

strategy. Implementation requires connectivity between and interoperability with disparate software and hardware systems for trade execution, a software design to encapsulate trading logic and order management and routing, and other processes such as optimization and data storage.

Costs associated with Stage 3 are people, programmers, leased lines, real time servers, software development and testing costs, office space, colocation costs. Stage 3 costs are incurred only for those projects that pass Gate 2. That is, the Stage 3 costs are multiplied by the probabilities that both Stage 1 and Stage 2 were completed successfully. The present value of the future costs for Stage 3 at $t = 0$ is:

$$PVCosts_3 = \bar{c}_3 \cdot \exp(-r_f \cdot (\bar{t}_1 + \bar{t}_2 + \bar{t}_3)) \cdot p_1 \cdot p_2 \tag{7}$$

Gate 3. As with Gate 2, continuation of the project beyond this gate is contingent upon finding the expected net present value to be higher than the termination value. Since the launch costs are all sunk by Gate 3, this calculation should reduce to the net present value. This sunk cost is assumed to be at fair value for the option, not at depreciated cost. Approval at this gate permits full trading of the system.

Stage 4: Manage Portfolio and Risk. Stage 4 is the portfolio and risk management stage. The goal of this stage is to monitor the performance outputs of the trading system relative to the backtested reference distribution of returns to ensure the system is performing to specification. Then at $t = 0$, the total present value of the costs over the four K|V stages is:

$$PVCosts_0 = \sum_{i=1}^{3} PVCosts_i \tag{8}$$

V. REAL OPTION VALUATION

Since the return of each trading day is assumed to be independent (there is no serial correlation), simulation results collapse to a simple expected value scenario (unlike DM). Thus, for option valuation, we can bypass simulation and proceed directly to a closed form solution by estimating the expected present value at ℓ directly. The present value of expected future cash flows at ℓ is:

$$PVCF_{\ell} = \bar{\pi}_d \cdot \bar{\tau} \cdot \exp(-r_d \cdot \bar{\tau}) \tag{9}$$

And, the present value of the expected future cash flows is multiplied by the probabilities that Stage 1, Stage 2 and Stage 3 were completed successfully:

$$PVCF_0 = PVCF_{\ell} \cdot \exp(-r_d \cdot (\bar{t}_1 + \bar{t}_2 + \bar{t}_3)) \cdot p_1 \cdot p_2 \cdot p_3 \tag{10}$$

Then, the option value C at $t = 0$, is:

$$C_0 = \max(PVCF_0 - PVCosts_0, 0) \tag{11}$$

Using this framework, the value of the project can be reassessed at each successive gate with new (and better) estimates of times, costs, probabilities of success, and expected returns. In this model risk reduction comes by way of the gate-based development costs. Risk is reduced because environmental regimes—economic climate, contract dry-ups, technology advances—may change during the R&D process that affect the profit opportunity.

We can gain intuition about R&D project risk using simulation. To seed the simulation, we use the parameter values shown in Table 1.

	Time in Days				**Cost in Dollars**				**Probability of Success**
	Min (a)	Mode (m)	Max (b)		Min (a)	Mode (m)	Max (b)		
Stage 1	5	7	10		5,000	10,000	15,000		**.3**
Stage 2	5	10	15		50,000	75,000	125,000		**.5**
Stage 3	10	15	25		75,000	100,000	200,000		**.9**

Table 1. Sample Estimates for Stage Times and Costs

To generate simulated daily profits, we draw variates from *GLD*(.003, 3.250, .006, .002) and multiply them by the (static) daily investment capital of $1,000,000[6]. Figure 6 shows the histogram of profits per day. For random time till τ, we use a minimum of two months (42 trading days), most likely of one year (252 days) and maximum of two years (504 days) (i.e. *Triangular*(42, 252, 504), a discount rate of .05 / 252[7], and a risk-free rate of .025 / 252.

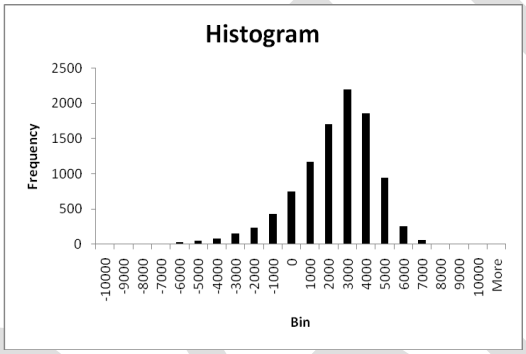
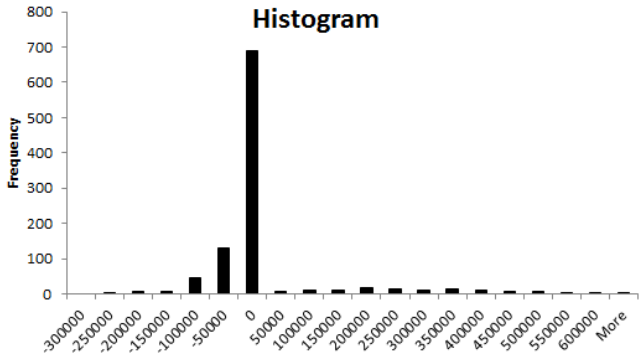


Figure 6. Histogram of Profits per Day

Using this data, the value of the call option on the R&D project is $6,556. (See Appendix 2 for replication.) Figure 7, however, shows the histogram of outcomes as $PVCF_0 - PVCosts_0$ for 1000 trials.



[6] Many trading strategies have a maximum level of scalability. Profits thus cannot be reinvested back into the strategy.
[7] 252 is the estimated number of trading days per year.

Figure 7. Histogram of Simulated $PVCF_0 - PVCosts_0$ at $t = 0$

As many of the trials were killed after K|V Stage 1, the result is a large number of small losses in the 0 to -$50,000 range. For trials that *were* killed, the average loss was $29,564. Fewer projects were killed at later gates, resulting in greater losses. In the left tail, the 1% loss (i.e. $Q(.01)$) was $197,565, and the 5% loss (i.e. $Q(.05)$) was $107,244. The largest single loss was $278,808. While only a few trials made it through to launch at K|V Gate 3, over the entire simulation the expected cash flows exceed the expected losses. Thus, strategy cycling works when R&D projects on losing trading ideas are terminated quickly.

VI. CONCLUSION

The real option method presented uses stage-based, real options approach to reduce R&D project risk and enable strategy cycling. Our model enables initial parameterization of automated trading system R&D projects.

In the finance literature, portfolio management refers to the allocation of resources between different financial instruments—usually stocks and bonds. The real telos is no longer a portfolio of financial assets, but rather a portfolio of intellectual ones, of trading systems where diversification comes not by asset class, but by algorithms, geography, instrument type, and holding periods. But, trading system R&D demands significant investment of time and money. The portfolio of assets to be optimized is now a portfolio of real options on trading system R&D projects, where each project has phases of research, development, launch and operation until failure.

Finance has moved towards new product development, where portfolio management means the allocation and prioritization of R&D projects. The goal of asset

allocation in a portfolio is still the best possible expected return/risk profile, but under the new scenario the portfolio manager must allocate resources among intellectual assets, prioritizing between existing trading strategies as well as the R&D of new ones.

REFERENCES

Adner, Ron. 2007. "Real Options and Resource Reallocation Processes," in Jeffrey J. Reuer, Tony W. Tong (eds.), *Real Options Theory (Advances in Strategic Management, Volume 24)*, Emerald Group Publishing Limited. pp.363-372.

Bilson, John F.O., Andrew Kumiega, and Ben Van Vliet. 2010. "Trading Model Uncertainty and Statistical Process Control." *Journal of Trading*. 5(3), pp. 39-50.

Boehm, Barry W. 1988. "A Spiral Model of Software Development and Enhancement." *Computer*. 21(5), pp. 61-72.

Cooper, Robert G. 2001. *Winning at New Products*. Basic Books. Cambridge, MA.

Cooper, Ricky and Ben Van Vliet. 2012. "Whole Distribution Statistical Process Control
in High Frequency Trading." *Journal of Trading*. 7(2), pp. 57-68.

Gunther McGrath, Rita, Walter J. Ferrier and Aubrey L. Mendelow. 2004. "Response: Real Options as Engines of Choice and Heterogeneity." *Academy of Management Review*. 29(1), pp. 86-101.

Hackett, Sean, and David Dilts. 2004. "A Real Options-Driven Theory of Business Incubation."
Journal of Technology Transfer. 29, pp. 41–54.

Klastorin, T. 2004. *Project Management: Tools and Trade-offs*. New York: John Wiley & Sons.

Kumiega, Andrew and Ben Van Vliet. 2008. *Quality Money Management*. Academic Press Elsevier. London.

Lint, Onno, and Enrico Pennings. 2001. "An option approach to the new product development process: a case study at Philips Electronics." *R&D Management*. 31(2), pp. 1467-9310.

Mathews, Scott, Vinay Datar, and Blake Johnson. 2007. "A Practical Method for Valuing Real Options: The Boeing Approach." *Journal of Applied Corporate Finance*. 19(2) pp. 1745-6622

Öztürk, A., & R. Dale. "Least squares estimation of the parameters of the Generalized Lambda Distribution." *Technometrics*. 27, pp. 81–84.

Ramberg, J. and Schmeiser, B. 1974. "An approximate method for generating asymmetric random variables." *Communications of the ACM*. 17(2), pp. 78-82.

Royce, Winston W. 1970. "Managing the Development of Large Software Systems." *Proceedings of IEEE WESCON*. pp. 1-9.

Tukey, J.W. 1960. "The Practical Relationship Between the Common Transformations of Percentages of Counts and of Amounts." *Technical Report 36*. Statistical Techniques Research Group, Princeton University.

APPENDIX 1

The four parameter Ramberg and Schmeiser (1974) GLD specification is an extension of Tukey's (1960) lambda distribution. It is most often represented by its inverse cumulative distribution, or quantile function:

$$Q(p) = \lambda_1 + \frac{p^{\lambda_3} - (1-p)^{\lambda_4}}{\lambda_2} \tag{14}$$

This leads to the derivation of its density function as:

$$f(x) = \frac{\lambda_2}{\lambda_3 \cdot p^{\lambda_3 - 1} + \lambda_4 (1-p)^{\lambda_4 - 1}} \qquad \text{at } x = Q(p). \tag{15}$$

Where λ_1 is a location parameter, λ_2 the scale parameter, and λ_3 and λ_4 determine the shape. For parameter estimation (if needed) we use the least squares method developed by Ozturk and Dale (1985). Goodness of fit can be checked using some variant of the Kolmogorov-Smirnov test.

The first two moments, mean and variance, are given by:

$$\mu = E(X) = \lambda_1 + \frac{A}{\lambda_2}$$

$$\sigma^2 = E\left[(X-\mu)^2\right] = \frac{B - A}{\lambda_2^2}$$

Where:

$$A = \frac{1}{1+\lambda_3} - \frac{1}{1+\lambda_4}$$

$$B = \frac{1}{1+2\lambda_3} + \frac{1}{1+2\lambda_4} - 2\beta(1+\lambda_3, 1+\lambda_4)$$

And β denotes the Beta function.

APPENDIX 2

	A	B
1	**Cash Flow Parameters**	
2	**Investment**	$ 1,000,000
3	**Lambda 1**	0.003
4	**Lambda 2**	3.250
5	**Lambda 3**	0.006
6	**Lambda 4**	0.002
7	**Time Till Failure (τ) Distribution**	
8	**Min (a)**	42
9	**Mode (m)**	252
10	**Max (b)**	504
11	**Daily Interest Rates**	
12	**Discount Rate**	0.0002
13	**Risk Free Rate**	0.0001

	D	E	F	G	H	I	J	K	L	M
1		**Time in Days**				**Cost in Dollars**				**Probability of Success**
2		Min	Mode	Max		Min	Mode	Max		
3		(a)	(m)	(b)		(a)	(m)	(b)		
4	**Stage 1**	5	7	10		5,000	10,000	15,000		0.3
5	**Stage 2**	5	10	15		50,000	75,000	125,000		0.5
6	**Stage 3**	10	15	25		75,000	100,000	200,000		0.9

	O	P	
1	**Expectations**		
2			
3	**Cash Flows**		
4	**E(τ)**	266	=FLOOR((B8+B9+B10)/3,1)
5	**E(π_d)**	1,779	=B2*(B3+(1/(1+B5)-1/(1+B6))/B4)
6	**PV E(CF) at l**	448,890	=P4*P5*EXP(-B12*P4)
7	**Expected Stage Costs**		
8	**E(c_1)**	10,000	=(I4+J4+K4)/3
9	**E(c_2)**	83,333	=(I5+J5+K5)/3
10	**E(c_3)**	125,000	=(I6+J6+K6)/3
11	**Expected Stage Times**		
12	**E(t_1)**	7	=(E4+F4+G4)/3
13	**E(t_2)**	10	=(E5+F5+G5)/3
14	**E(t_3)**	17	=(E6+F6+G6)/3
15	**Discounted Expected Costs**		
16	**E($costs_1$)**	9,993	=P8*EXP(-B13*P12)
17	**E($costs_2$)**	24,957	=P9*EXP(-B13*(P13+P12))*M4
18	**E($costs_3$)**	18,687	=P10*EXP(-B13*(P14+P13+P12))*M4*M5
19	**Option Value at t = 0**		
20	**PV E(CF)**	60,193	=P6*EXP(-(P12+P13+P14))*M4*M5*M6
21	**PV E(Costs)**	53,637	=SUM(P16:P18)
22	**Option Value**	**6,556**	=MAX(P20-P21,0)