The Long Bond

Ian Martin (Stanford)

Steve Ross (MIT)

August, 2013

Outline

- A simple framework for thinking about the properties of the long end of the yield curve
- Connects the “Recovery Theorem” of Ross (2013) to earlier work of Backus, Gregory and Zin (1989), Kazemi (1992), Bansal and Lehmann (1997), Alvarez and Jermann (2005) and Hansen and Scheinkman (2009)

Outline

- What do we learn by observing the (infinitely) long yield?
- What do we learn by observing the realized returns on the (infinitely) long bond?
- What shape is the yield curve, on average?
- What information can we learn from interest-rate options?
- What is the expected return on the long bond?

The framework

1. The fixed-income market is complete
    - The fixed-income derivatives market is the most well-developed of all derivatives markets
2. State variable relevant for interest rates follows a Markov chain
    - For example, the state variable might be a vector:
    (short rate, yield curve spread, yield curve curvature, . . . ,
    VIX index, "animal spirits", state of the business cycle, . . .)
    - We will not take a stance on the details of the state variable
    - But it *is* important that all elements of the vector can reasonably be thought of as stationary

The framework

- The system moves randomly from state to state; we do not make any assumptions about how the system moves around
- Instead, we can infer it from the yield curve and other interest-rate derivatives prices
- There are objective probabilities

$$\mathbf{\Pi} = \begin{pmatrix} \pi(1,1) & \pi(1,2) & \cdots & \pi(1,m) \\ \pi(2,1) & \pi(2,2) & & \vdots \\ \vdots & & \ddots & \vdots \\ \pi(m,1) & \cdots & \cdots & \pi(m,m) \end{pmatrix}$$

- $\pi(i,j)$ is the probability of moving from state i to state j

The framework

- We summarize fixed-income asset prices in a matrix, $\boldsymbol{A}$, of Arrow–Debreu prices

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}(1,1) & \boldsymbol{A}(1,2) & \cdots & \boldsymbol{A}(1,m) \\ \boldsymbol{A}(2,1) & \boldsymbol{A}(2,2) & & \vdots \\ \vdots & & \ddots & \vdots \\ \boldsymbol{A}(m,1) & \cdots & \cdots & \boldsymbol{A}(m,m) \end{pmatrix}$$

- $\boldsymbol{A}(i,j)$ is the price, in state i, of the Arrow–Debreu security that pays off \$1 if state j materializes next period
- No arbitrage requires that the entries of $\boldsymbol{A}$ are nonnegative

The framework

- We think of the matrix $\boldsymbol{A}$ as our *data*
- All the information about fixed-income markets that we can conceivably extract from asset prices is contained in $\boldsymbol{A}$
- For example, if the matrix of Arrow–Debreu prices is

$$\boldsymbol{A} = \begin{pmatrix} 0.44 & 0.20 & 0.30 \\ 0.50 & 0.30 & 0.16 \\ 0.10 & 0.10 & 0.78 \end{pmatrix}$$

then the price of a bond in state 2 is 0.96, so the interest rate in state 2 is about 4%

Recovery

- Assume (for now) that $\boldsymbol{A}$ is directly observable
- What can we learn from $\boldsymbol{A}$?
- In particular, can we learn the transition probabilities?
- In the finance lingo, asset prices tell us $\mathbb{Q}$; but can we learn $\mathbb{P}$?

Recovery

- If there were a utility-maximizing investor with time discount factor ϕ, marginal utility $u'(i)$ in state i, and transition probabilities $\pi(i,j)$, then the Arrow–Debreu prices would satisfy

$$\boldsymbol{A}(i,j) = \phi\pi(i,j)\frac{u'(j)}{u'(i)}$$

- In terms of inverse marginal utility, $v(i) \equiv 1/u'(i)$,

$$\boldsymbol{A}(i,j) = \phi v(i)\pi(i,j)/v(j)$$

Recovery

- This can be written concisely as the matrix equation

$$\boldsymbol{A} = \phi \boldsymbol{D} \boldsymbol{\Pi} \boldsymbol{D}^{-1}$$

- $\boldsymbol{D}$ is a diagonal matrix with positive diagonal elements $\{v(i)\}$
- $\boldsymbol{\Pi} = \{\pi(i,j)\}$ is a *stochastic* matrix, i.e. a matrix whose row sums all equal 1 (because they sum over the probabilities of moving from the current state to any other state)
- We will not need the existence of such a utility-maximizing agent; instead we will use…

The pseudo-representative agent

Result (The pseudo-representative agent)

Given the asset price data $\boldsymbol{A}$*, there exists a unique decomposition*

$$\boldsymbol{A} = \phi \boldsymbol{D} \boldsymbol{\Pi} \boldsymbol{D}^{-1}$$

where $\boldsymbol{D}$ *is a diagonal matrix and* $\boldsymbol{\Pi}$ *is a transition matrix.*

This can be interpreted as establishing existence and uniqueness of a "pseudo-representative agent" (ϕ *and* $\boldsymbol{D}$*) and probability measure (*$\boldsymbol{\Pi}$*) that rationalize the observable asset prices (*$\boldsymbol{A}$*).*

The pseudo-representative agent

Proof.

The Perron–Frobenius theorem guarantees that $\boldsymbol{A}$ has a unique eigenvector $\boldsymbol{v}$ and associated eigenvalue ϕ—so that $\boldsymbol{A}\boldsymbol{v} = \phi\boldsymbol{v}$—with the properties that (i) the entries of $\boldsymbol{v}$ are positive; and (ii) ϕ is real and positive (and the largest, in absolute value, of all the eigenvalues of $\boldsymbol{A}$). Let $\boldsymbol{D}$ be the diagonal matrix with $\boldsymbol{v}$ along its diagonal, so that $\boldsymbol{v} = \boldsymbol{D}\boldsymbol{e}$, where $\boldsymbol{e}$ is the vector of ones. Then, $\frac{1}{\phi}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}\boldsymbol{e} = \frac{1}{\phi}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{v} = \boldsymbol{D}^{-1}\boldsymbol{v} = \boldsymbol{e}$. This shows that $\frac{1}{\phi}\boldsymbol{D}^{-1}\boldsymbol{A}\boldsymbol{D}$ is a stochastic matrix, which we will call $\boldsymbol{\Pi}$, so that $\boldsymbol{A} = \phi\boldsymbol{D}\boldsymbol{\Pi}\boldsymbol{D}^{-1}$. (Uniqueness follows easily.) $\square$

A hypothesis

$$\boldsymbol{A} = \phi \boldsymbol{D}\boldsymbol{\Pi}\boldsymbol{D}^{-1}$$

- In general, no guarantee that $\boldsymbol{\Pi}$ is the true probability distribution
- Maintained hypothesis: $\boldsymbol{\Pi}$ *is* the true probability distribution
- A more general hypothesis than the assumption that fixed-income assets are priced (as if) by a utility-maximizing investor
- Ultimately an empirical question: to settle it, must construct $\boldsymbol{A}$. We skip this (hard) problem and simply develop some implications of the hypothesis, if true
- Results that are conditional on the hypothesis are in pink

Recovery

- In our earlier example, the matrix of Arrow–Debreu prices

$$\boldsymbol{A} = \begin{pmatrix} 0.44 & 0.20 & 0.30 \\ 0.50 & 0.30 & 0.16 \\ 0.10 & 0.10 & 0.78 \end{pmatrix}$$

can be uniquely decomposed as

$$\boldsymbol{A} = \underbrace{0.97}_{\phi} \underbrace{\begin{pmatrix} 0.56 & & \\ & 0.57 & \\ & & 0.61 \end{pmatrix}}_{\boldsymbol{v} \text{ along the diagonal}} \underbrace{\begin{pmatrix} 0.45 & 0.21 & 0.34 \\ 0.51 & 0.31 & 0.18 \\ 0.10 & 0.10 & 0.80 \end{pmatrix}}_{\text{transition probabilities, } \boldsymbol{\Pi}} \begin{pmatrix} 0.56 & & \\ & 0.57 & \\ & & 0.61 \end{pmatrix}^{-1}$$

Recovery

- The eigenvector $\boldsymbol{v}$ represents the vector of (inverse) marginal utilities across states
- The eigenvalue ϕ represents the pure time discount factor
- We can use these, together with $\boldsymbol{A}$, to calculate the transition probability matrix $\boldsymbol{\Pi}$
- But it may be hard, in practice, to calculate the original matrix $\boldsymbol{A}$
- We need to observe the yield curve and prices of a rich range of interest-rate derivatives—swaptions, caps, floors, ...

Back to the yield curve

- The price, in state 1, of a two-year zero-coupon bond can be calculated by squaring $\boldsymbol{A}$...

$$\begin{pmatrix} 0.44 & 0.20 & 0.30 \\ 0.50 & 0.30 & 0.16 \\ 0.10 & 0.10 & 0.78 \end{pmatrix}^2 = \begin{pmatrix} 0.32 & 0.18 & 0.40 \\ 0.39 & 0.21 & 0.32 \\ 0.17 & 0.13 & 0.65 \end{pmatrix}$$

  ...and adding up the entries of the first row, giving 0.90, so the annualized two-year yield is about 5% in state 1

- The n-year yield can be calculated by raising the matrix $\boldsymbol{A}$ to the power of n

The yield curve

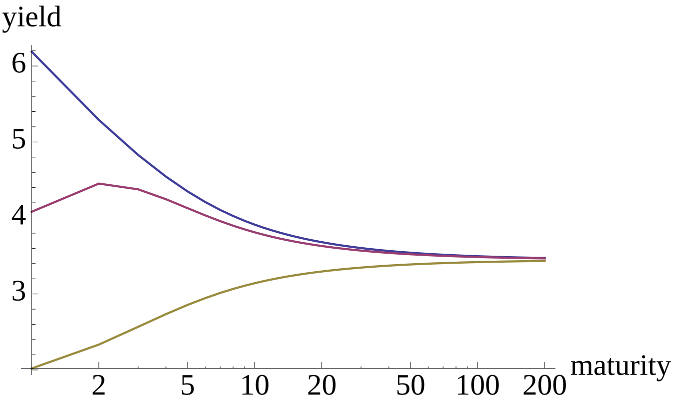

Figure: The yield curve in the three states of the world

The yield curve

Result (The long bond yield reveals ϕ)

The yield and unconditional expected log return of the long bond are determined by ϕ, the largest eigenvalue of $\boldsymbol{A}$: $y_\infty = \bar{r}_\infty = -\log \phi$.

- The long bond: the infinitely long zero-coupon bond
- So, in our setting, the long end of the yield curve directly reveals the time discount factor of the pseudo-representative agent
- No need to compute $\boldsymbol{A}$
- The long yield is constant (cf Dybvig, Ingersoll and Ross 1996)

The yield curve

- We now have a direct link between the eigenvalue and something we can observe in the market
- Want to do the same for the eigenvector, $\boldsymbol{v}$, which summarizes marginal utilities across states
- Idea: interpret $\boldsymbol{v}$ as the payoffs on an asset that pays $v(1)$ in state 1, $v(2)$ in state 2, $\ldots$
- This asset has a nice interpretation

The yield curve

Result (The long bond return reveals $\boldsymbol{v}$)

Returns on the $\mathbf{v}$-asset replicate the returns on the long bond:

$$R_\infty(i,j) = \frac{v(j)}{\phi v(i)} = \frac{\textit{payoff of } \mathbf{v}\textit{-asset in state } j}{\textit{price of } \mathbf{v}\textit{-asset in state } i}$$

- Cross-sectional approach: observe many derivatives prices, back out $\boldsymbol{A}$ and hence $\boldsymbol{v}$
- Time-series approach: observe the long bond over time and hence back out $\boldsymbol{v}$

Links with earlier literature

- These results link the recovery theorem to an earlier literature
- Long yield equals maximum eigenvalue (Backus, Gregory, Zin 1989—first use of the Perron–Frobenius theorem in asset pricing?)
- The inverse of the return on the long bond is a stochastic discount factor (Kazemi 1992—in a diffusion model)
- More recently, Alvarez and Jermann 2005; Hansen and Scheinkman 2009

The matrix A in Ross (2013)

The State Price Transition Matrix, P												
	Sigmas	-5	-4	-3	-2	-1	0	1	2	3	4	5
Sigmas	$S_0 \setminus S_T$	0.368	0.449	0.549	0.67	0.819	1	1.221	1.492	1.822	2.226	2.718
-5	0.368	0.671	0.241	0.053	0.005	0.001	0.001	0.001	0.001	0.001	0.000	0.000
-4	0.449	0.280	0.396	0.245	0.054	0.004	0.000	0.000	0.000	0.000	0.000	0.000
-3	0.549	0.049	0.224	0.394	0.248	0.056	0.004	0.000	0.000	0.000	0.000	0.000
-2	0.67	0.006	0.044	0.218	0.390	0.250	0.057	0.003	0.000	0.000	0.000	0.000
-1	0.819	0.006	0.007	0.041	0.211	0.385	0.249	0.054	0.002	0.000	0.000	0.000
0	1	0.005	0.007	0.018	0.045	0.164	0.478	0.276	0.007	0.000	0.000	0.000
1	1.221	0.001	0.001	0.001	0.004	0.040	0.204	0.382	0.251	0.058	0.005	0.000
2	1.492	0.001	0.001	0.001	0.002	0.006	0.042	0.204	0.373	0.243	0.055	0.004
3	1.822	0.002	0.001	0.001	0.002	0.003	0.006	0.041	0.195	0.361	0.232	0.057
4	2.226	0.001	0.000	0.000	0.001	0.001	0.001	0.003	0.035	0.187	0.347	0.313
5	2.718	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.032	0.181	0.875

- Implies interest rates of 2.5%, 2.1%, 2.5%, 3.3%, 4.6%, 0.0%, 5.4%, 7.0%, 10.4%, 11.8%, -8.4% in states $-5, -4, \ldots, 5$

Implied yield curves in Ross (2013)

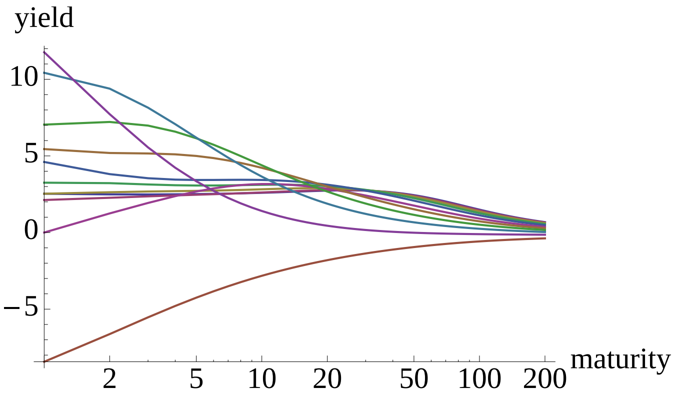

Figure: Implied yield curves

Comparison with Ross (2013)

- If the riskless rate is constant, pricing is risk-neutral
- Confronted with high risk premia in equity markets, the framework is forced to conclude that the riskless rate varies wildly

Long and not-so-long bonds

- The long bond is a theoretically appealing construct. But how closely it can be approximated by bonds of finite maturity?
- Introduce the quantity

$$Q \equiv \log \frac{\max_k v(k)}{\min_k v(k)}$$

- Q indexes the importance of risk for bond pricing
- Maximum and minimum possible log returns on the long bond are

$$\min_{i,j} r_\infty(i,j) = y_\infty - Q \qquad \text{and} \qquad \max_{i,j} r_\infty(i,j) = y_\infty + Q$$

Long and not-so-long bonds

Result

The T-period yield cannot diverge too far from the long yield:

$$|y_T(i) - y_\infty| \le \frac{Q}{T}$$

For any two states i and j, the T_1-period and T_2-period yields satisfy

$$|y_{T_1}(i) - y_{T_2}(j)| \le Q\left(\frac{1}{T_1} + \frac{1}{T_2}\right)$$

- Either yield volatility (across states) or yield curve slope (across maturities) signals high Q, and hence important bond risk premia

A property of the long bond

Result

Within the class of fixed-income assets, the long bond is growth-optimal; that is, in every state the long bond has the highest expected log return of all fixed-income assets.

- Related: Kazemi (1992), Alvarez and Jermann (2005)

The average shape of the yield curve

Result

The yield curve cannot always slope down. Nor can it always slope up. But it must slope up on average.

- This is an empirical regularity
- Implies that $\phi < e^{-\bar{y}_1}$: average short rates provide an upper bound on the subjective time discount factor

How much does the long yield vary?

- It doesn't! This is the wrong question
- Long-dated bonds have tiny yield changes, but enormous duration (i.e., sensitivity to yield changes)

$$r_t(i,j) = \underbrace{y_t(i)}_{\text{yield}} - \underbrace{(t-1)}_{\text{duration}} \times \underbrace{(y_{t-1}(j) - y_t(i))}_{\text{yield change}}$$

- So, to detect small changes in the long end of the yield curve, better to look at expected returns, which *can* vary across states

What is the expected return on the long bond?

Result

Options on the long bond reveal its conditional expected excess return:

$$\overline{R}_\infty(i) - R_f(i) = 2\left\{ \int_0^{R_f(i)} \mathrm{put}(K;i)\,dK + \int_{R_f(i)}^{\infty} \mathrm{call}(K;i)\,dK \right\}$$

where $\mathrm{call}(K;i)$ *is the price, in state i, of a call option with strike K, maturing next period, on the long bond return, and* $\mathrm{put}(K;i)$ *is the corresponding put price.*

What is the expected return on the long bond?

- A simple formula that expresses a forward-looking expected return in terms of option prices
- The long bond earns a nonnegative risk premium in our framework
- For the risk premium to be constant, we would need the volatility index to be constant over time
- Moral: derivatives prices help to reveal the compensation that market participants require for interest-rate risk

What's going on?

- In one respect our assumptions are very weak: we do not even assume that the concept of utility is well-defined
- There may be very rare and unpleasant states of the world
- Nonetheless, interest rates are stationary and bounded
- These properties are not satisfied in other models, which can have yield curves that are downward-sloping on average (eg, Vasicek, Cox–Ingersoll–Ross, Bansal–Yaron, Gollier, . . .)

What's going on?

- A fundamental result of asset-pricing is the relationship

$$\mathbb{E}\left[M_1 R_1 \cdots M_T R_T\right] = 1$$

- Along sample paths, $M_1 R_1 \cdots M_T R_T$ has an interesting property: in the case of a long-dated asset (large T) almost all of the "value" of the asset comes from rare states of the world in which $M_1 R_1 \cdots M_T R_T$ explodes

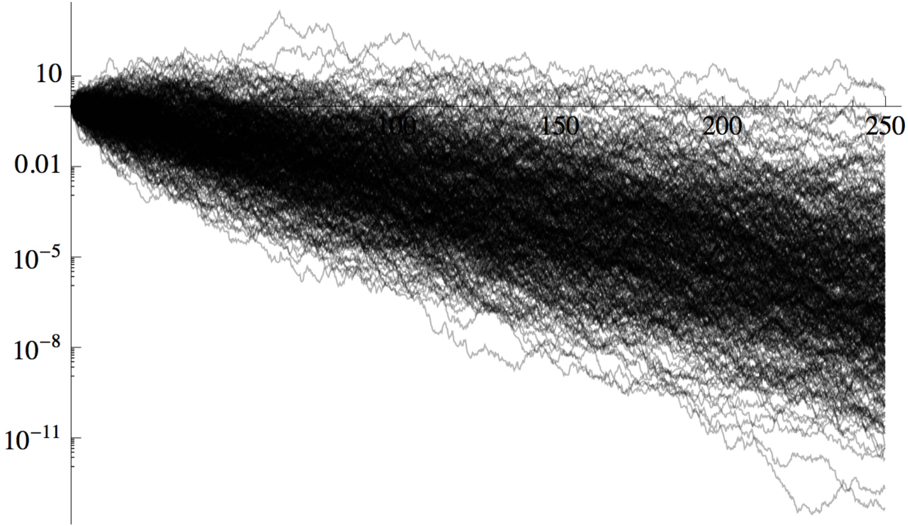

Figure: 400 sample paths of $M_1R_1 \cdots M_TR_T$ over a 250-year horizon

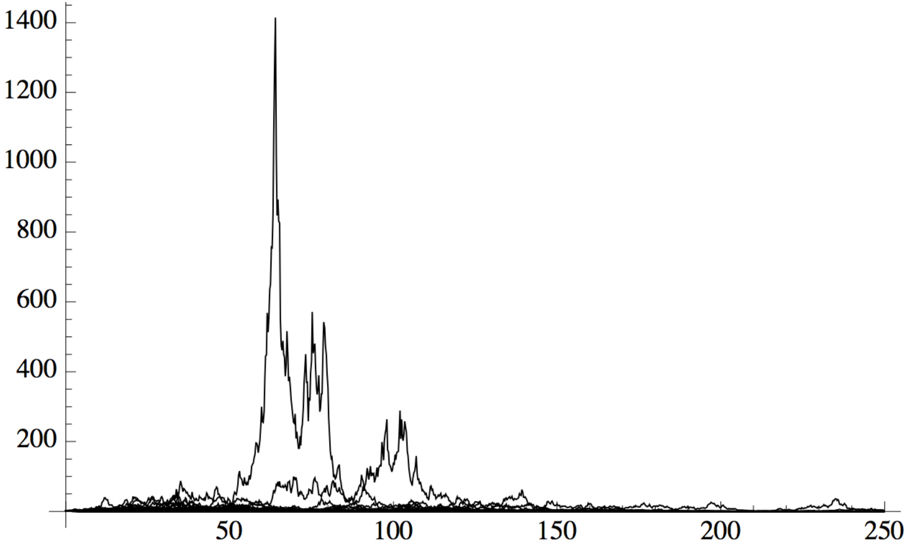

Figure: 400 sample paths of $M_1R_1 \cdots M_TR_T$ over a 250-year horizon

What's going on?

- Whether or not the economy features disasters in the Rietz–Barro sense, *extreme outcomes matter for pricing long-dated assets*
- This is true for all assets except the growth-optimal asset
- But, in our setting, the long bond *is* growth-optimal, so it is the unique fixed-income asset that can escape this conclusion

Multidimensional Breeden–Litzenberger

Extremely preliminary

- To test the hypothesis, we need to construct $\boldsymbol{A}$
- It's not enough to observe the yield curve; have to exploit derivatives prices
- If the state variable is one-dimensional (eg, short rate), use Breeden–Litzenberger
- If state variable is $(y_{1,t}, y_{10,t})'$, Breeden–Litzenberger doesn't work
- Finding a multidimensional version of Breeden–Litzenberger is interesting more generally—but hard

Multidimensional Breeden–Litzenberger

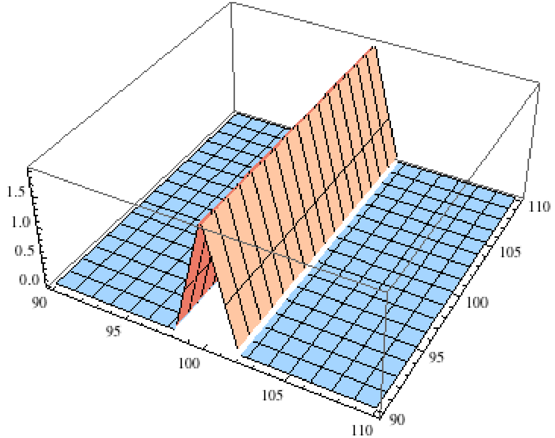

Figure: Butterfly spread using options on 1yr rate

Multidimensional Breeden–Litzenberger

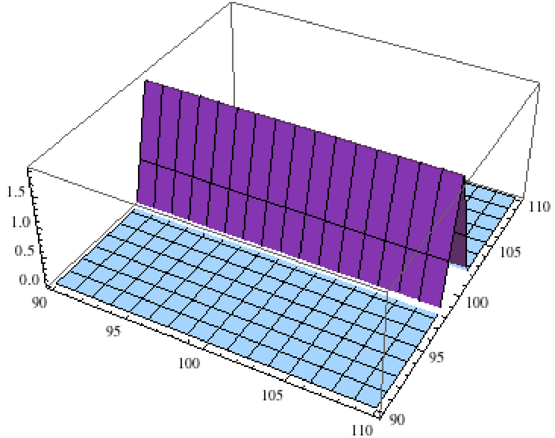

Figure: Butterfly spread using options on 10yr rate

Multidimensional Breeden–Litzenberger

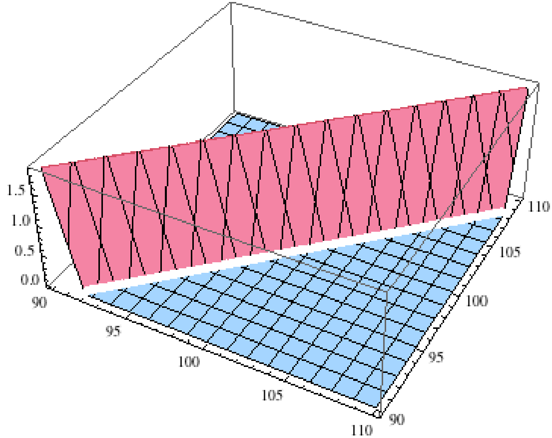

Figure: Butterfly spread using spread options

Multidimensional Breeden–Litzenberger

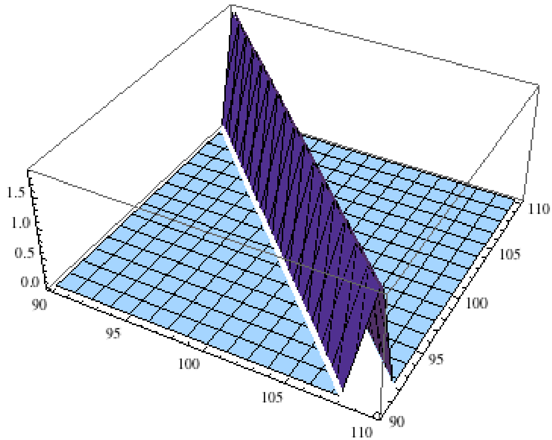

Figure: Butterfly spread using average rate options

Multidimensional Breeden–Litzenberger

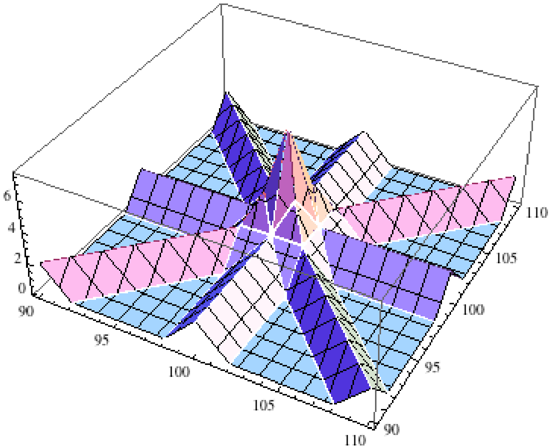

Figure: All together

Multidimensional Breeden–Litzenberger

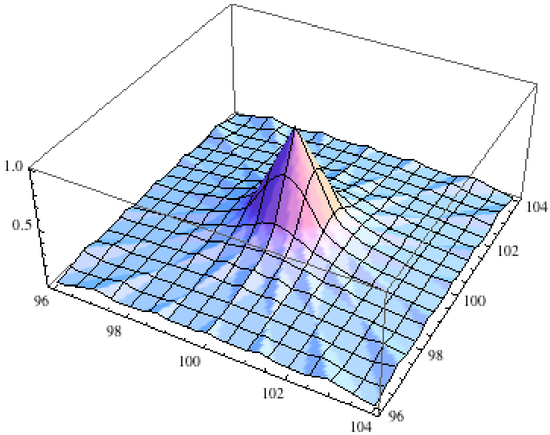

Figure: All together